



HLA alleles, haplotypes frequencies, and their association with hematological disorders: a report from 1550 families whose patients underwent allogeneic bone marrow transplantation in Egypt

Yasser ElNahass^{1,3} · Nourhan Mekky⁴ · Nabil M. Abdelfattah³ · Raafat Abdelfattah^{1,3} · Mohamed Samra^{1,3} · Omar A. Fahmy^{2,3} · Gamal Fathy³ · Wafaa Elmetnawy² · Salwa Sabet⁴ · Heba Bassiouny⁴ · Heba Nader⁵ · Alaa ElHaddad^{1,3} · H.K. Mahmoud^{1,3}

Received: 17 April 2024 / Accepted: 15 June 2024 / Published online: 21 June 2024
© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2024

Abstract

HLA alleles are representative of ethnicities and may play important roles in predisposition to hematological disorders. We analyzed DNA samples for HLA-A, -B, -C, -DRB1, and -DQB1 loci, from 1550 patients and 4450 potential related donors by PCR-SSO (Polymerase chain reaction sequence-specific oligonucleotides) and estimated allele frequencies in donors and patients from 1550 families who underwent bone marrow transplantation (BMT) in Egypt. We also studied the association between HLA allele frequencies and incidence of acute myeloid leukemia, acute lymphoblastic leukemia, and severe aplastic anemia. The most frequently observed HLA class I alleles were HLA- A*01:01 (16.9%), A*02:01 (16.1%), B*41:01 (8.7%), B*49:01 (7.3%), C*06:02 (25.1%), and C*07:01 (25.1%), and the most frequently observed class II alleles were HLA-DRB1*11:01 (11.8%), DRB1*03:01 (11.6%), DQB1*03:01 (27.5%), and DQB1*05:01 (18.9%). The most frequently observed haplotypes were A*33:01~B*14:02~DRB1*01:02 (2.35%) and A*01:01~B*52:01~DRB1*15:01 (2.11%). HLA-DRB1*07:01 was associated with higher AML odds (OR, 1.26; 95% CI, 1.02–1.55; $p=0.030$). Only HLA-B38 antigen showed a trend towards increased odds of ALL (OR, 1.52; 95% CI, 1.00–2.30; $p=0.049$) HLA-A*02:01, -B*14:02, and -DRB1*15:01 were associated with higher odds of SAA (A*02:01: OR, 1.35; 95% CI, 1.07–1.70; $p=0.010$; B*14:02: OR, 1.43; 95% CI, 1.06–1.93; $p=0.020$; DRB1*15:01: OR, 1.32; 95% CI, 1.07–1.64; $p=0.011$). This study provides estimates of HLA allele and haplotype frequencies and their association with hematological disorders in an Egyptian population.

Keywords HLA alleles · Egyptians · Allo-SCT · Hematological disorders

Introduction

Human leukocyte antigens (HLA) are encoded by the major histocompatibility complex (MHC) on the short arm of chromosome 6 (p21). The human MHC loci are classified into Class I (HLA-A, -B, and -C) and Class II (HLA-DR, -DP, and -DQ) (Gfeller and Bassani-Sternberg 2018).

Donor-recipient histocompatibility is critical for the success of allogeneic stem cell transplantation (allo-SCT) (Simpson and Dazzi 2019).

Allo-SCT is the standard of care for many hematological disorders. HLA allele matching plays a key role in the etiology of allogeneic immune responses following allo-SCT, such as graft-versus-host disease (GVHD) (Park and Seo 2012), and significantly impacts the transplant outcome (Petersdorf 2017). The most used source of hematopoietic stem cells is bone marrow (BM). However, peripheral blood stem cell harvesting is the common practice in our BMT unit due to the relatively easier and safer procedure compared to BM harvesting (Mahmoud et al. 2008).

Since there is no donor registry in Egypt, therefore, allo-SCT transplants are only performed from related donors -chiefly siblings and parents. Unfortunately, about 50% of our patients do not find a fully matched HLA donor.

✉ Nabil M. Abdelfattah
nabil.mahmoud20@pharma.asu.edu.eg

¹ National Cancer Institute, Cairo University, Cairo, Egypt
² Faculty of Medicine, Cairo University, Cairo, Egypt
³ Nasser Institute, Ministry of Health, Cairo, Egypt
⁴ Faculty of Science, Cairo University, Cairo, Egypt
⁵ Genome Onco-center, Cairo, Egypt

Haploidentical stem cell transplants (Haplo-SCT) were recently introduced into practice (Chang et al. 2021). However; matched unrelated donor (MUD) transplants have better outcomes and are usually the preferred alternative to Haplo-SCT (Gooptu et al. 2021).

HLA polymorphic regions have functional significance for disease susceptibility due to their association with peptide-binding grooves. Certain HLA antigens and alleles are linked to the pathophysiology and outcomes of hematological disorders (Kazemi et al. 2021) and could be critical factors for disease predisposition. Abnormal HLA expression is associated with the occurrence of several hematological disorders such as lymphomas, acute and chronic leukemias, severe aplastic anemia (SAA), hemoglobinopathies (e.g., B-thalassemia, sickle cell disease), and congenital immunodeficiencies (Villemagne et al. 2005; Huang et al. 2012; Gragert et al. 2014; Zaimoku et al. 2017).

The frequencies of HLA alleles and haplotypes are unique to each ethnic group and their diversity reflects the extent of variation between different populations. The information about the frequencies and distributions of HLA alleles and haplotypes among Egyptians is currently limited (Elshakankiry et al. 2017). In this study, we retrospectively collected HLA typing data performed on HLA- loci A, B, C, DRB1, and DQB1 for patients and donors at the Histocompatibility Laboratory of Nasser Institute for Health and Research (NIH) and estimated HLA allele and haplotype frequencies in Egyptians in an attempt to establish a National stem cell donor registry, which will help transplant candidates to find a compatible stem cell donor. Furthermore, we studied the association between HLA alleles and incidence of the most frequently allografted hematological disorders in our BMT Unit; mainly acute myeloid leukemia (AML), acute lymphoblastic leukemia (ALL), and severe aplastic anemia (SAA) to unravel risky HLA alleles for each hematological disorder.

Patients and methods

Patients and donors' selection

Between 2016 and 2022, we received peripheral blood samples from 1550 allo-SCT candidates with various hematological disorders -including hematological malignancies, bone marrow failure syndromes, primary immunodeficiencies, and hemoglobinopathies (Table 1). We also typed blood samples from 4450 potential donors.

DNA extraction and HLA typing method

DNA was extracted using the QiAamp Blood Mini Kit® (Qiagen, Hilden, Germany) according to the manufacturer's

Table 1 Baseline demographics and characteristics of patients and donors

Characteristic	n/N (%)
Donors — no. (%)	4450 / 6000 (74)
Age — yr	
Median (IQR)	30 (21, 40)
Range	0.06–74
Patients — no. (%)	1550 / 6000 (26)
Hematological condition — no. (%)	
AML, MS	466 / 1550 (30)
SAA	438 / 1550 (28)
ALL	265 / 1550 (17)
BTM	95 / 1550 (6.1)
CML	87 / 1550 (5.6)
MDS, CMML	56 / 1550 (3.6)
Others ^a	143 / 1550 (9.2)
Patients (AML, ALL, SAA) born to consanguineous parents — no. (%)	367 / 1550 (23.4)
HLA-A, -B, and -DRB1 homozygous ^b — no. (%)	47 / 2117 (2.2)
Patients who had at least one donor — no. (%)	832 / 1550 (53.7)
Total number of matched donors ^c — no. (%)	861 / 1550 (55.5)
One locus (5/6) mismatch ^d — no. (%)	62 / 1139 (5.4)
Donor-recipient relation — no. (%)	
Sibling	3874 / 4450 (87)
Mother	170 / 4450 (3.8)
Father	152 / 4450 (3.4)
Cousin	63 / 4450 (1.4)
Daughter	38 / 4450 (0.9)
Son	37 / 4450 (0.8)
Uncle (Paternal)	22 / 4450 (0.5)
Uncle (Maternal)	21 / 4450 (0.5)
Half-Brother	13 / 4450 (0.3)
Half-Sister	11 / 4450 (0.2)
Husband	10 / 4450 (0.2)
Other	39 / 4450 (0.8)

ALL Acute Lymphoblastic Leukemia, AML Acute Myeloid Leukemia, BTM Beta-Thalassemia Major, CML Chronic Myeloid Leukemia, CMML Chronic Myelomonocytic Leukemia, IQR Interquartile range, MDS Myelodysplastic syndrome, MS Myeloid Sarcoma, SAA Severe Aplastic Anemia

^aHeterogeneous group of rare benign and malignant hematological conditions, and inborn errors of metabolism

^bHLA-A, -B, and -DRB1 typing was completed for 1171 donors and 946 patients

^cIn 29 families, there were two matched donors per family

^dHLA-A, -B, and -DRB1 mismatch analysis was completed for 1139 donors and patients only

protocol. HLA-A, -B, -C, -DRB1, and -DQB1 were typed using the polymerase chain reaction sequence-specific oligonucleotide probes (PCR-SSO). The detection method is based on PCR amplification principles (exons 1–4 of HLA-A locus, exons 2–4 of HLA -B locus, exon 2–3 of

HLA-C locus, exon 2 of HLA-DRB1 locus, and exon 2–3 of HLA-DQB1 locus). Amplified DNA was then hybridized with specific oligonucleotide probes immobilized as parallel lines on membrane-based strips. HLA allele polymorphisms were genotyped using INNO-LIPA (FUJIREBIO Europe N.V. Gent, Belgium). The INNO-LIPA kit detects the HLA allele groups listed in IMGT/HLA version reports 3.52. The results were analyzed by using the LiPA interpretation software LIRAS.

Statistical analysis and modeling

For each locus, only fully typed subjects were included in the analysis. Antigen and allele frequencies were estimated by direct counting and were estimated for each locus and the studied subgroups (i.e. HLA-A, -B, -C, -DRB1, and -DQB1 in patients with ALL, AML, SAA, and donors independently). Standard errors were calculated as the square root of $AF(1-AF)/n^a$, where AF is the allele frequency, and n^a is the number of alleles given the locus and the subgroup. Antigen and allele frequencies were calculated using R software (R Core Team 2022) version 4.2.0.

Haplotype frequencies were estimated in subjects with full 4-digit typing of HLA-A, -B, and -DRB1. Haplotype phasing and haplotype frequency were estimated using the expectation–maximization algorithm implemented in Python for Population Genomics (PyPOP) software (Lancaster et al. 2007) version 1.0.0b7, which was installed using the package installer for Python package management system (pip) version 23.2.1 and executed in Python (Van Rossum and Drake 2009) version 3.11.5-. Haplotype frequencies standard errors were calculated as the square root of $HF(1-HF)/2N$, where HF is the haplotype frequency, and N is the number of subjects with complete haplotype data. Haplotype frequencies were estimated in donors and patients and the frequencies were compared using Fisher's exact test.

A univariate multinomial regression analysis was performed to determine the association between HLA antigens, alleles, and the odds of ALL, AML, and SAA, taking donors as the reference control group. All additive, dominant, and recessive genetic models were considered, and all the odds ratios were age- and sex-adjusted. The regression analysis was performed using the 'nnet' package (Venables and Ripley 2002) version 7.3–17 in R software. The alleles included in the regression analysis satisfied the following conditions: (1) A frequency of at least 0.03 in all the patient groups and donors, (2) A count of at least 5 in all of the selected patient subgroups and 5 in donors, (3) For the dominant and recessive models, for each antigen group/allele, there must be at least 5 heterozygous patients and 5 heterozygous donors for the dominant model, 5 homozygous patients and 5 homozygous donors for the recessive model (the latter condition was not applied for

the additive model, as the genotypes are coded and analyzed as numerical variables). P-values were adjusted for false-detection rate (FDR) using Benjamini-Hochberg's procedure. Missing data were omitted from the analysis.

Tests for deviations from the Hardy-Weinberg equilibrium (HWE) were performed using Guo and Thomson's Chi-squared test (Guo and Thompson 1992), and testing for selective neutrality was performed using Slatkin's test (Slatkin 1994). Both tests were performed for each locus in donors using PyPO: Two-locus pairwise linkage disequilibrium (LD) analysis was performed using the same package. Haplotypes were compared between donors and recipients in each disease using Fisher's exact test, and p-values were also adjusted for FDR.

Allele frequencies of Egyptian donors were compared to other populations using the package 'midasHLA' (Hammer and Migdal, 2022) version 1.4.0 in R, which fetches allele frequency data from the Allele frequency net database (AFND, www.allelefrequencies.net). We included only alleles of frequencies > 0.03 and populations with missing frequency data of no more than 5%. We used agglomerative clustering to examine the ancestral relation of the included populations, with correlation as a measure of genetic distance, and Ward's method for clustering. Allele frequencies were transformed using the arcsine square root transformation before clustering. Heatmaps were plotted using the package 'pheatmap' (Kolde 2019) version 1.0.12.

Results

Demographics, antigen, and allele frequencies

Patients represented 26% of the whole cohort (1550/6000). The median age of donors was 30 years, and the median age of patients was 27 years. The most frequent hematological conditions eligible for allogeneic stem cell transplantation (allo-SCT) were AML (30%), SAA (28%), and ALL (17%). Fifty-three percent of patients had at least one fully matched donor per family, and 1.9% had more than one fully matched donor per family. Siblings represented 87% of potential donors while only 7.2% were parents (Table 1).

Patients born to consanguineous parents represented 23.4% of study population. The highest rates of consanguinity were observed among families of patients with Fanconi anemia (83%) and Beta-Thalassemia (67%). MDS patients were the most likely to find a fully matched donor (66.1% match probability), while patients with inherited immunodeficiency syndromes were the least likely (15.4%) (Table 2).

Table 2 Age, gender, consanguinity, and probability of finding a matched donor for each hematological disorder

	ALL	AML	BTM	CML	FA	HL	ID	NHL/CLL	MDS/CMML	MF	SAA	Others	Overall
<i>N</i> (%)	265 (17.1)	466 (30.1)	95 (6.1)	87 (5.6)	24 (1.5)	14 (0.9)	13 (0.8)	34 (2.2)	56 (3.6)	18 (1.2)	438 (28.3)	40 (2.6)	1550 (100)
Age at presentation – yr													
Median (IQR)	26 (21, 33)	30 (23, 40)	5 (3–2, 8.5)	34 (28, 44)	8 (4, 12.5)	24 (20, 28)	1 (0.83, 2)	28 (20, 39)	35 (27, 45)	40 (34, 44)	27 (21, 37)	9 (4, 24)	27 (20, 37)
Range	3–64	2–62	0.3–34	10–60	2–20	3–37	0.2–8	13–55	0.8–56	13–55	0.7–59	0.3–41	0.2–64
Male sex – no. (%)	177 (66.8)	290 (62.2)	58 (61.1)	47 (54)	15 (62.5)	9 (64.3)	9 (69.2)	25 (73.5)	34 (60.7)	13 (72.2)	270 (61.6)	20 (50)	967 (62.4)
Positive consanguinity ^a – no. (%)	41 (15)	68 (15)	64 (67)	13 (15)	20 (83)	3 (21)	7 (54)	6 (18)	10 (18)	5 (28)	108 (25)	22 (55)	367 (23.4)
At least one matched donor ^b – no. (%)	135 (50.9)	265 (56.9)	37 (38.9)	42 (48.3)	11 (45.8)	8 (57.1)	2 (15.4)	19 (55.9)	37 (66.1)	11 (61.1)	246 (56.2)	19 (47.5)	832 (53.7)

ALL Acute Lymphoblastic Leukemia, AML Acute Myeloid Leukemia, BTM Beta-Thalassemia Major, CLL Chronic Lymphocytic Leukemia, CML Chronic Myelogenous Leukemia, CMML Chronic Myelomonocytic Leukemia, FA Fanconi Anemia, HL Hodgkin's Lymphoma, ID Immunodeficiency, MDS Myelodysplastic Syndrome, MF Myelofibrosis, NHL Non-Hodgkin's Lymphoma SAA Severe Aplastic anemia

^aPercentage of patients born to consanguineous parents

^bOne or more matched donors per family

HLA antigen and allele frequencies in donors

For class I loci, we observed 24 unique HLA-A antigens, 37 -B antigens, and 14 -C antigens. The most frequent HLA-A antigens were A2 (23.2%), A1 (18.2%), and A30 (9.28%). For HLA-B, the most frequent were B41 (10.7%), B35 (9.47%), and B14 (7.87%). For HLA-C the most frequent antigens were Cw7 (35.7%), Cw6 (20.7%), and Cw4 (16.6%). For HLA class II loci, we observed 13 unique DR antigens and 5 unique DQ antigens; the most frequent were DR13 (16.9%), DR4 (16.9%), and DR3 (14.2%). For HLA-DQ, the most frequent antigens were DQ3 (36.2%), DQ6 (25.8%), and DQ5 (20.6%) (ST1).

For HLA class I alleles, we observed 52 unique -A alleles, 76 -B alleles, and 24 -C alleles; the most frequent A alleles were A*01:01 (16.9%), A*02:01 (16.1%), and A*03:01 (7.31%); the most frequent B alleles were B*41:01 (8.71%), B*49:01 (7.31%), and B*14:02 (6.81%), and the most frequent HLA-C alleles were C*06:02 (25.1%), C*07:01 (25.1%), and C*04:01 (17.1%). For HLA class II alleles, we observed 37 DRB1 alleles and 14 DQB1 alleles; the most frequent DRB1 alleles were DRB1*11:01 (11.8%), DRB1*03:01 (11.6%), and DRB1*13:01 (11.4%), and the most frequent DQB1 alleles were DQB1*03:01 (27.5%), DQB1*05:01 (18.9%), and DQB1*06:01 (13.8%) (Table 3). Agglomerative hierarchical clustering suggested a distant ancestry between Egyptian donors and the common ancestor of Tunisians and Saudis (Fig. 1).

Hardy-Weinberg equilibrium and selective neutrality testing in donors

Except for HLA-DQ, there was a significant deviation from expected genotype frequencies under HWE in all loci. Additionally, excessive homozygosity was observed in all loci. Slatkin's test showed that there was a significant departure from the infinite allele model in the HLA-B and HLA-DRB1 loci, with negative *F_{nd}* values suggesting balancing selection. Furthermore, the expected homozygosity under HWE was significantly lower than expected under the infinite allele model (Table 4).

HLA antigen and allele frequencies in patients

The most common HLA class I antigens in ALL patients were A2 (22.2%), B41 (10.7%), and Cw7 (37.7%). The most common alleles were A*02:01 (17.5%), B*41:01 (8.3%), C*04:01 (19.2%) and C*07:01 (19.2%). The most common HLA class II antigens were DR13 (20.2%), DQ3 (25%), and DQ6 (25%), while the most common alleles were DRB1*13:01 (13.6%) and DQB1*04:02 (18.9%) (ST2 and Table 5).

Table 3 HLA Allele frequency in donors

Locus	Allele	N	AF ± SE	Locus	Allele	N	AF ± SE	Locus	Allele	n	AF ± SE
HLA-A	A*01:01	485	0.16923 ± 0.00700	HLA-DRB1	B*07:02	60	0.02070 ± 0.00265	HLA-DRB1	C*01:02	5	0.01174 ± 0.00522
	A*02:01	462	0.16120 ± 0.00687		B*58:01	60	0.02070 ± 0.00265		C*02:06	4	0.00939 ± 0.00467
	A*03:01	210	0.07327 ± 0.00487		B*15:03	58	0.02001 ± 0.00260		C*11:01	3	0.00704 ± 0.00405
	A*24:02	150	0.05234 ± 0.00416		B*13:02	53	0.01829 ± 0.00249		C*12:03	3	0.00704 ± 0.00405
	A*33:01	147	0.05129 ± 0.00412		B*57:01	50	0.01725 ± 0.00242		C*16:01	3	0.00704 ± 0.00405
	A*68:02	146	0.05094 ± 0.00411		B*35:08	49	0.01691 ± 0.00239		C*03:01	2	0.00469 ± 0.00331
	A*30:01	141	0.04920 ± 0.00404		B*41:02	41	0.01415 ± 0.00219		C*07:02	2	0.00469 ± 0.00331
	A*23:01	121	0.04222 ± 0.00376		B*14:01	31	0.01070 ± 0.00191		C*15:01	2	0.00469 ± 0.00331
	A*02:05	94	0.03280 ± 0.00333		B*07:05	24	0.00828 ± 0.00168		C*16:02	2	0.00469 ± 0.00331
	A*32:01	94	0.03280 ± 0.00333		B*27:03	24	0.00828 ± 0.00168		C*07:04	1	0.00235 ± 0.00234
	A*30:04	91	0.03175 ± 0.00328		B*39:01	24	0.00828 ± 0.00168		C*12:01	1	0.00235 ± 0.00234
	A*11:01	87	0.03036 ± 0.00320		B*40:01	22	0.00759 ± 0.00161		C*14:02	1	0.00235 ± 0.00234
	A*02:02	76	0.02652 ± 0.00300		B*55:01	22	0.00759 ± 0.00161		C*14:03	1	0.00235 ± 0.00234
	A*29:01	75	0.02617 ± 0.00298		B*37:01	20	0.00690 ± 0.00154		C*15:02	1	0.00235 ± 0.00234
	A*26:01	63	0.02198 ± 0.00274		B*51:08	20	0.00690 ± 0.00154		<i>r^a</i>	426	
	A*69:01	60	0.02094 ± 0.00267		B*15:16	19	0.00656 ± 0.00150		DRB1*11:01	873	0.11842 ± 0.00376
	A*30:02	58	0.02024 ± 0.00263		B*27:07	19	0.00656 ± 0.00150		DRB1*03:01	854	0.11584 ± 0.00373
	A*24:03	46	0.01605 ± 0.00235		B*47:01	19	0.00656 ± 0.00150		DRB1*13:01	840	0.11394 ± 0.00370
	A*31:01	45	0.01570 ± 0.00232		B*15:01	18	0.00621 ± 0.00146		DRB1*07:01	776	0.10526 ± 0.00357
	A*03:02	41	0.01431 ± 0.00222		B*13:01	16	0.00552 ± 0.00138		DRB1*15:01	752	0.10201 ± 0.00353
	A*68:01	30	0.01047 ± 0.00190		B*57:03	15	0.00518 ± 0.00133		DRB1*04:03	509	0.06905 ± 0.00295
	A*66:01	24	0.00837 ± 0.00170		B*39:03	14	0.00483 ± 0.00129		DRB1*01:02	435	0.05901 ± 0.00274
	A*74:01	19	0.00663 ± 0.00152		B*73:01	13	0.00449 ± 0.00124		DRB1*13:03	375	0.05087 ± 0.00256
	A*01:03	12	0.00419 ± 0.00121		B*27:02	10	0.00345 ± 0.00109		DRB1*04:02	369	0.05005 ± 0.00254
	A*36:01	10	0.00349 ± 0.00110		B*51:05	10	0.00345 ± 0.00109		DRB1*10:01	283	0.03839 ± 0.00224
	A*34:02	9	0.00314 ± 0.00105		B*35:03	9	0.00311 ± 0.00103		DRB1*04:05	233	0.03161 ± 0.00204
A*80:01	9	0.00314 ± 0.00105	B*40:02	9	0.00311 ± 0.00103	DRB1*14:01	190	0.02577 ± 0.00185			
A*02:17	8	0.00279 ± 0.00099	B*42:02	7	0.00242 ± 0.00091	DRB1*03:02	167	0.02265 ± 0.00173			
A*24:01	8	0.00279 ± 0.00099	B*07:01	6	0.00207 ± 0.00084	DRB1*08:02	148	0.02008 ± 0.00163			
A*01:02	7	0.00244 ± 0.00092	B*27:05	6	0.00207 ± 0.00084	DRB1*04:04	112	0.01519 ± 0.00142			
A*02:08	4	0.00140 ± 0.00070	B*27:01	5	0.00173 ± 0.00077	DRB1*01:01	74	0.01004 ± 0.00116			
A*23:04	4	0.00140 ± 0.00070	B*40:06	5	0.00173 ± 0.00077	DRB1*11:02	74	0.01004 ± 0.00116			
A*02:06	3	0.00105 ± 0.00060	B*57:02	5	0.00173 ± 0.00077	DRB1*12:01	52	0.00705 ± 0.00097			
A*26:07	3	0.00105 ± 0.00060	B*41:05	4	0.00138 ± 0.00069	DRB1*09:01	51	0.00692 ± 0.00097			
A*28:01	3	0.00105 ± 0.00060	B*51:07	4	0.00138 ± 0.00069	DRB1*16:01	36	0.00488 ± 0.00081			
A*01:06	2	0.00070 ± 0.00049	B*54:01	4	0.00138 ± 0.00069	DRB1*13:05	31	0.00421 ± 0.00075			

Table 3 (continued)

Locus	Allele	N	AF ± SE	Locus	Allele	N	AF ± SE	Locus	Allele	n	AF ± SE
	A*10:01	2	0.00070 ± 0.00049		B*15:02	3	0.00104 ± 0.00060		DRB1*03:07	28	0.00380 ± 0.00072
	A*25:01	2	0.00070 ± 0.00049		B*15:08	3	0.00104 ± 0.00060		DRB1*04:01	24	0.00326 ± 0.00066
	A*74:03	2	0.00070 ± 0.00049		B*44:01	3	0.00104 ± 0.00060		DRB1*11:03	22	0.00298 ± 0.00064
	A*02:03	1	0.00035 ± 0.00035		B*08:04	2	0.00069 ± 0.00049		DRB1*08:01	21	0.00285 ± 0.00062
	A*02:04	1	0.00035 ± 0.00035		B*15:09	2	0.00069 ± 0.00049		DRB1*16:02	12	0.00163 ± 0.00047
	A*02:07	1	0.00035 ± 0.00035		B*29:02	2	0.00069 ± 0.00049		DRB1*14:04	10	0.00136 ± 0.00043
	A*03:08	1	0.00035 ± 0.00035		B*39:06	2	0.00069 ± 0.00049		DRB1*01:03	5	0.00068 ± 0.00030
	A*04:02	1	0.00035 ± 0.00035		B*44:05	2	0.00069 ± 0.00049		DRB1*08:03	3	0.00041 ± 0.00023
	A*11:02	1	0.00035 ± 0.00035		B*47:03	2	0.00069 ± 0.00049		DRB1*14:06	3	0.00041 ± 0.00023
	A*24:06	1	0.00035 ± 0.00035		B*81:01	2	0.00069 ± 0.00049		DRB1*12:02	2	0.00027 ± 0.00019
	A*24:07	1	0.00035 ± 0.00035		B*14:03	1	0.00035 ± 0.00035		DRB1*13:07	2	0.00027 ± 0.00019
	A*29:02	1	0.00035 ± 0.00035		B*14:06	1	0.00035 ± 0.00035		DRB1*15:02	2	0.00027 ± 0.00019
	A*32:09	1	0.00035 ± 0.00035		B*27:08	1	0.00035 ± 0.00035		DRB1*03:04	1	0.00014 ± 0.00014
	A*34:01	1	0.00035 ± 0.00035		B*37:04	1	0.00035 ± 0.00035		DRB1*13:04	1	0.00014 ± 0.00014
	A*66:03	1	0.00035 ± 0.00035		B*46:01	1	0.00035 ± 0.00035		DRB1*14:02	1	0.00014 ± 0.00014
	A*69:03	1	0.00035 ± 0.00035		B*48:01	1	0.00035 ± 0.00035		DRB1*14:09	1	0.00014 ± 0.00014
	r ^a	2866			B*51:02	1	0.00035 ± 0.00035		r ^a	7372	
HLA-B	B*41:01	252	0.08696 ± 0.00523		B*51:04	1	0.00035 ± 0.00035	HLA-DQB1	DQB1*03:01	16	0.27586 ± 0.05869
	B*49:01	213	0.07350 ± 0.00485		B*73:02	1	0.00035 ± 0.00035		DQB1*05:01	11	0.18966 ± 0.05148
	B*14:02	198	0.06832 ± 0.00469		B*78:01	1	0.00035 ± 0.00035		DQB1*06:01	8	0.13793 ± 0.04528
	B*52:01	167	0.05763 ± 0.00433		B*78:02	1	0.00035 ± 0.00035		DQB1*04:02	4	0.06897 ± 0.03327
	B*50:01	147	0.05072 ± 0.00408		B*82:02	1	0.00035 ± 0.00035		DQB1*02:02	3	0.05172 ± 0.02908
	B*51:01	133	0.04589 ± 0.00389		B*83:01	1	0.00035 ± 0.00035		DQB1*03:05	3	0.05172 ± 0.02908
	B*38:01	131	0.04520 ± 0.00386		r ^a	2898			DQB1*06:03	3	0.05172 ± 0.02908
	B*35:01	126	0.04348 ± 0.00379	HLA-C	C*06:02	107	0.25117 ± 0.02101		DQB1*02:01	2	0.03448 ± 0.02396
	B*44:02	104	0.03589 ± 0.00346		C*07:01	107	0.25117 ± 0.02101		DQB1*03:02	2	0.03448 ± 0.02396
	B*18:01	100	0.03451 ± 0.00339		C*04:01	73	0.17136 ± 0.01826		DQB1*06:02	2	0.03448 ± 0.02396
	B*15:17	84	0.02899 ± 0.00312		C*08:02	42	0.09859 ± 0.01444		DQB1*02:03	1	0.01724 ± 0.01709
	B*08:01	83	0.02864 ± 0.00310		C*02:02	25	0.05869 ± 0.01139		DQB1*05:02	1	0.01724 ± 0.01709
	B*44:03	77	0.02657 ± 0.00299		C*03:04	15	0.03521 ± 0.00893		DQB1*06:04	1	0.01724 ± 0.01709
	B*53:01	75	0.02588 ± 0.00295		C*17:01	8	0.01878 ± 0.00658		DQB1*06:09	1	0.01724 ± 0.01709
	B*35:02	69	0.02381 ± 0.00283		C*05:01	6	0.01408 ± 0.00571		r ^a	58	
	B*45:01	65	0.02243 ± 0.00275		C*12:02	6	0.01408 ± 0.00571				
	B*42:01	63	0.02174 ± 0.00271		C*17:03	6	0.01408 ± 0.00571				

AF Allele frequency, r^a total number of alleles, SE Standard error

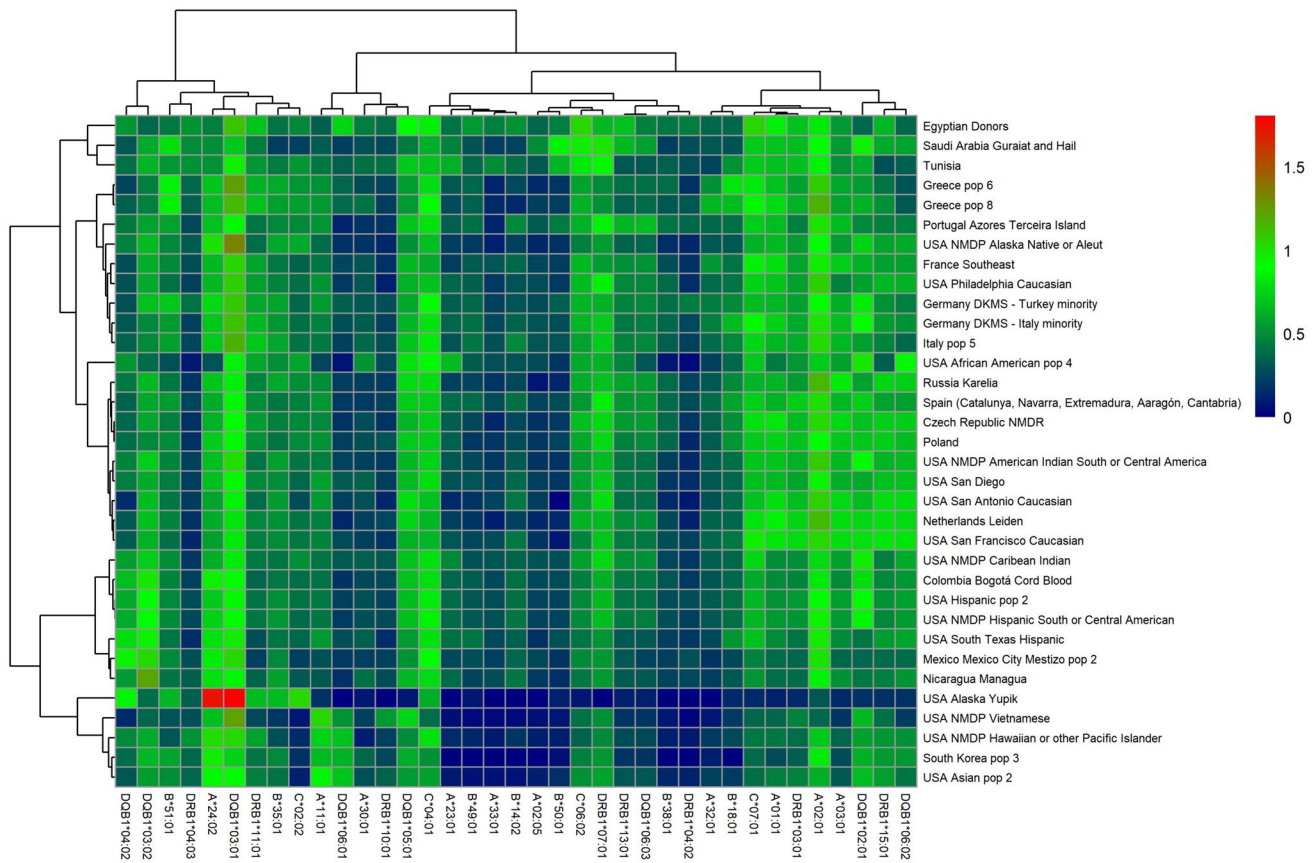


Fig. 1 Heatmap showing clustered and transformed allele frequencies in Egyptian donors and selected populations

In AML patients, the most common HLA class I antigens were A2 (22.8%), B41 (11.3%), and Cw7 (32.4%). The most common alleles were A*01:01 (19.5%), B*41:01 (9.6%), C*04:01 (23%), and C*06:02 (23%). The most common HLA class II antigens were DR13 (18.2%), DQ2 (30%), DQ3 (30%), and DQ4 (30%), while the most common alleles were DRB1*07:01 (12.9%), DQB1*02:01 (30%) and DQB1*04:01 (30%) (ST2 and Table 5).

In SAA patients, the most common HLA class I antigens were A2 (29.6%), B41 (11.3%), and Cw7 (40%). The most common alleles were A*02:01 (19.8%), B*41:01 (9.76%), B*14:02 (9.76%), and C*04:01 (18.5%). The most common HLA class II antigens were DR4 (17.9%) and DQ3 (30%), while the most common alleles were DRB1*03:01 (13.2%), DRB1*15:01 (13.1%), and DQB1*03:01 (25%) (ST2 and Table 5).

Table 4 Hardy-Weinberg and neutrality testing for each locus in donors

Locus	n_{typed}	n_{gametes}	k	f_{obs}	f_{HWE}	f_{IAM}	F_{nd}^{a}	P_{GT}	P_{Slatkin}
HLA-A	1433	2866	52	0.10677	0.08000	0.09950	-0.5950	0.00092	0.29730
HLA-B	1449	2898	76	0.08834	0.03990	0.06520	-1.35810	<0.00001	0.01310
HLA-C	213	426	24	0.46479	0.17140	0.15470	0.30230	<0.00001	0.72140
HLA-DRB1	3686	7372	37	0.10608	0.07990	0.16420	-1.33530	<0.00001	0.00640
HLA-DQB1	29	58	14	0.24138	0.14860	0.15890	-0.22170	0.10433	0.52050

f_{IAM} expected homozygosity under mutation-drift equilibrium (infinite allele model), f_{HWE} expected homozygosity under Hardy-Weinberg equilibrium, F_{nd} normalized deviate of the homozygosity, f_{obs} observed homozygosity, k number of unique alleles, n_{typed} number of patients/donors who have full genotype data for a given locus, n_{gametes} total number of alleles ($2 \times n_{\text{typed}}$), P_{GT} Guo and Thomson’s Hardy-Weinberg Chi-squared test p-value, P_{Slatkin} Slatkin’s neutrality test p-value

^aNegative F_{nd} values imply balancing selection, while positive values imply directional selection

Table 5 HLA class I and II allele frequency in ALL, AML, and SAA patients

Locus	Allele	ALL		Allele	AML		Allele	SAA	
		<i>n</i>	AF ± SE		<i>n</i>	AF ± SE		<i>n</i>	AF ± SE
HLA-A	A*02:01	60	0.17544 ± 0.02057	A*01:01	117	0.19500 ± 0.01617	A*02:01	117	0.19831 ± 0.01642
	A*01:01	55	0.16082 ± 0.01986	A*02:01	102	0.17000 ± 0.01534	A*01:01	78	0.13220 ± 0.01394
	A*03:01	24	0.07018 ± 0.01381	A*03:01	39	0.06500 ± 0.01006	A*03:01	37	0.06271 ± 0.00998
	A*33:01	20	0.05848 ± 0.01269	A*30:01	38	0.06333 ± 0.00994	A*33:01	37	0.06271 ± 0.00998
	A*68:02	18	0.05263 ± 0.01207	A*24:02	37	0.06167 ± 0.00982	A*68:02	35	0.05932 ± 0.00973
	A*30:04	17	0.04971 ± 0.01175	A*32:01	26	0.04333 ± 0.00831	A*02:05	35	0.05932 ± 0.00973
	A*24:02	16	0.04678 ± 0.01142	A*33:01	26	0.04333 ± 0.00831	A*23:01	25	0.04237 ± 0.00829
	A*32:01	15	0.04386 ± 0.01107	A*30:04	23	0.03833 ± 0.00784	A*29:01	23	0.03898 ± 0.00797
	A*30:01	14	0.04094 ± 0.01071	A*68:02	22	0.03667 ± 0.00767	A*30:01	20	0.03390 ± 0.00745
	A*11:01	13	0.03801 ± 0.01034	A*02:02	19	0.03167 ± 0.00715	A*02:02	20	0.03390 ± 0.00745
	A*02:05	12	0.03509 ± 0.00995	A*69:01	19	0.03167 ± 0.00715	A*11:01	20	0.03390 ± 0.00745
	A*26:01	12	0.03509 ± 0.00995	A*11:01	18	0.03000 ± 0.00696	A*24:02	18	0.03051 ± 0.00708
	A*31:01	8	0.02339 ± 0.00817	A*29:01	18	0.03000 ± 0.00696	A*69:01	14	0.02373 ± 0.00627
	A*23:01	7	0.02047 ± 0.00766	A*02:05	17	0.02833 ± 0.00677	A*26:01	14	0.02373 ± 0.00627
	A*30:02	7	0.02047 ± 0.00766	A*23:01	15	0.02500 ± 0.00637	A*32:01	13	0.02203 ± 0.00604
	A*24:03	7	0.02047 ± 0.00766	A*26:01	11	0.01833 ± 0.00548	A*30:04	11	0.01864 ± 0.00557
	A*02:02	6	0.01754 ± 0.00710	A*30:02	11	0.01833 ± 0.00548	A*68:01	10	0.01695 ± 0.00531
	A*29:01	6	0.01754 ± 0.00710	A*03:02	7	0.01167 ± 0.00438	A*30:02	9	0.01525 ± 0.00505
	A*69:01	5	0.01462 ± 0.00649	A*24:03	5	0.00833 ± 0.00371	A*24:03	9	0.01525 ± 0.00505
	A*03:02	5	0.01462 ± 0.00649	A*31:01	5	0.00833 ± 0.00371	A*03:02	8	0.01356 ± 0.00476
	A*68:01	3	0.00877 ± 0.00504	A*34:02	4	0.00667 ± 0.00332	A*31:01	8	0.01356 ± 0.00476
	A*74:01	3	0.00877 ± 0.00504	A*66:01	3	0.00500 ± 0.00288	A*01:03	5	0.00847 ± 0.00377
	A*23:04	2	0.00585 ± 0.00412	A*01:02	2	0.00333 ± 0.00235	A*34:02	3	0.00508 ± 0.00293
	A*66:01	1	0.00292 ± 0.00292	A*02:06	2	0.00333 ± 0.00235	A*66:01	3	0.00508 ± 0.00293
	A*01:02	1	0.00292 ± 0.00292	A*36:01	2	0.00333 ± 0.00235	A*80:01	3	0.00508 ± 0.00293
	A*24:01	1	0.00292 ± 0.00292	A*68:01	2	0.00333 ± 0.00235	A*02:17	3	0.00508 ± 0.00293
	A*02:17	1	0.00292 ± 0.00292	A*80:01	2	0.00333 ± 0.00235	A*74:01	3	0.00508 ± 0.00293
	A*24:07	1	0.00292 ± 0.00292	A*01:03	1	0.00167 ± 0.00167	A*36:01	2	0.00339 ± 0.00239
	A*25:01	1	0.00292 ± 0.00292	A*02:04	1	0.00167 ± 0.00167	A*25:01	2	0.00339 ± 0.00239
	A*29:02	1	0.00292 ± 0.00292	A*02:07	1	0.00167 ± 0.00167	A*74:03	2	0.00339 ± 0.00239
	A*34:02	0	0.00000 ± 0.00000	A*02:08	1	0.00167 ± 0.00167	A*02:08	1	0.00169 ± 0.00169
	A*02:06	0	0.00000 ± 0.00000	A*23:04	1	0.00167 ± 0.00167	A*28:01	1	0.00169 ± 0.00169
A*36:01	0	0.00000 ± 0.00000	A*24:01	1	0.00167 ± 0.00167	A*34:01	1	0.00169 ± 0.00169	
A*80:01	0	0.00000 ± 0.00000	A*24:06	1	0.00167 ± 0.00167	A*01:02	0	0.00000 ± 0.00000	
A*01:03	0	0.00000 ± 0.00000	A*25:02	1	0.00167 ± 0.00167	A*02:06	0	0.00000 ± 0.00000	
A*02:04	0	0.00000 ± 0.00000	A*02:17	0	0.00000 ± 0.00000	A*02:04	0	0.00000 ± 0.00000	
A*02:07	0	0.00000 ± 0.00000	A*24:07	0	0.00000 ± 0.00000	A*02:07	0	0.00000 ± 0.00000	
HLA-A	A*02:08	0	0.00000 ± 0.00000	A*25:01	0	0.00000 ± 0.00000	A*23:04	0	0.00000 ± 0.00000
	A*24:06	0	0.00000 ± 0.00000	A*28:01	0	0.00000 ± 0.00000	A*24:01	0	0.00000 ± 0.00000
	A*25:02	0	0.00000 ± 0.00000	A*29:02	0	0.00000 ± 0.00000	A*24:06	0	0.00000 ± 0.00000
	A*28:01	0	0.00000 ± 0.00000	A*34:01	0	0.00000 ± 0.00000	A*25:02	0	0.00000 ± 0.00000
	A*34:01	0	0.00000 ± 0.00000	A*74:01	0	0.00000 ± 0.00000	A*24:07	0	0.00000 ± 0.00000
	A*74:03	0	0.00000 ± 0.00000	A*74:03	0	0.00000 ± 0.00000	A*29:02	0	0.00000 ± 0.00000
	<i>n</i> ^a	342			600			590	
HLA-B	B*41:01	29	0.08286 ± 0.01473	B*41:01	58	0.09571 ± 0.01195	B*41:01	58	0.09764 ± 0.01218
	B*14:02	24	0.06857 ± 0.01351	B*49:01	39	0.06436 ± 0.00997	B*14:02	58	0.09764 ± 0.01218
	B*52:01	24	0.06857 ± 0.01351	B*14:02	36	0.05941 ± 0.00960	B*49:01	44	0.07407 ± 0.01075
	B*38:01	24	0.06857 ± 0.01351	B*52:01	36	0.05941 ± 0.00960	B*52:01	35	0.05892 ± 0.00966

Table 5 (continued)

Locus	Allele	ALL		Allele	AML		Allele	SAA	
		<i>n</i>	AF ± SE		<i>n</i>	AF ± SE		<i>n</i>	AF ± SE
HLA-B	B*49:01	22	0.06286 ± 0.01297	B*35:01	30	0.04950 ± 0.00881	B*50:01	26	0.04377 ± 0.00839
	B*50:01	20	0.05714 ± 0.01241	B*50:01	29	0.04785 ± 0.00867	B*51:01	25	0.04209 ± 0.00824
	B*15:17	16	0.04571 ± 0.01116	B*38:01	27	0.04455 ± 0.00838	B*35:01	22	0.03704 ± 0.00775
	B*35:01	15	0.04286 ± 0.01083	B*51:01	27	0.04455 ± 0.00838	B*38:01	21	0.03535 ± 0.00758
	B*18:01	15	0.04286 ± 0.01083	B*18:01	24	0.03960 ± 0.00792	B*44:03	18	0.03030 ± 0.00703
	B*44:02	13	0.03714 ± 0.01011	B*44:02	24	0.03960 ± 0.00792	B*18:01	17	0.02862 ± 0.00684
	B*42:01	11	0.03143 ± 0.00933	B*15:17	22	0.03630 ± 0.00760	B*35:02	16	0.02694 ± 0.00664
	B*51:01	10	0.02857 ± 0.00891	B*44:03	20	0.03300 ± 0.00726	B*58:01	16	0.02694 ± 0.00664
	B*53:01	10	0.02857 ± 0.00891	B*53:01	19	0.03135 ± 0.00708	B*15:17	14	0.02357 ± 0.00622
	B*45:01	10	0.02857 ± 0.00891	B*13:02	18	0.02970 ± 0.00690	B*45:01	14	0.02357 ± 0.00622
	B*58:01	8	0.02286 ± 0.00799	B*08:01	17	0.02805 ± 0.00671	B*15:03	14	0.02357 ± 0.00622
	B*08:01	7	0.02000 ± 0.00748	B*45:01	14	0.02310 ± 0.00610	B*44:02	13	0.02189 ± 0.00600
	B*39:01	7	0.02000 ± 0.00748	B*35:02	11	0.01815 ± 0.00542	B*42:01	13	0.02189 ± 0.00600
	B*13:02	6	0.01714 ± 0.00694	B*41:02	11	0.01815 ± 0.00542	B*07:02	13	0.02189 ± 0.00600
	B*35:08	6	0.01714 ± 0.00694	B*42:01	11	0.01815 ± 0.00542	B*08:01	12	0.02020 ± 0.00577
	B*44:03	5	0.01429 ± 0.00634	B*58:01	11	0.01815 ± 0.00542	B*35:08	11	0.01852 ± 0.00553
	B*41:02	5	0.01429 ± 0.00634	B*07:05	8	0.01320 ± 0.00464	B*13:02	10	0.01684 ± 0.00528
	B*57:01	5	0.01429 ± 0.00634	B*27:07	8	0.01320 ± 0.00464	B*53:01	9	0.01515 ± 0.00501
	B*07:02	4	0.01143 ± 0.00568	B*35:08	8	0.01320 ± 0.00464	B*55:01	9	0.01515 ± 0.00501
	B*14:01	4	0.01143 ± 0.00568	B*07:02	7	0.01155 ± 0.00434	B*57:01	9	0.01515 ± 0.00501
	B*37:01	4	0.01143 ± 0.00568	B*14:01	7	0.01155 ± 0.00434	B*07:05	7	0.01178 ± 0.00443
	B*55:01	4	0.01143 ± 0.00568	B*15:03	7	0.01155 ± 0.00434	B*40:01	7	0.01178 ± 0.00443
	B*27:03	4	0.01143 ± 0.00568	B*51:08	7	0.01155 ± 0.00434	B*41:02	6	0.01010 ± 0.00410
	B*35:02	3	0.00857 ± 0.00493	B*15:01	5	0.00825 ± 0.00367	B*37:01	6	0.01010 ± 0.00410
	B*15:03	3	0.00857 ± 0.00493	B*37:01	5	0.00825 ± 0.00367	B*39:01	6	0.01010 ± 0.00410
	B*51:08	3	0.00857 ± 0.00493	B*55:01	5	0.00825 ± 0.00367	B*35:03	6	0.01010 ± 0.00410
	B*42:02	3	0.00857 ± 0.00493	B*57:01	5	0.00825 ± 0.00367	B*27:03	5	0.00842 ± 0.00375
	B*27:07	2	0.00571 ± 0.00403	B*13:01	4	0.00660 ± 0.00329	B*51:05	5	0.00842 ± 0.00375
	B*15:01	2	0.00571 ± 0.00403	B*15:16	4	0.00660 ± 0.00329	B*14:01	4	0.00673 ± 0.00336
	B*13:01	2	0.00571 ± 0.00403	B*39:01	4	0.00660 ± 0.00329	B*51:08	4	0.00673 ± 0.00336
	B*40:01	2	0.00571 ± 0.00403	B*39:03	4	0.00660 ± 0.00329	B*13:01	4	0.00673 ± 0.00336
	B*40:02	2	0.00571 ± 0.00403	B*27:02	3	0.00495 ± 0.00285	B*39:03	3	0.00505 ± 0.00291
	B*44:05	2	0.00571 ± 0.00403	B*27:03	3	0.00495 ± 0.00285	B*57:03	3	0.00505 ± 0.00291
	B*27:05	2	0.00571 ± 0.00403	B*35:03	3	0.00495 ± 0.00285	B*73:01	3	0.00505 ± 0.00291
	B*07:05	1	0.00286 ± 0.00285	B*40:01	3	0.00495 ± 0.00285	B*40:02	3	0.00505 ± 0.00291
	B*15:16	1	0.00286 ± 0.00285	B*57:03	3	0.00495 ± 0.00285	B*47:01	3	0.00505 ± 0.00291
	B*27:02	1	0.00286 ± 0.00285	B*73:01	3	0.00495 ± 0.00285	B*41:05	3	0.00505 ± 0.00291
	B*57:03	1	0.00286 ± 0.00285	B*27:01	2	0.00330 ± 0.00233	B*27:07	2	0.00337 ± 0.00238
	B*73:01	1	0.00286 ± 0.00285	B*51:05	2	0.00330 ± 0.00233	B*15:16	2	0.00337 ± 0.00238
	B*57:02	1	0.00286 ± 0.00285	B*57:02	2	0.00330 ± 0.00233	B*27:02	2	0.00337 ± 0.00238
	B*47:01	1	0.00286 ± 0.00285	B*07:01	1	0.00165 ± 0.00165	B*40:06	2	0.00337 ± 0.00238
	B*15:02	1	0.00286 ± 0.00285	B*14:06	1	0.00165 ± 0.00165	B*47:03	2	0.00337 ± 0.00238
B*15:08	1	0.00286 ± 0.00285	B*40:02	1	0.00165 ± 0.00165	B*51:07	2	0.00337 ± 0.00238	
B*37:04	1	0.00286 ± 0.00285	B*40:06	1	0.00165 ± 0.00165	B*27:08	1	0.00168 ± 0.00168	
B*41:05	1	0.00286 ± 0.00285	B*42:02	1	0.00165 ± 0.00165	B*39:06	1	0.00168 ± 0.00168	
B*44:01	1	0.00286 ± 0.00285	B*44:05	1	0.00165 ± 0.00165	B*73:02	1	0.00168 ± 0.00168	
B*39:03	0	0.00000 ± 0.00000	B*47:01	1	0.00165 ± 0.00165	B*78:02	1	0.00168 ± 0.00168	
B*35:03	0	0.00000 ± 0.00000	B*48:01	1	0.00165 ± 0.00165	B*81:01	1	0.00168 ± 0.00168	

Table 5 (continued)

Locus	Allele	ALL		Allele	AML		Allele	SAA	
		<i>n</i>	AF ± SE		<i>n</i>	AF ± SE		<i>n</i>	AF ± SE
HLA-B	B*27:01	0	0.00000 ± 0.00000	B*51:04	1	0.00165 ± 0.00165	B*82:02	1	0.00168 ± 0.00168
	B*51:05	0	0.00000 ± 0.00000	B*62:01	1	0.00165 ± 0.00165	B*83:01	1	0.00168 ± 0.00168
	B*07:01	0	0.00000 ± 0.00000	B*15:02	0	0.00000 ± 0.00000	B*15:01	0	0.00000 ± 0.00000
	B*14:06	0	0.00000 ± 0.00000	B*15:08	0	0.00000 ± 0.00000	B*27:01	0	0.00000 ± 0.00000
	B*40:06	0	0.00000 ± 0.00000	B*27:05	0	0.00000 ± 0.00000	B*57:02	0	0.00000 ± 0.00000
	B*48:01	0	0.00000 ± 0.00000	B*27:08	0	0.00000 ± 0.00000	B*07:01	0	0.00000 ± 0.00000
	B*51:04	0	0.00000 ± 0.00000	B*37:04	0	0.00000 ± 0.00000	B*14:06	0	0.00000 ± 0.00000
	B*62:01	0	0.00000 ± 0.00000	B*39:06	0	0.00000 ± 0.00000	B*42:02	0	0.00000 ± 0.00000
	B*27:08	0	0.00000 ± 0.00000	B*41:05	0	0.00000 ± 0.00000	B*44:05	0	0.00000 ± 0.00000
	B*39:06	0	0.00000 ± 0.00000	B*44:01	0	0.00000 ± 0.00000	B*48:01	0	0.00000 ± 0.00000
	B*47:03	0	0.00000 ± 0.00000	B*47:03	0	0.00000 ± 0.00000	B*51:04	0	0.00000 ± 0.00000
	B*51:07	0	0.00000 ± 0.00000	B*51:07	0	0.00000 ± 0.00000	B*62:01	0	0.00000 ± 0.00000
	B*73:02	0	0.00000 ± 0.00000	B*73:02	0	0.00000 ± 0.00000	B*15:02	0	0.00000 ± 0.00000
	B*78:02	0	0.00000 ± 0.00000	B*78:02	0	0.00000 ± 0.00000	B*15:08	0	0.00000 ± 0.00000
	B*81:01	0	0.00000 ± 0.00000	B*81:01	0	0.00000 ± 0.00000	B*27:05	0	0.00000 ± 0.00000
	B*82:02	0	0.00000 ± 0.00000	B*82:02	0	0.00000 ± 0.00000	B*37:04	0	0.00000 ± 0.00000
B*83:01	0	0.00000 ± 0.00000	B*83:01	0	0.00000 ± 0.00000	B*44:01	0	0.00000 ± 0.00000	
<i>n</i> ^a	350			606			594		
HLA-C	C*04:01	5	0.19231 ± 0.07729	C*04:01	12	0.23077 ± 0.05843	C*04:01	10	0.18519 ± 0.05286
	C*07:01	5	0.19231 ± 0.07729	C*06:02	12	0.23077 ± 0.05843	C*07:01	9	0.16667 ± 0.05072
	C*06:02	4	0.15385 ± 0.07076	C*07:01	10	0.19231 ± 0.05465	C*08:02	9	0.16667 ± 0.05072
	C*02:02	4	0.15385 ± 0.07076	C*02:02	4	0.07692 ± 0.03695	C*06:02	8	0.14815 ± 0.04834
	C*12:02	2	0.07692 ± 0.05226	C*05:01	4	0.07692 ± 0.03695	C*02:02	5	0.09259 ± 0.03945
	C*03:04	2	0.07692 ± 0.05226	C*08:02	3	0.05769 ± 0.03233	C*01:02	2	0.03704 ± 0.02570
	C*05:01	1	0.03846 ± 0.03771	C*01:02	2	0.03846 ± 0.02667	C*17:01	2	0.03704 ± 0.02570
	C*03:01	1	0.03846 ± 0.03771	C*07:02	2	0.03846 ± 0.02667	C*12:03	2	0.03704 ± 0.02570
	C*12:03	1	0.03846 ± 0.03771	C*12:02	1	0.01923 ± 0.01904	C*17:03	2	0.03704 ± 0.02570
	C*17:03	1	0.03846 ± 0.03771	C*16:01	1	0.01923 ± 0.01904	C*16:01	1	0.01852 ± 0.01835
	C*08:02	0	0.00000 ± 0.00000	C*17:01	1	0.01923 ± 0.01904	C*03:04	1	0.01852 ± 0.01835
	C*01:02	0	0.00000 ± 0.00000	C*03:01	0	0.00000 ± 0.00000	C*11:01	1	0.01852 ± 0.01835
	C*07:02	0	0.00000 ± 0.00000	C*03:04	0	0.00000 ± 0.00000	C*14:02	1	0.01852 ± 0.01835
	C*16:01	0	0.00000 ± 0.00000	C*11:01	0	0.00000 ± 0.00000	C*16:02	1	0.01852 ± 0.01835
	C*17:01	0	0.00000 ± 0.00000	C*12:03	0	0.00000 ± 0.00000	C*05:01	0	0.00000 ± 0.00000
	C*11:01	0	0.00000 ± 0.00000	C*14:02	0	0.00000 ± 0.00000	C*07:02	0	0.00000 ± 0.00000
	C*14:02	0	0.00000 ± 0.00000	C*16:02	0	0.00000 ± 0.00000	C*12:02	0	0.00000 ± 0.00000
	C*16:02	0	0.00000 ± 0.00000	C*17:03	0	0.00000 ± 0.00000	C*03:01	0	0.00000 ± 0.00000
<i>n</i> ^a	26			52			54		
HLA-DRB1	DRB1*13:01	67	0.13563 ± 0.01540	DRB1*07:01	113	0.12929 ± 0.01135	DRB1*03:01	110	0.13221 ± 0.01174
	DRB1*03:01	54	0.10931 ± 0.01404	DRB1*13:01	112	0.12815 ± 0.01131	DRB1*15:01	109	0.13101 ± 0.01170
	DRB1*07:01	53	0.10729 ± 0.01392	DRB1*03:01	101	0.11556 ± 0.01081	DRB1*07:01	89	0.10697 ± 0.01072
	DRB1*11:01	52	0.10526 ± 0.01381	DRB1*11:01	98	0.11213 ± 0.01067	DRB1*11:01	88	0.10577 ± 0.01066
	DRB1*15:01	41	0.08300 ± 0.01241	DRB1*15:01	82	0.09382 ± 0.00986	DRB1*13:01	82	0.09856 ± 0.01033
	DRB1*04:03	41	0.08300 ± 0.01241	DRB1*04:02	50	0.05721 ± 0.00786	DRB1*04:03	59	0.07091 ± 0.00890
	DRB1*13:03	28	0.05668 ± 0.01040	DRB1*04:03	47	0.05378 ± 0.00763	DRB1*01:02	58	0.06971 ± 0.00883
	DRB1*10:01	28	0.05668 ± 0.01040	DRB1*13:03	43	0.04920 ± 0.00732	DRB1*04:02	38	0.04567 ± 0.00724
	DRB1*01:02	27	0.05466 ± 0.01023	DRB1*10:01	37	0.04233 ± 0.00681	DRB1*10:01	35	0.04207 ± 0.00696
	DRB1*04:02	23	0.04656 ± 0.00948	DRB1*01:02	34	0.03890 ± 0.00654	DRB1*13:03	32	0.03846 ± 0.00667
DRB1*08:02	20	0.04049 ± 0.00887	DRB1*04:05	31	0.03547 ± 0.00626	DRB1*04:05	32	0.03846 ± 0.00667	

Table 5 (continued)

Locus	Allele	ALL		Allele	AML		Allele	SAA	
		<i>n</i>	AF ± SE		<i>n</i>	AF ± SE		<i>n</i>	AF ± SE
	DRB1*14:01	9	0.01822 ± 0.00602	DRB1*14:01	30	0.03432 ± 0.00616	DRB1*04:04	16	0.01923 ± 0.00476
	DRB1*03:02	9	0.01822 ± 0.00602	DRB1*03:02	17	0.01945 ± 0.00467	DRB1*03:02	15	0.01803 ± 0.00461
	DRB1*04:05	7	0.01417 ± 0.00532	DRB1*04:04	16	0.01831 ± 0.00453	DRB1*08:02	14	0.01683 ± 0.00446
	DRB1*04:04	5	0.01012 ± 0.00450	DRB1*08:02	11	0.01259 ± 0.00377	DRB1*11:02	10	0.01202 ± 0.00378
	DRB1*01:01	5	0.01012 ± 0.00450	DRB1*11:02	8	0.00915 ± 0.00322	DRB1*01:01	10	0.01202 ± 0.00378
	DRB1*13:05	5	0.01012 ± 0.00450	DRB1*12:01	8	0.00915 ± 0.00322	DRB1*14:01	8	0.00962 ± 0.00338
	DRB1*11:02	3	0.00607 ± 0.00350	DRB1*01:01	6	0.00686 ± 0.00279	DRB1*12:01	4	0.00481 ± 0.00240
	DRB1*12:01	3	0.00607 ± 0.00350	DRB1*08:01	5	0.00572 ± 0.00255	DRB1*04:01	4	0.00481 ± 0.00240
	DRB1*04:01	3	0.00607 ± 0.00350	DRB1*09:01	5	0.00572 ± 0.00255	DRB1*08:01	3	0.00361 ± 0.00208
	DRB1*08:01	2	0.00405 ± 0.00286	DRB1*16:02	4	0.00458 ± 0.00228	DRB1*09:01	3	0.00361 ± 0.00208
	DRB1*09:01	2	0.00405 ± 0.00286	DRB1*04:01	3	0.00343 ± 0.00198	DRB1*13:05	3	0.00361 ± 0.00208
	DRB1*03:07	2	0.00405 ± 0.00286	DRB1*13:05	3	0.00343 ± 0.00198	DRB1*16:02	2	0.00240 ± 0.00170
	DRB1*16:01	2	0.00405 ± 0.00286	DRB1*11:03	2	0.00229 ± 0.00162	DRB1*03:07	2	0.00240 ± 0.00170
	DRB1*16:02	1	0.00202 ± 0.00202	DRB1*14:04	2	0.00229 ± 0.00162	DRB1*14:04	1	0.00120 ± 0.00120
	DRB1*12:02	1	0.00202 ± 0.00202	DRB1*01:03	1	0.00114 ± 0.00114	DRB1*16:01	1	0.00120 ± 0.00120
	DRB1*14:06	1	0.00202 ± 0.00202	DRB1*03:07	1	0.00114 ± 0.00114	DRB1*03:04	1	0.00120 ± 0.00120
	DRB1*11:03	0	0.00000 ± 0.00000	DRB1*13:07	1	0.00114 ± 0.00114	DRB1*08:03	1	0.00120 ± 0.00120
	DRB1*14:04	0	0.00000 ± 0.00000	DRB1*14:02	1	0.00114 ± 0.00114	DRB1*12:02	1	0.00120 ± 0.00120
	DRB1*01:03	0	0.00000 ± 0.00000	DRB1*15:02	1	0.00114 ± 0.00114	DRB1*13:04	1	0.00120 ± 0.00120
	DRB1*13:07	0	0.00000 ± 0.00000	DRB1*16:01	1	0.00114 ± 0.00114	DRB1*11:03	0	0.00000 ± 0.00000
	DRB1*14:02	0	0.00000 ± 0.00000	DRB1*03:04	0	0.00000 ± 0.00000	DRB1*01:03	0	0.00000 ± 0.00000
	DRB1*15:02	0	0.00000 ± 0.00000	DRB1*08:03	0	0.00000 ± 0.00000	DRB1*13:07	0	0.00000 ± 0.00000
	DRB1*03:04	0	0.00000 ± 0.00000	DRB1*12:02	0	0.00000 ± 0.00000	DRB1*14:02	0	0.00000 ± 0.00000
	DRB1*08:03	0	0.00000 ± 0.00000	DRB1*13:04	0	0.00000 ± 0.00000	DRB1*15:02	0	0.00000 ± 0.00000
	DRB1*13:04	0	0.00000 ± 0.00000	DRB1*14:06	0	0.00000 ± 0.00000	DRB1*14:06	0	0.00000 ± 0.00000
	<i>n</i> ^a	494			874			832	
HLA-DQB1	DQB1*04:02	3	0.18750 ± 0.09758	DQB1*02:01	3	0.30000 ± 0.14491	DQB1*03:01	5	0.25000 ± 0.09682
	DQB1*03:01	3	0.18750 ± 0.09758	DQB1*04:02	3	0.30000 ± 0.14491	DQB1*02:01	3	0.15000 ± 0.07984
	DQB1*05:01	3	0.18750 ± 0.09758	DQB1*03:01	2	0.20000 ± 0.12649	DQB1*05:01	3	0.15000 ± 0.07984
HLA-DQB1	DQB1*02:01	1	0.06250 ± 0.06052	DQB1*03:05	1	0.10000 ± 0.09487	DQB1*02:02	2	0.10000 ± 0.06708
	DQB1*06:03	1	0.06250 ± 0.06052	DQB1*06:03	1	0.10000 ± 0.09487	DQB1*06:01	2	0.10000 ± 0.06708
	DQB1*02:03	1	0.06250 ± 0.06052	DQB1*02:02	0	0.00000 ± 0.00000	DQB1*06:02	2	0.10000 ± 0.06708
	DQB1*03:02	1	0.06250 ± 0.06052	DQB1*02:03	0	0.00000 ± 0.00000	DQB1*04:02	1	0.05000 ± 0.04873
	DQB1*06:01	1	0.06250 ± 0.06052	DQB1*03:02	0	0.00000 ± 0.00000	DQB1*06:03	1	0.05000 ± 0.04873
	DQB1*06:02	1	0.06250 ± 0.06052	DQB1*05:01	0	0.00000 ± 0.00000	DQB1*03:02	1	0.05000 ± 0.04873
	DQB1*06:04	1	0.06250 ± 0.06052	DQB1*06:01	0	0.00000 ± 0.00000	DQB1*03:05	0	0.00000 ± 0.00000
	DQB1*03:05	0	0.00000 ± 0.00000	DQB1*06:02	0	0.00000 ± 0.00000	DQB1*02:03	0	0.00000 ± 0.00000
	DQB1*02:02	0	0.00000 ± 0.00000	DQB1*06:04	0	0.00000 ± 0.00000	DQB1*06:04	0	0.00000 ± 0.00000
	<i>n</i> ^a	16			10			20	

AF Allele frequency, ALL Acute lymphoblastic leukemia, AML Acute myeloid leukemia, *n*^a total number of alleles, SAA Severe aplastic anemia, SE Standard error

Table 6 Most common HLA-A~B~DRB1 haplotypes in donors

Haplotype	Count (2N = 2342 ^a)	HF ± SE
A*33:01~B*14:02~DRB1*01:02	55	0.02351 ± 0.00314
A*01:01~B*52:01~DRB1*15:01	50	0.02117 ± 0.00298
A*01:01~B*08:01~DRB1*03:01	21	0.00887 ± 0.00194
A*11:01~B*52:01~DRB1*15:01	20	0.00847 ± 0.00190
A*30:01~B*42:01~DRB1*03:02	19	0.00811 ± 0.00186
A*01:01~B*49:01~DRB1*13:01	17	0.00736 ± 0.00177
A*02:01~B*52:01~DRB1*15:01	17	0.00730 ± 0.00176
A*02:01~B*41:01~DRB1*07:01	17	0.00713 ± 0.00174
A*24:02~B*35:02~DRB1*11:01	16	0.00675 ± 0.00169
A*02:05~B*50:01~DRB1*07:01	15	0.00630 ± 0.00164

HF Haplotype frequency, SE standard error

^aHLA-A, -B, and -DRB1 full typing was complete for 1171 donors

Haplotype frequencies and association with hematological conditions

The EM algorithm identified 1230 unique haplotypes; the most frequent were HLA-A*33:01~B*14:02~DRB1*01:02 (2.35%), HLA-A*01:01~B*52:01~DRB1*15:01 (2.11%), and HLA-A*01:01~B*08:01~DRB1*03:01 (0.89%) (Table 6). The most common haplotypes in

ALL patients were A*01:01~B*52:01~DRB1*15:01 (2.45%), A*33:01~B*14:02~DRB1*01:02 (2.45%), and A*32:01~B*41:01~DRB1*03:01 (1.84%). The most common haplotypes in AML and myeloid sarcoma (MS) were A*01:01~B*52:01~DRB1*15:01 (2.62%), A*33:01~B*14:02~DRB1*01:02 (2.27%), and A*02:01~B*41:01~DRB1*07:01 (1.92%). The most common haplotypes in SAA were A*01:01~B*52:01~DRB1*15:01 (2.32%), A*33:01~B*14:02~DRB1*01:02 (2.14%), and A*33:01~B*14:02~DRB1*11:01 (1.25%). The haplotype A*02:01~B*18:01~DRB1*03:01 frequency was significantly higher in patients with ALL than donors (1.22% vs. 0.13%). Additionally, the frequencies of the haplotypes A*33:01~B*14:02~DRB1*11:01 and A*02:01~B*14:02~DRB1*01:02 (1.25% and 1.07% respectively) were significantly higher than donors (0.17% and 0.26% respectively) (Table 7).

Pairwise linkage disequilibrium (LD) in donors

The haplotypes A*31:01~C*06:02, B*13:01~C*06:02, and B*44:02~C*05:01, were in complete linkage disequilibrium. A strong LD was found between B*50:01~C*06:02, B*52:01~C*12:02, B*14:02~C*08:02,

Table 7 HLA-A~B~DRB1 haplotype frequencies in patients and donors

Haplotype ^{a,b}	Patients (freq)	Donors (freq)	P_{FET}	Q_{FET}^b	OR (95% CI)
ALL					
A*01:01~B*52:01~DRB1*15:01	8/326 (0.02453)	50/2342 (0.02134)	0.684	0.821	1.15 (0.47–2.48)
A*33:01~B*14:02~DRB1*01:02	8/326 (0.02453)	55/2342 (0.02348)	0.846	0.846	1.05 (0.43–2.24)
A*32:01~B*41:01~DRB1*03:01	6/326 (0.01840)	14/2342 (0.00600)	0.027	0.081	3.11 (0.97–8.70)
A*02:01~B*52:01~DRB1*15:01	4/326 (0.01227)	17/2342 (0.00726)	0.313	0.470	1.70 (0.41–5.25)
A*02:01~B*18:01~DRB1*03:01	4/326 (0.01227)	3/2342 (0.00128)	0.006	0.036	9.67 (1.63–66.4)
A*01:01~B*38:01~DRB1*13:01	4/326 (0.01227)	8/2342 (0.00342)	0.049	0.098	3.62 (0.79–13.6)
AML/MS					
A*01:01~B*52:01~DRB1*15:01	15/572 (0.02622)	50/2342 (0.02134)	0.527	0.878	1.23 (0.64–2.25)
A*33:01~B*14:02~DRB1*01:02	13/572 (0.02273)	55/2342 (0.02348)	1.000	1.000	0.97 (0.48–1.80)
A*02:01~B*41:01~DRB1*07:01	11/572 (0.01923)	17/2342 (0.00726)	0.015	0.060	2.68 (1.13–6.11)
A*01:01~B*13:02~DRB1*07:01	7/572 (0.01224)	9/2342 (0.00384)	0.024	0.060	3.20 (1.01–9.73)
A*01:01~B*08:01~DRB1*03:01	6/572 (0.01049)	21/2342 (0.00700)	0.807	1.000	1.17 (0.39–3.02)
SAA					
A*01:01~B*52:01~DRB1*15:01	13/560 (0.02321)	50/2342 (0.02134)	0.748	0.876	1.09 (0.54–2.05)
A*33:01~B*14:02~DRB1*01:02	12/560 (0.02143)	55/2342 (0.02348)	0.876	0.876	0.91 (0.44–1.74)
A*33:01~B*14:02~DRB1*11:01	7/560 (0.01250)	4/2342 (0.00171)	0.002	0.010	7.39 (1.87–34.6)
A*02:01~B*52:01~DRB1*15:01	6/560 (0.01069)	17/2342 (0.00726)	0.425	0.708	1.48 (0.48–3.96)
A*02:01~B*14:02~DRB1*01:02	6/560 (0.01069)	6/2342 (0.00256)	0.016	0.040	4.21 (1.21–15.8)

ALL acute lymphoblastic leukemia, AML acute myeloid leukemia, FET Fisher's exact test, freq frequency, MS Myeloid Sarcoma, SAA severe aplastic anemia

^aHaplotypes with frequency < 0.01 in patients are ignored

^bBenjamini-Hochberg's method for controlling FDR

C*12:02~DRB1*15:01, DRB1*11:01~DQB1*03:01, and DRB1*10:01~DQB1*05:01 (ST3).

HLA antigens and alleles association with ALL

Only HLA-B38 showed a trend towards increased odds of ALL in the additive (OR, 1.52; 95% CI, 1.00–2.30; $p=0.049$; $q=0.729$) and the dominant models (OR, 1.64; 95% CI, 1.06–2.54; $p=0.026$, $q=0.392$) (ST4).

HLA antigens and alleles association with AML

HLA-DR1 showed a trend towards lower odds of AML in the additive (OR, 0.67; 95% CI, 0.48–0.92; $p=0.013$, $q=0.157$) and dominant models (OR, 0.66; 95% CI, 0.47–0.93; $p=0.018$, q -value = 0.214). On the other hand, DR7 showed a trend towards higher AML odds (OR, 1.25; 95% CI, 1.02–1.54; $p=0.033$, $q=0.200$). HLA-DRB1*01:02 showed a trend towards lower AML odds (OR, 0.64; 95% CI, 0.45–0.91; $p=0.014$, $q=0.214$) in the additive model. HLA-DRB1*07:01 was associated with higher AML odds in both the additive (OR, 1.26; 95% CI, 1.02–1.55; $p=0.030$, $q=0.296$) and dominant models (OR, 1.27; 95% CI, 1.00–1.62; $p=0.045$, $q=0.454$) (ST4 and Table 8).

HLA antigens and alleles association with SAA

HLA-A2 showed a significant association with higher SAA odds in the additive (OR, 1.38; 95% CI, 1.16–1.64; $p<0.001$, $q=0.007$), dominant (OR, 1.44; 95% CI, 1.15–1.81; $p=0.002$, $q=0.041$), and recessive models (OR, 1.71; 95% CI, 1.16–2.51; $p=0.006$, $q=0.037$). Furthermore, HLA-B14 and -DR15 showed a trend towards higher SAA odds in the additive (B14: OR, 1.41; 95% CI, 1.09–1.83; $p=0.008$, $q=0.250$; DR15: OR, 1.31; 95% CI, 1.06–1.62; $p=0.012$, $q=0.157$) and dominant models (B14: OR, 1.43; 95% CI, 1.06–1.92; $p=0.018$, $q=0.392$; DR15: OR, 1.38; 95% CI, 1.08–1.75; $p=0.009$, $q=0.210$). On the other hand, HLA-A30 and -DR13 showed a trend toward lower odds of SAA in both the additive (A30: OR, 0.72; 95% CI, 0.53–0.98; $p=0.036$, $q=0.349$; DR13: OR, 0.79; 95% CI, 0.65–0.97; $p=0.025$, $q=0.200$) and recessive models (A30: OR, 0.70; 95% CI, 0.50–0.97; $p=0.034$, $q=0.413$; DR13: OR, 0.79; 95% CI, 0.63–1.00; $p=0.047$, $q=0.281$). Finally, Cw6 showed a trend towards lower SAA odds in the additive model (OR, 0.67; 95% CI, 0.46–0.98; $p=0.039$, $q=0.530$) (ST4).

HLA-A*02:01, -B*14:02, and -DRB1*15:01 alleles were associated with higher odds of SAA in the additive (A*02:01: OR, 1.35; 95% CI, 1.07–1.70; $p=0.010$, $q=0.242$; B*14:02: OR, 1.43; 95% CI, 1.06–1.93; $p=0.020$, $q=0.425$; DRB1*15:01: OR, 1.32; 95% CI, 1.07–1.64; $p=0.011$, $q=0.214$) and dominant models (A*02:01: OR, 1.36; 95% CI, 1.03–1.80; $p=0.032$, $q=0.411$; B*14:02: OR,

1.44; 95% CI, 1.01–2.07; $p=0.046$, $q=0.818$; DRB1*15:01: OR, 1.38; 95% CI, 1.09–1.76; $p=0.008$, $q=0.246$). Furthermore, A*02:01 showed a trend towards increased SAA odds in the recessive model (OR, 1.90; 95% CI, 1.07–3.38; $p=0.029$, $q=0.176$). On the other hand, A*24:02 showed a trend towards decreased SAA odds (OR, 0.59; 95% CI, 0.36–0.96; $p=0.035$, $q=0.373$) (Table 8).

Discussion

The BMT unit at NIH, Cairo became operational in 1997, and the transplant rates increased dramatically since that time. The challenges that face the clinical hematopoietic stem cell transplantation program in Egypt are financial, technological, and administrative (Mahmoud et al. 2008; Mahmoud et al. 2020). However, this is our first report on HLA-A, -B, -C, -DRB1, and -DQB1 allele and haplotype frequencies from HLA data of 6000 individuals from different regions within Egypt. The history of Egypt reflects its geographical location at the crossroads of several cultures. In addition, Egypt took part in many regional empires throughout its history, which explains the complexity and heterogeneity of its genetic make-up. Allele and haplotype frequencies encountered in this study were similar to those reported by AFND (Gonzalez-Galarza et al. 2020).

In this study, 52, 76, 24, 37, and 14 HLA-A, -B, -C, -DRB1, and -DQB1 alleles were observed in Egyptian donors. Among the HLA-A alleles, A*01:01 (16.9%) and A*02:01 (16.1%) showed the highest frequency and represented 33% of HLA -A allelic diversity. HLA-B*41:01 (8.7%), B*49:01 (7.4%), B*14:02 (6.8%), B*52:01 (5.8%), and B*50:01 (5.1%) accounted for more than 33% of genetic diversity at the B locus. Although HLA-B*50:01 was reported to be frequent in Arabs (Jawdat et al. 2014) Caucasians, North Africans, and West-South Asians (Gonzalez-Galarza et al. 2020), B*41:01 was the most frequent B allele observed in Egyptians.

The most frequent HLA-C alleles were C*06:02 (25.1%), C*07:01 (25.1%), and C*04:01 (17.1%), accounting for 63% of HLA-C diversity. For HLA-DRB1, DRB1*11:01 (11.8%), DRB1*03:01 (11.6%), and DRB1*13:01 (11.4%) showed a frequency of more than 33%, followed by DRB1*07:01 (10.5%), and DRB1*15:01 (10.2%), which collectively accounted for 55% of HLA-DRB1 diversity in this study. Both HLA-DRB1*07:01:01G and DRB1*03:01:01G are common in Arabs of Central and Eastern regions of Saudi (Jawdat et al. 2020), and are frequent in Arabs from Tunisia and Jordan (Gonzalez-Galarza et al. 2020). For HLA-DQB1, five alleles showed a frequency of almost 70%. These alleles include DQB1*03:01 (27.5%), DQB1*05:01 (18.9%), DQB1*06:01 (13.8%), DQB1*04:02

Table 8 Association between HLA alleles and the occurrence of ALL, AML, and SAA

Model	Locus	Allele ^a	ALL			AML			SAA					
			OR	95% CI	p-value	q-value ^b	OR	95% CI	p-value	q-value ^b	OR	95% CI	p-value	q-value ^b
Additive	HLA-A	A*01:01	0.95	0.69–1.30	0.742	0.847	1.14	0.91–1.44	0.263	0.693	0.76	0.59–1.00	0.047	0.373
		A*02:01	1.13	0.84–1.52	0.418	0.693	1.05	0.83–1.32	0.707	0.847	1.35	1.07–1.70	0.010	0.242
		A*03:01	0.96	0.62–1.49	0.869	0.907	0.85	0.60–1.21	0.363	0.693	0.87	0.61–1.25	0.465	0.693
		A*24:02	0.89	0.53–1.50	0.664	0.838	1.16	0.80–1.68	0.427	0.693	0.59	0.36–0.96	0.035	0.373
		A*30:01	0.83	0.48–1.44	0.507	0.693	1.28	0.89–1.84	0.190	0.651	0.70	0.43–1.12	0.136	0.651
		A*33:01	1.19	0.73–1.92	0.491	0.693	0.82	0.53–1.26	0.359	0.693	1.31	0.90–1.90	0.163	0.651
	HLA-B	A*68:02	1.06	0.63–1.77	0.832	0.907	0.70	0.44–1.11	0.133	0.651	1.23	0.83–1.82	0.295	0.693
		A*11:01	1.30	0.71–2.38	0.388	0.693	0.98	0.58–1.65	0.945	0.945	1.18	0.71–1.95	0.520	0.693
		B*14:02	0.95	0.61–1.45	0.798	0.979	0.85	0.59–1.21	0.371	0.924	1.43	1.06–1.93	0.020	0.425
		B*35:01	0.87	0.51–1.49	0.623	0.924	1.11	0.74–1.66	0.605	0.924	0.80	0.51–1.26	0.330	0.924
		B*38:01	1.39	0.88–2.20	0.152	0.924	0.99	0.64–1.51	0.949	0.979	0.73	0.46–1.18	0.197	0.924
		B*41:01	0.86	0.57–1.29	0.456	0.924	1.10	0.81–1.50	0.539	0.924	1.09	0.80–1.48	0.575	0.924
HLA-C	B*49:01	0.77	0.49–1.23	0.275	0.924	0.85	0.59–1.22	0.384	0.924	0.98	0.69–1.39	0.926	0.979	
	B*50:01	1.03	0.64–1.66	0.898	0.979	0.91	0.61–1.37	0.660	0.924	0.82	0.54–1.26	0.368	0.924	
	B*52:01	1.11	0.71–1.74	0.646	0.924	1.02	0.70–1.49	0.912	0.979	1.00	0.68–1.46	0.979	0.979	
	C*04:01	1.34	0.46–3.93	0.589	0.643	1.64	0.78–3.44	0.192	0.576	1.20	0.55–2.61	0.643	0.643	
	C*07:01	0.71	0.27–1.89	0.496	0.643	0.76	0.37–1.54	0.444	0.643	0.60	0.29–1.24	0.168	0.576	
	DRB1*01:02	0.92	0.62–1.38	0.701	0.871	0.64	0.45–0.91	0.014	0.214	1.20	0.90–1.60	0.210	0.533	
	DRB1*03:01	0.93	0.70–1.24	0.643	0.871	1.00	0.81–1.24	0.996	0.996	1.15	0.93–1.41	0.190	0.533	
Dominant	HLA-A	DRB1*04:02	0.91	0.59–1.39	0.658	0.871	1.15	0.85–1.55	0.365	0.689	0.90	0.64–1.26	0.530	0.864
		DRB1*04:03	1.16	0.84–1.59	0.368	0.689	0.79	0.59–1.06	0.121	0.533	1.00	0.77–1.31	0.991	0.996
		DRB1*07:01	1.02	0.76–1.36	0.903	0.967	1.26	1.02–1.55	0.030	0.296	1.01	0.81–1.28	0.899	0.967
		DRB1*10:01	1.42	0.96–2.11	0.081	0.533	1.10	0.78–1.56	0.576	0.864	1.07	0.75–1.52	0.726	0.871
		DRB1*11:01	0.87	0.65–1.17	0.352	0.689	0.94	0.75–1.17	0.568	0.864	0.88	0.69–1.11	0.265	0.611
		DRB1*13:01	1.19	0.91–1.55	0.213	0.533	1.16	0.94–1.44	0.162	0.533	0.83	0.65–1.06	0.132	0.533
	HLA-B	DRB1*13:03	1.08	0.73–1.60	0.697	0.871	0.96	0.69–1.32	0.783	0.904	0.74	0.51–1.06	0.099	0.533
		DRB1*15:01	0.80	0.58–1.11	0.182	0.533	0.92	0.72–1.16	0.469	0.828	1.32	1.07–1.64	0.011	0.214
		A*01:01	0.94	0.65–1.35	0.734	0.840	1.17	0.89–1.54	0.253	0.760	0.74	0.55–1.00	0.051	0.411
		A*02:01	1.10	0.77–1.58	0.587	0.828	1.06	0.80–1.40	0.680	0.840	1.36	1.03–1.80	0.032	0.411
		A*03:01	1.04	0.65–1.66	0.865	0.865	0.85	0.58–1.24	0.393	0.825	0.87	0.59–1.29	0.499	0.825
		A*11:01	1.33	0.72–2.45	0.362	0.825	0.93	0.54–1.60	0.805	0.840	1.20	0.72–2.00	0.486	0.825
HLA-C	A*24:02	0.93	0.54–1.62	0.805	0.840	1.13	0.76–1.69	0.550	0.825	0.59	0.36–0.99	0.047	0.411	
	A*30:01	0.87	0.49–1.56	0.641	0.840	1.35	0.91–1.99	0.138	0.552	0.68	0.41–1.13	0.138	0.552	
	A*33:01	1.18	0.70–1.97	0.534	0.825	0.84	0.54–1.31	0.443	0.825	1.30	0.87–1.94	0.197	0.676	
	A*68:02	1.07	0.63–1.81	0.793	0.840	0.67	0.41–1.08	0.100	0.552	1.22	0.81–1.82	0.342	0.825	

Table 8 (continued)

Model	Locus	Allele ^a	ALL				AML				SAA			
			OR	95% CI	p-value	q-value ^b	OR	95% CI	p-value	q-value ^b	OR	95% CI	p-value	q-value ^b
	HLA-B	B*14:02	1.05	0.65–1.68	0.852	0.895	0.80	0.53–1.21	0.300	0.891	1.44	1.01–2.07	0.046	0.818
		B*35:01	0.81	0.44–1.48	0.491	0.891	1.19	0.77–1.85	0.442	0.891	0.85	0.53–1.39	0.522	0.891
		B*38:01	1.54	0.95–2.50	0.081	0.818	1.00	0.63–1.58	>0.999	>0.999	0.72	0.44–1.18	0.195	0.818
		B*41:01	0.88	0.56–1.38	0.587	0.891	1.12	0.79–1.59	0.511	0.891	1.09	0.77–1.54	0.630	0.891
		B*49:01	0.71	0.43–1.18	0.187	0.818	0.88	0.60–1.30	0.525	0.891	0.95	0.65–1.39	0.806	0.891
		B*50:01	1.11	0.67–1.86	0.679	0.891	0.92	0.60–1.43	0.723	0.891	0.71	0.44–1.15	0.164	0.818
		B*52:01	1.09	0.66–1.78	0.745	0.891	1.06	0.71–1.58	0.794	0.891	1.06	0.71–1.59	0.762	0.891
	HLA-DRB1	DRB1*01:02	0.90	0.59–1.37	0.632	0.861	0.64	0.45–0.92	0.017	0.260	1.19	0.88–1.61	0.253	0.632
		DRB1*03:01	0.93	0.67–1.28	0.659	0.861	0.99	0.77–1.26	0.907	0.938	1.10	0.86–1.40	0.436	0.818
		DRB1*04:02	0.95	0.61–1.48	0.810	0.930	1.18	0.85–1.62	0.322	0.689	0.91	0.64–1.30	0.602	0.861
		DRB1*04:03	1.20	0.84–1.73	0.312	0.689	0.80	0.58–1.10	0.174	0.577	0.93	0.69–1.27	0.669	0.861
		DRB1*07:01	1.06	0.77–1.47	0.703	0.861	1.27	1.00–1.62	0.045	0.454	0.97	0.75–1.26	0.837	0.930
		DRB1*10:01	1.44	0.95–2.20	0.087	0.577	1.11	0.77–1.60	0.577	0.861	1.00	0.68–1.46	0.990	0.990
		DRB1*11:01	0.81	0.58–1.13	0.211	0.577	0.93	0.72–1.18	0.534	0.861	0.85	0.66–1.10	0.212	0.577
		DRB1*13:01	1.24	0.92–1.67	0.159	0.577	1.12	0.88–1.42	0.364	0.727	0.82	0.63–1.06	0.136	0.577
		DRB1*13:03	1.10	0.73–1.67	0.644	0.861	0.97	0.70–1.37	0.883	0.938	0.73	0.50–1.07	0.108	0.577
		DRB1*15:01	0.80	0.56–1.13	0.205	0.577	0.95	0.74–1.23	0.717	0.861	1.38	1.09–1.76	0.008	0.246
Recessive	HLA-A	A*01:01	0.95	0.37–2.44	0.911	0.938	1.15	0.58–2.28	0.681	0.938	0.66	0.28–1.58	0.355	0.711
		A*02:01	1.49	0.69–3.22	0.313	0.711	1.03	0.51–2.07	0.938	0.938	1.90	1.07–3.38	0.029	0.176
	HLA-DRB1	DRB1*11:01	1.45	0.57–3.67	0.438	0.978	0.99	0.42–2.32	0.978	0.978	1.03	0.44–2.43	0.941	0.978

ALL Acute lymphoblastic leukemia, AML Acute myeloid leukemia, CI confidence interval, OR odds ratio, SAA Severe aplastic anemia

^aAlleles with frequencies of at least 0.03 and counts of at least 5 in all patient groups and donors are included in the analysis

^bAdjusted p-values for false-detection rate using Benjamini-Hochberg's method

(6.9%), and DQB1*02:02 (5.1%). HLA-DQB1*02:01:01 and DQB1*03:01:01 are frequent in Arabs, while HLA-DQB1*03:02:01 is frequent in Arabs, Chinese, Indians, and Caucasians (Gonzalez-Galarza et al. 2020).

The most frequent HLA alleles in this study were A*01:01, DRB1*11:01, DRB1*03:01, DRB1*13:01, DRB1*07:01, DQB1*03:01, DQB1*05:01, DQB1*06:01, and DQB1*04:02. These alleles are listed in the catalog from the European Federation for Immunogenetics (EFI) as common “COM” alleles within all European sub-regions (Sanchez-Mazas et al. 2017). The frequencies of A*02:01, A*01:01, A*03:01, C*07:01, C*04:01, DRB1*15:01, DRB1*07:01, and DRB1*03:01 in the European American population (Creary et al. 2019) were similar in their frequencies to Egyptians.

Regarding allelic frequencies, HLA-A*01:01 (16.9%) frequency was similar to Caucasians of Australia (18.7%), Belgians (15.5%), Austrians (14.6%), peoples of North-western England (20.8%) and Southeastern France (15%), Czechs (16%), Palestinian Arabs of Gaza (17.8%), French BM donors (13%), and Germans (15.3%). HLA-A*02:01 frequency was 16.6%, which is similar to Brazilians (19.2%) and Chinese (18.7%), but lower than Argentinians (20–40%), Austrians (29.4%), Belgians (26.6%), German donors (28.3%), and the French BM donor registry (29%).

HLA-B*41:01 frequency was 8.7% in this analysis, which is higher than that of the Western population (1–2%). HLA-B*49:01 (7.1%) was similar in frequency to Armenians (9.5%), lower than Israeli Ethiopian Jews (18.6%), and higher than Western population (3%). HLA-B*14:02 (6.9%) was similar in frequency to Armenians (6%), lower than Ashkenazi and Polish Jews (10%), and higher than French and Germans (2–3%).

HLA-C*07:01 (25.1%) was similar in frequency to Cameroonian Baka Pygmies (25%), Southern Irish (21%), Southern Italians (20.6%), and Germans of Essen (20.9%), but higher than Burkina Faso Fulani (17.4%), peoples of North-western England (19%) and Southeastern France (17.3%), Brazilians (15.6%), and Australia New South Wales Caucasians (17.3%). HLA-C*06:02 (25.1%) was higher in frequency than the populations of the USA (2–9%), Burkinabes (5.3%), South-Asian Indians (13.9%), Cameroonians (17%), Czechs (16%), Germans (11%), and Southeastern French (10.6%), but similar to Indians (22.2%), Kenyans (21.7%), Moroccans (23.6%), Pakistanis (21.4%), Saudi Arabians (21.6%), South Africans (20.1%), and Tanzanians (19.2%).

HLA-DRB1*11:01 frequency (11.8%) was similar to Brazil's South Caucasians (11.1%), Chinese (11.7%), Gabonese (13.5%), Georgians (13%), German-Italian minority (10.4%), Greece (11.4%), African-Americans (11.3%), Swedes (11.6%), Indians (12%), Israel-Arabs (10.7%), Italy central (12.8%), but lower than Irani Parsis (21.9%), Lebanese (30.2%), Russian Siberians (16.7%),

and Ukrainians (13.5%). HLA-DRB1*03:01 frequency (11.6%) was similar to Algerians (13.9%), Argentinians (10.9%), Austrians (11.8%), Belgians (15.7%), mixed-race Brazilians (11.6%), Chinese (13.1%), Colombians (15%), Czechs (11%), Croatians (12%), Danes (10.2%), British (12.7%), French (15.2%), German donors (10.6%), Israel Ethiopia Jews (12.5%) but higher in Italy Sardinians (55.7%). DRB1*13:01 (11.4%) was similar in frequency to Cameroonians and central Africans, Georgians, Indians, Iranians, Portuguese, Saudi Arabians, and Spanish (10–15%), but higher in Russian Siberians (23.8%). HLA-DQB1*03:01 (27.5%) was slightly higher in frequency than Algerians, Austrians, Belgians, and French (16–21%), but lower than in Brazilians (60%). HLA-DQB1*05:01 (18.9%) was similar in frequency to Austrians, Belgians, and the people of South France and Europe (12–25%) (Gonzalez-Galarza et al. 2020).

The most prevalent HLA-A~B~DRB1 haplotype in our donors were A*33:01~B*14:02~DRB1*01:02 (2.35%), A*01:01~B*52:01~DRB1*15:01 (2.12%), and A*01:01~B*08:01~DRB1*03:01 (0.89%). None of them included the A*02:01:01 allele unlike the Saudi population (Jawdat et al. 2020). HLA-A*33:01~B*14:02~DRB1*01:02 frequency was similar to Israeli Americans, Ashkenazi Jews, Italians, and Poles (1.7–2.3%), but is higher in Americans, Europeans, Russians, Saudis, and Croatians (41–44%). HLA-A*01:01~B*52:01~DRB1*15:01 was found at a similar frequency in Germans, Turks, and Poles (1.2%), but this haplotype frequency was higher in Israelis, Iraqi Jews (40.3%), and Indians (28.7%) (Gonzalez-Galarza et al. 2020). The haplotype A*01:01~B*08:01~DRB1*03:01 (0.89%) was the third most common HLA haplotype in our study; this haplotype was also seen in Caucasian and Hispanic ethnicities reported in the 16th IHIW, with frequencies of 0.07, 0.01, and 0.10. This haplotype ranked first in Europeans, Hispanics, and Dutch (Askar et al. 2013).

The allele frequencies that we reported in our study are consistent with a previous study on 286 Egyptian donors (Elshakankiry et al. 2017), the most common HLA class I antigens in their study were HLA-A2 (19.2%), A1 (15.8%), B41 (10.9%), and B35 (10.2%). However, they reported different frequencies of HLA-DRB1*13:01 (18.7%), DRB1*04:01 (15%), DRB1*11:01 (14.7%), and DRB1*03:01 (13.2%). These differences might be attributed to the small sample size.

The deviations from the HWE were expected a priori. Consanguinity rates among Egyptians were estimated to be between 29 and 39% in 2012 (Temtamy and Aglan 2012). In 2017, it was estimated to be over 40% depending on the region and urbanization (Ahmed 2017). Inbreeding increases with consanguinity, resulting in higher homozygosity. However, we cannot attribute the observed deviation in HWE solely to inbreeding given the small sample size of HLA-C

and HLA-DQ alleles and the evidence of balancing selection in HLA-B and -DR loci.

We studied the association between HLA antigen groups and alleles with the ALL, AML, and SAA since they were the most prevalent. Most of our findings were trends toward significance, and generally, no remarkable antigen groups or alleles correlated with the odds of ALL or AML. However, only HLA-A2 showed a statistically significant and positive association with SAA. Our findings are consistent with a report by Gluckman et al. where they found an excess of homozygous HLA-A2 in SAA patients compared to healthy controls (Gluckman et al. 1981). In our study population, the frequency of HLA-A2 was 23.3% in donors and 29.6% in SAA patients. HLA-A2 correlated with higher odds of SAA in all the genetic models we conducted. The exact allele that is correlated with SAA could not be identified precisely in our study. However, we think that it is most likely to be HLA-A*02:01 since it showed a strong trend towards higher SAA odds.

We also found unprecedented significant associations between the haplotypes A*33:01~B*14:02~DRB1*11:01 and A*02:01~B*14:02~DRB1*01:02 with SAA. It is noteworthy that HLA-A*33:01, -B*14:02, and -A*02:01 showed strong trends toward increased SAA odds. The noted association with haplotype could be a result of the higher power of haplotype analysis compared to analysis of the individual antigens or alleles. Also, we found an unprecedented association between A*02:01~B*18:01~DRB1*03:01 and the odds of ALL.

The main limitation of this study is that we used PCR-SSO instead of sequencing due to the limited budget provided by the National Insurance for HLA typing. We perform HLA typing by screening patients and their family members for HLA-DRB1. We only proceed to class I (AB) typing for identical patients and donors on class II or for patients capable of affording the cost of a haploidentical transplant, which until now is not fully covered by insurance, and explains why the number of HLA-DRB1 typed samples is more than twice the number of HLA class I (AB) in our study.

Knowledge of population-specific allele and haplotype frequencies increases the chance of finding matched donors. Unfortunately, less than 3% of donors listed in international registries are from oriental origins, which makes it difficult to find a match for a patient with no sibling donor indicated for BMT in Egypt (Besse et al. 2016).

Conclusion

The results of this study present valuable information about HLA genes in Egyptians that can be used for hematopoietic stem cell unrelated donor recruitment, and selection, and as a helpful resource for population genetic studies and HLA disease associations.

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1007/s00251-024-01343-x>.

Author contributions Yasser ElNahass, Nabil M Abdelfattah and Nourhan Mekky formulated the research question, study design, and collected the data. Yasser ElNahass and Nabil M Abdelfattah conducted the statistical analysis and wrote the manuscript. All authors reviewed, edited, and provided guidance throughout all the phases of this study and writing the manuscript.

Funding This research did not receive any specific grant from funding agencies in the public, commercial, or non-profit sectors, and no funding was received to assist with the preparation of this manuscript.

Data availability Data available on request from the authors.

Declarations

Ethical approval All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki Declaration and its later amendments. All patients and donors gave written informed consent and the study was approved by the National Cancer Institutional Review Board number 2301-302-053.

Competing interests The authors have no competing interests to declare that are relevant to the content of this article.

References

- Ahmed SM (2017) Consanguineous marriage among Egyptian youth: secondary analysis of survey of young people in Egypt, 2014. *Egypt J Commun Med* 35(2):85–94. <https://doi.org/10.21608/EJCM.2017.3571>
- Askar M et al (2013) 16(th) IHIW: global distribution of extended HLA haplotypes. *Int J Immunogenet* 40(1):31–38. <https://doi.org/10.1111/IJI.12029>
- Besse K et al (2016) On modeling human leukocyte antigen-identical sibling match probability for allogeneic hematopoietic cell transplantation: estimating the need for an unrelated donor source, biology of blood and marrow transplantation. *J Am Soc Blood Marrow Trans* 22(3):410–417. <https://doi.org/10.1016/J.BBMT.2015.09.012>
- Chang YJ, Zhao XY, Huang XJ (2021) Haploidentical stem cell transplantation for acute myeloid leukemia: current therapies challenges and future prospective. *Front Oncol*. <https://doi.org/10.3389/FONC.2021.758512>
- Creary LE et al (2019) Next-generation sequencing reveals new information about HLA allele and haplotype diversity in a large European American population. *Hum Immunol* 80(10):807–822. <https://doi.org/10.1016/J.HUMIMM.2019.07.275>
- Elshakankiry NH et al (2017) P227 determination of HLA -A, -B and -DRB1 alleles and HLA-A -B haplotype frequencies in Egyptians based on family study. *Hum Immunol* 78:222. <https://doi.org/10.1016/J.HUMIMM.2017.06.287>
- Gfeller D, Bassani-Sternberg M (2018) Predicting antigen presentation—what could we learn from a million peptides? *Front Immunol* 9:1716. <https://doi.org/10.3389/FIMMU.2018.01716/BIBTEX>
- Gluckman E et al (1981) HLA markers in patients suffering from aplastic anaemia. *Haematologia* 14(2):165–172
- Gonzalez-Galarza FF et al (2020) Allele frequency net database (AFND) 2020 update: gold-standard data classification, open

- access genotype data and new query tools. *Nucleic Acids Res* 48(D1):D783–D788. <https://doi.org/10.1093/NAR/GKZ1029>
- Goopu M et al (2021) HLA-haploidentical vs matched unrelated donor transplants with posttransplant cyclophosphamide-based prophylaxis. *Blood* 138(3):273–282. <https://doi.org/10.1182/BLOOD.2021011281>
- Grager L et al (2014) Fine-mapping of HLA associations with chronic lymphocytic leukemia in US populations. *Blood* 124(17):2657–2665. <https://doi.org/10.1182/BLOOD-2014-02-558767>
- Guo SW, Thompson EA (1992) Performing the exact test of Hardy-Weinberg proportion for multiple alleles. *Biometrics* 48(2):361. <https://doi.org/10.2307/2532296>
- Hammer C, Migdal M (2022) midasHLA: R package for immunogenomics data handling and association analysis
- Huang X et al (2012) HLA associations in classical Hodgkin lymphoma: EBV status matters. *PLoS ONE* 7(7):39986. <https://doi.org/10.1371/JOURNAL.PONE.0039986>
- Jawdat D et al (2014) HLA-B50 polymorphism in the Saudi population. *Int J Immunogenet* 41(2):95–97. <https://doi.org/10.1111/IJI.12096>
- Jawdat D et al (2020) HLA-A, -B, -C, -DRB1, -DQB1, and -DPB1 Allele and Haplotype Frequencies of 28,927 Saudi Stem Cell Donors Typed by Next-Generation Sequencing. *Front Immunol*. <https://doi.org/10.3389/FIMMU.2020.544768>
- Kazemi MH et al (2021) Association of HLA alleles with hematologic malignancies. *Gene Rep* 25:101346. <https://doi.org/10.1016/J.GENREP.2021.101346>
- Kolde R (2019) pheatmap: pretty heatmaps
- Lancaster AK et al (2007) PyPop update – a software pipeline for large-scale multilocus population genomics. *Tissue Antigens* 69(SUPPL. 1):192–197. <https://doi.org/10.1111/J.1399-0039.2006.00769.X>
- Mahmoud HK et al (2008) Hematopoietic stem cell transplantation in Egypt. *Bone Marrow Transplant*. <https://doi.org/10.1038/BMT.2008.136>
- Mahmoud HK et al (2020) Hematopoietic stem cell transplantation in Egypt: challenges and opportunities. *Mediterr J Hematol Infect Dis* 12(1):2020023. <https://doi.org/10.4084/MJHID.2020.023>
- Park M, Seo JJ (2012) Role of HLA in hematopoietic stem cell transplantation. *Bone Marrow Res* 2012:1–7. <https://doi.org/10.1155/2012/680841>
- Petersdorf EW (2017) Which factors influence the development of GVHD in HLA-matched or mismatched transplants? *Best Pract Res Clin Haematol* 30(4):333–335. <https://doi.org/10.1016/J.BEHA.2017.09.003>
- R Core Team (2022) R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>
- Sanchez-Mazas A et al (2017) Common and well-documented HLA alleles over all of Europe and within European sub-regions: a catalogue from the European Federation for Immunogenetics. *HLA* 89(2):104–113. <https://doi.org/10.1111/TAN.12956>
- Simpson E, Dazzi F (2019) Bone marrow transplantation 1957–2019. *Front Immunol*. <https://doi.org/10.3389/FIMMU.2019.01246>
- Slatkin M (1994) An exact test for neutrality based on the Ewens sampling distribution. *Genet Res* 64(1):71–74. <https://doi.org/10.1017/S0016672300032560>
- Temtamy S, Aglan M (2012) Consanguinity and genetic disorders in Egypt. *Middle East J Med Genet* 1(1):12–17. <https://doi.org/10.1097/01.MXE.0000407744.14663.D8>
- Van Rossum G, Drake FL (2009) Python 3 reference manual. CreateSpace, Scotts Valley, CA
- Venables WN, Ripley BD (2002) Modern applied statistics with S. Fourth. Springer, New York
- Villemagne B et al (2005) Two new cases of familial hairy cell leukemia associated with HLA haplotypes A2, B7, Bw4, Bw6. *Leuk Lymphoma* 46(2):243–245. <https://doi.org/10.1080/10428190400013589>
- Zaimoku Y et al (2017) Identification of an HLA class I allele closely involved in the autoantigen presentation in acquired aplastic anemia. *Blood* 129(21):2908–2916. <https://doi.org/10.1182/BLOOD-2016-11-752378>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.