Genetic similarity of chromosome 6 between patients receiving hematopoietic stem cell transplantation and HLA matched sibling donors

Hannu Turpeinen,[1] Liisa Volin,[2] Lauri Nikkinen,[3] Pauli Ojala,[1] Aarno Palotie,[4-8] Janna Saarela,[5,9] and Jukka Partanen[1]

[1]Research and Development, Finnish Red Cross Blood Service, Helsinki, Finland; [2]Department of Medicine, Helsinki University Central Hospital, Helsinki, Finland; [3]Blood Component Expertise, Finnish Red Cross Blood Service, Helsinki, Finland; [4]Department of Clinical Chemistry, University of Helsinki, Finland; [5]The Finnish Institute for Molecular Medicine (FIMM); [6]Finnish Genome Center, University of Helsinki, Finland; [7]The Wellcome Trust Sanger Institute, Hinxton, Cambridge, UK; [8]The Broad Institute of MIT and Harvard, Cambridge, MA, USA, and [9]Department of Molecular Medicine, National Public Health Institute, Helsinki, Finland

ABSTRACT

Background

Matching for HLA genes located on chromosome 6 is required in hematopoietic stem cell transplantation to reduce the incidence of graft-versus-host disease. However, a considerable proportion of patients still suffer from it, obviously due to genetic differences outside the HLA gene region.

Design and Methods

We studied the similarity of almost 4,000 single nucleotide polymorphisms on chromosome 6 between patients receiving hematopoietic stem cell transplantation and their HLA-matched sibling donors.

Results

We observed that as a result of routine HLA matching the siblings in fact shared surprisingly long chromosomal fragments with similar single nucleotide polymorphism genotypes – from 11.65 Mb to 134.66 Mb. The number of genes mapped on these shared fragments varied from 402 to 1,302. Considering the whole chromosome 6, the HLA-matched siblings were apparently identical for 65.2-97.8% of the single nucleotide polymorphisms.

Conclusions

Potentially, genes similar in some transplantation pairs while different in others might have a significant role in determining the outcome after hematopoietic stem cell transplantation.

Key words: hematopoietic stem cell transplantation, graft-versus-host disease, chromosome 6, histocompatibility, genetic similarity.

Introduction

In allogeneic hematopoietic stem cell transplantation (HSCT), matching for HLA genes on chromosome 6 between patient and donor is regarded as an important prerequisite for good clinical success. Therefore all patients and donor candidates are routinely genotyped for HLA alleles. Mismatches in HLA are seldom accepted if a genetically similar donor is available. However, even patients with perfectly HLA-matched donors suffer regularly from life-threatening graft-versus-host disease (GvHD). This can be explained to a large extent by genetic polymorphisms in genes locating outside the HLA. Several good candidates, including e.g. genes for minor histocompatibility antigens, have been proposed to be responsible for this non-HLA effect in GvHD.[1,2]

Related persons share more genome than unrelated persons. This fact favors transplantations from sibling donors over a registry-based, unrelated donor. Sibling donor candidates are routinely matched at least for *HLA-A, -B* and *-DRB1* genes, while registry donor candidates need to be genotyped for a denser set of HLA genes. Based on the relatively small expectation of recombination events within the HLA region in any specific meiosis, it can be assumed that siblings who share identical *HLA-A, -B* and *-DRB1* genes, have actually also inherited other HLA genes identical-by-descent (IBD). Other chromosomes behave independently of matched HLA region and on them siblings share on average 50% of genetic IBD material. However, it is of note that this 50% sharing is only an average and siblings quite regularly share either none or all DNA IBD on smaller chromosomes.[3]

Organization of physically closely linked genetic variants into haploblocks is a hallmark for human genome. Genetic material within these haploblocks is inherited jointly as a unit and the blocks are broken off by recombination events.[4] Linkage disequilibrium is known to be relatively strong in the HLA region. It has been estimated that the recombination frequency for the HLA region would be only <40% of that averaged in other parts of the genome.[5] Therefore, in transplantations from a sibling donor, patient and donor matched for the major HLA genes can be assumed to share the entire HLA region.

The overall genetic similarity of large genomic regions on any chromosomes between HSCT patient and donor has scarcely been studied in a systematic manner. Kikuchi *et al*. have reported an association of some markers with acute GvHD after systematic genotyping of 155 microsatellite markers on chromosome 22 but did not address the overall genetic similarity between recipients and donors.[6] Malkki *et al*. however, have reported that unrelated, HLA-matched transplantation donors differ significantly from their recipients in microsatellite alleles locating next to the matched HLA genes.[7] Average matching between patients and donors ranged from 11% to 85%.

To address the question of overall genetic similarity of chromosome 6 between HSCT patients and their HLA-matched sibling donors we analyzed approximately 4,000 single nucleotide polymorphism (SNP) markers mapping along the whole chromosome 6. We show that the identical fragments siblings share differ greatly in length and position. Therefore, also the number of identical genes between patient and donor vary among transplantation pairs. Potentially, these genes identical in some pairs while different in others could partly explain different histocompatibility and risk of GvHD or relapse after HLA matched HSCT between siblings.

Design and Methods

Hematopoietic stem cell transplantation patients and donors

Thirty-three sibling pairs (33 recipients, 33 donors) who had undergone HSCT were included in the analyses of genetic similarity. All transplantations were performed in a single center (Division of Hematology, Department of Medicine, Helsinki University Central Hospital, Helsinki, Finland) between 1993 and 2004. One pair had one allele mismatch [towards both directions, e.g. towards graft-versus-host (GvH) and host-versus-graft (HvG)] at HLA-B locus; all the other patients and their donors were matched for *HLA-A, -B, -C* and *-DRB1* genes. In addition, most of the pairs were also matched for *HLA-DPB1* gene. One pair was identified to be a monozygotic twin. The detailed patient clinical characteristics are described in Table 1 and the *Online Supplementary Table S1*.

In addition to these 33 sibling pairs used in genetic similarity analyses, 7 new patients (altogether 33+7=40 patients) and 8 new donors (altogether 33+8=41 donors) were included in SNP genotype association analyses. These extra samples originated from transplantation pairs in which genotyping for either patient or donor did not meet the quality control limits set for the study and were therefore excluded from further analyses resulting thus solitary samples with no sibling pair. In SNP genotype association analyses, pairs with patient having grade III-IV acute GvHD or limited/extensive chronic GvHD were regarded as an acute/chronic GvHD positive group respectively.[8] Their controls were the pairs with patient having no acute GvHD or no chronic GvHD, respectively. χ^2 and Fisher's tests were used to analyze the statistical significance between groups. Altogether, 19 patients had acute GvHD, while 13 patients suffered from chronic GvHD and 14 patients had neither form of the disease.

The study was approved by the Ethical Review Board of the Helsinki University Hospital, Helsinki, Finland.

Single nucleotide polymorphisms genotyped and definition for similarity

Altogether more than 50,000 SNPs were genotyped for all the samples. We used the GeneChip Human Mapping 50K array (Xba) (Affymetrix, Santa Clara, California, USA) and followed manufacturer's instructions for genotyping. Altogether there are 3,962 SNPs on chromosome 6; this study concentrated only on these SNPs on chromosome 6 and SNPs on other chromosomes were used only for illustrative purposes. Genotyping of 9/3,962 (0.2%) SNPs failed in all samples and of 39/3,962 (1.0%) SNPs failed in more than 10% (7/66 samples) of the samples. Two hundred and forty-

one (6.1%) SNPs were monomorphic and a vast majority of SNPs (2,838/3,962, 71.6%) had a proper minor allele frequency of 0.1-0.5.

As assumed for adult HSCT material, no parental genotypes were available for the study. We therefore studied genetic similarity merely as identity-by-state (IBS). This approach considers genotypic differences between patient and his/her donor and ignores the fact that similar-looking genotypes may actually originate from different parental chromosomes, which is respected in the case of IBD. For classification of genetic similarity, the limit of two SNPs with different genotypes within the sliding window of 20 SNPs was used, when the starting and ending points of similarity regions were examined. A single sporadic SNP with different genotype between siblings may easily originate from genotyping error or sporadic mutation and not actually represent divergent inherited chromosomal regions.

Table 1. The observed regions of IBS on chromosome 6 in 33 HSCT sibling pairs.

	GvHD		Bordering SNP				IBS fragment size, Mb	of marker area	# genes in IBS fragment	Overall SNP similarity	IBS fragments combined, Mb	Antigen presentation related GeneGo objects	p value for enrichment of antigen presenting GeneGo objects
Pair	Acute	Chronic	Start	End	Start, bp	End, bp							
1[a]	No	No	RS1418708	RS1474642	150610	170770193	170.62	1.00	1580	0.99	170.62	27 / 197	2.75E-07
2	No	No	RS9296488	RS1410676	12957290	72529170	59.57	0.35	789	0.98	164.64	27 / 197	1.95E-07
3	No	NA	RS9295612	RS827978	24118243	158773875	134.66	0.79	1302	0.94	143.76	27 / 197	1.88E-08
4	No	Lim.	RS6927090	RS7768389	197145	44271508	44.07	0.26	747	0.91	126.76	24 / 197	8.40E-08
5	3	Lim.	RS644532	RS976482	10579023	40549850	29.97	0.18	567	0.90	122.26	22 / 197	3.61E-07
6	No	No	RS10484380	RS2246086	19142390	102472246	83.33	0.49	899	0.86	109.88	20 / 197	9.86E-06
7	3	Ext.	RS434509	RS502156	4506829	53578630	49.07	0.29	786	0.85	97.02	27 / 197	9.74E-06
8[b]	No	No	RS707937	RS1934593	31838993	106666295	74.83	0.44	620	0.84	96.07	13 / 197	1.62E-03
9	No	No	RS9295589	RS7768389	23553309	44271508	20.72	0.12	562	0.83	91.60	24 / 197	4.22E-08
10	3	NA	RS9328381	RS648549	6618518	88730020	82.11	0.48	935	0.83	93.55	21 / 197	6.06E-07
11	No	Ext.	RS5008746	RS7739888	3835729	44632151	40.80	0.24	713	0.83	98.92	20 / 197	1.95E-06
12	No	Lim.	RS3757092	RS9320409	2713210	97637567	94.92	0.56	1024	0.82	102.67	20 / 197	4.52E-06
13	4	NA	RS6922313	RS2207022	24546698	104005966	79.46	0.47	873	0.82	92.77	20 / 197	3.11E-06
14	No	No	RS1321055	RS10498894	12372985	74507950	62.13	0.36	811	0.81	75.87	20 / 197	3.53E-06
15	No	No	RS2817719	RS898893	25003866	107089425	82.09	0.48	886	0.80	96.70	20 / 197	3.36E-06
16	No	No	RS2754777	RS2295884	24898441	39976014	15.08	0.09	450	0.80	71.50	25 / 197	2.99E-08
17	3	Ext.	RS6904345	RS1266922	24805403	51930256	27.12	0.16	608	0.79	71.17	24 / 197	4.69E-10
18	3	Ext.	RS4712580	RS10484878	21173805	52344773	31.17	0.18	634	0.77	62.66	21 / 197	1.41E-08
19	No	ext.	RS2483766	RS2076472	20234635	41137320	20.90	0.12	486	0.77	62.73	17 / 197	1.54E-06
20	No	No	RS548987	RS718233	25977350	49692418	23.72	0.14	573	0.76	73.80	26 / 197	1.92E-11
21	4	NA	RS952578	RS6919087	21992498	37807134	15.81	0.09	448	0.76	63.40	19 / 197	1.44E-07
22	No	No	RS6917813	RS10498844	24898998	66814023	41.92	0.25	677	0.76	58.02	20 / 197	1.49E-07
23	4	NA	RS1322535	RS10484622	19147265	62737005	43.59	0.26	694	0.75	60.91	20 / 197	3.21E-07
24	3	Lim.	RS234420	RS1150870	4568811	53445640	48.88	0.29	784	0.75	78.92	20 / 197	3.21E-07
25	No	No	RS9296195	RS10484844	11998048	46994911	35.00	0.21	665	0.75	69.97	26 / 197	2.82E-11
26	No	Ext.	RS10498681	RS3798477	12916270	37059039	24.14	0.14	501	0.75	46.58	22 / 197	1.67E-11
27	4	NA	RS3757092	RS2237093	2713210	36143332	33.43	0.20	570	0.75	74.78	23 / 197	2.01E-10
28	No	No	RS9328514	RS919875	896816	40519678	39.62	0.23	650	0.74	69.40	18 / 197	1.83E-0
29	3	No	RS10484451	RS10498756	11425301	44046909	32.62	0.19	639	0.72	68.05	18 / 197	1.83E-07
30	3	Lim.	RS2013807	RS10498754	14196281	43264138	29.07	0.17	593	0.70	50.97	18 / 197	1.64E-07
31	3	No	RS6922313	RS7757672	24546698	36193303	11.65	0.07	402	0.69	38.85	19 / 197	2.12E-09
32	No	No	RS1590702	RS9296158	20395075	35675060	15.28	0.09	410	0.67	42.23	23 / 197	2.53E-12
33	4	Ext.	RS1418708[c]	RS7764126	150610	34897040	34.75	0.20	569	0.65	35.77	17 / 197	1.08E-08
All pairs			RS3887266	RS2296360	25951725	34911536	8.96	0.05	354				

Similarity around HLA — columns: Bordering SNP (Start, End, Start bp, End bp), IBS fragment size Mb, of marker area, # genes in IBS fragment.
Similarity along whole chromosome 6 — columns: Overall SNP similarity, IBS fragments combined Mb, Antigen presentation related GeneGo objects, p value for enrichment of antigen presenting GeneGo objects.

[a]Identical twins; [b]HLA-B mismatch; [c]HLA IBS fragment from the beginning of marker area; Lim.: limited; Ext.: extensive.
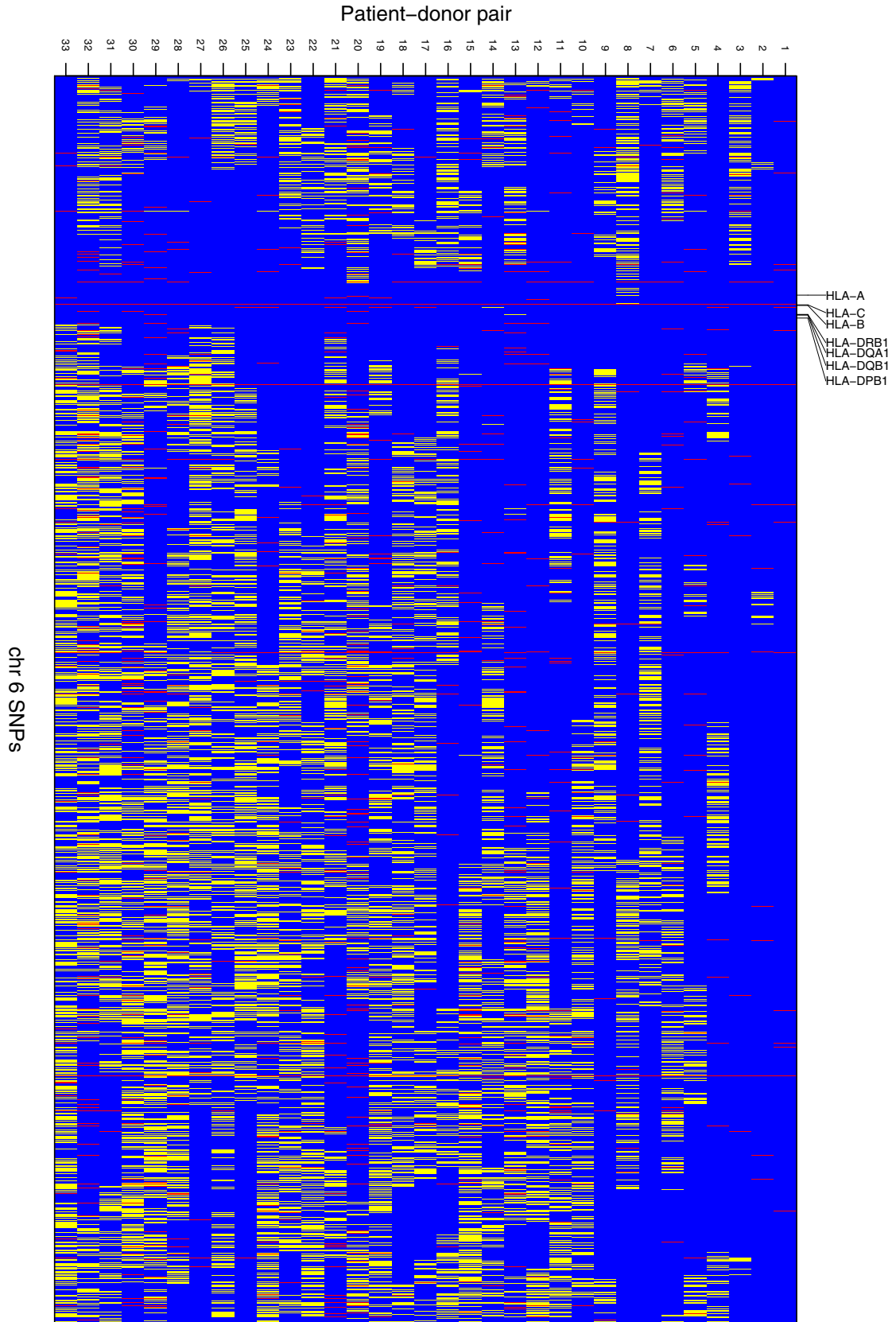
Patient–donor pair

Figure 1. The SNPs on chromosome 6 IBS between HSCT sibling pairs. Blue stripes represent similar genotypes and yellow stripes different genotypes between siblings. Red stripes are SNPs where genotyping of either patient or donor was not successful and thus no similarity could be determined. Note: Markers in Figure 1 are not in physical scale; i.e. each blue, yellow or red *stripe* represents one SNP, and the distance between neighbouring SNPs differs along the chromosome.

Results

Altogether 3,962 SNPs on chromosome 6 were analyzed in 33 sibling pairs who underwent hematopoietic stem cell transplantation. The similarity of patient and donor genotypes, as well as the association of patients' and donors' genotypes with GvHD incidence was inspected.

To obtain an overall picture of our sibling data we first calculated the similarity of genotypes for all autosomes in siblings. All genotypic differences between a recipient and a donor were counted as difference; no distinction was made between 1 or 2 allele differences or differences toward GvH or HvG direction. The observed average similarity per chromosome varied between 71.4% (for chromosome 12) and 79.8% (for chromosome 6). After HLA matching the average overall similarity of genotypes on chromosome 6 was naturally higher than that on other chromosomes and varied in siblings between 65.2% and 97.8% (identical twin pair excluded).

We then addressed specifically for SNPs on chromosome 6 the questions: (i) how long continuous chromosomal blocks were actually matched on chromosome 6 as a result of standard HLA matching, (ii) which genes were located on these blocks and (iii) whether the chromosomal length shared had any effect on the outcome after HSCT. The total length of chromosome 6 is approximately 170.90 Mb and the distance between the 2 most distal SNPs in our study set is approximately 170.62 Mb. Although the HLA matching encompassed only approximately 3.1 Mb (from HLA-A to HLA-DPB1), the actual length of chromosomal matching around the HLA observed with the present high-density SNP mapping was revealed to vary from 11.65 Mb (6.8% of the entire chromosome 6) to 134.66 Mb (78.8% of the entire length of chromosome 6). On average of all 33 pairs, the length for the matched region was 45.67 Mb, i.e. 26.7% of the entire chromosome 6. This observation clearly indicates that the current HLA matching protocol for sibling donors results in significantly longer chromosomal matching than merely the HLA region. The length of the common chromosomal fragment around HLA which was shared IBS by all 33 siblings pairs was 8.96 Mb long, in other words 5.2% of the entire chromosome 6 (Table 1, Figure 1). The identical twin pair, as assumed, was identical for the total length of the 170.62 Mb marker fragment.

To illustrate the effect of HLA matching on the observed patient-donor chromosomal similarity we made a similar analysis for chromosomes 5 and 7, which are both of substantially similar size to chromosome 6 and therefore subjected to an approximately similar level of recombination in meiosis. Relatively long regions (maximally >76 Mb in chr 5 and >87 Mb in chr 7) with no genotypic differences between siblings were also observed in the present data. However, contrary to chromosome 6, these regions were not centralized around any specific location on the chromosomes but were relatively evenly distributed. Furthermore, it is of note that some siblings only shared very short fragments IBS in these chromosomes (*Online Supplementary Figures S1A and S1B*).

To obtain information on whole chromosome 6 similarity we also summed up all long identical (excluding limit of 2 SNPs with different genotypes within the sliding window of 20 consecutive SNPs) chromosomal fragments along the whole chromosome 6. The sums varied between 35.8 Mb (20.9% of chromosome 6) and 164.6 Mb (96.3% of chromosome 6) (excluding the identical twin pair). The average identity of 81.6 Mb was observed. As chromosome 6 is known to harbor the most important genetic region in the human genome with respect to (auto)immunity, we also analyzed the enrichment of functional biological processes in these identical chromosomal fragments. This was performed against the manually curated systems biology, systems medicine and chemoinformatics platforms (*www.genego.com*)[9] using three parallel and compensating gene association conventions as an input (HGNC, Swissprot and Entrez). As a result, it was verified in the GeneGo processes database that out of >2,000 established classes, the most enriched functional annotation for all sibling pairs except one was 'Immune: antigen presentation'. Genome-widely, 197 GeneGo network objects are included in this category out of which 27 have been mapped to chromosome 6. Based on the identity of chromosome 6, the sibling pairs shared 13-27 (median 20) identical GeneGo's antigen presentation network objects in this crucial class (Table 1).

Altogether 1,580 genes have been mapped on the chromosome 6 region bordered by the distal SNPs of our study set (*www.ensembl.org/biomart*). The patient-donor pair with the longest consecutive shared fragment around the HLA genes, that of 134.66 Mb, potentially shared 1,302 of these genes identical. This number was 402 genes for the pair with the shortest shared fragment as compared to 128 genes which locate between the matched *HLA-A* and *HLA-DRB1* genes. Table 1 shows the number of shared genes within these regions in each pair. There were altogether 70 genes (out of the 1,302 which located in the longest consecutive region of IBS) which could be designated to Gene Ontology (GO) process of 'immune system process' (GO:0002376). Based on our SNP analyses, 45 of them were also shared IBS by the pair with the shortest consecutive stretch of IBS around the HLA region. This leaves 25 immune related genes locating in the region shared IBS by one pair but not by another one. These genes are listed in the *Online Supplementary Table S2*.

In addition to genetic similarity analyses, we analyzed the traditional case-control association of all SNPs on chromosome 6 with acute and chronic GvHD. We compared separately the distribution of patient and donor genotypes between GvHD cases and controls. In addition, we analyzed the association of SNPs that differ toward GvH (i.e. patient has an allele that donor is missing) between patients and donors. Altogether 850 SNPs (out of 23,772 tests performed) resulted in a nominal association of $p<0.05$. However, none of the individual results remained significant after correction for multiple testing. The *Online Supplementary Table S3* lists all the nominally statistically significant SNPs with p and OR

values.

Regardless of being aware that recipients and donors are related and thus not fully independent genetically we still pointed out the SNPs that resulted in significant association with GvHD in more than one analyzed type. Altogether 34 and 39 SNPs conferred a statistically significant association with acute and chronic GvHD respectively, in both recipient and donor genotypes. Seven of these were statistically significantly associated with both acute and chronic GvHD in both recipients and donors (*Online Supplementary Table S3*).

We also studied the possible association between the number of SNPs IBS between patient and his/her donor and the outcome after transplantation. Association of chromosome 6 overall similarity (% of SNPs identical between siblings and combined length of long IBS regions) and the length of genomic region around the HLA genes with total similarity between patient and his/her donor were analyzed against acute and chronic GvHD. The non-parametric Mann-Whitney test was used. No statistically significant result was observed (*details not shown*).

The Finnish population is often regarded to be genetically more homogeneous than most other populations.[10,11] However, recent findings suggest that although certain subisolates in Finland are very homogeneic these isolates might differ from each other considerably reducing the overall homogeneity among Finns.[12] To analyze the effect of possible population homogeneity on our observations of chromosome 6 similarities between siblings we also analyzed the similarity of SNPs between all these 33 unrelated donor-donor pairs. Thirty-three donors make altogether 528 different donor-donor pairs. These donor-donor pairs demonstrated three facts: i) most of the SNPs were similar even between 2 totally unrelated Finnish individuals, ii) regardless of this the lengths of IBS regions with consecutive similar genotypes were rather short and iii) no accumulation of any kind for similarity blocks was observed in or around the HLA region. The similarity rates of SNPs on chromosome 6 for donor-donor pairs varied between 0.506 and 0.592. The longest region with consecutive similar SNPs consisted of only 70 SNPs representing ~5.5 Mb region of IBS. For illustrative purposes, an example picture of genotypic similarities of chromosome 6 markers for 33 randomly selected donor-donor pairs is provided in the *Online Supplementary Figure S2*.

Discussion

By genotyping almost 4,000 SNPs on chromosome 6 in HSCT pairs with related sibling donors we have shown that regardless of HLA identity in all pairs the overall genetic similarity siblings share varies notably. The consecutive region of similar genotypes varied from 11.65 Mb to 134.66 Mb around the HLA genes and overall similarity of all markers on chromosome 6 between 65.2% and 97.8% among siblings. These findings clarify the objective of genetic matching between any 2 individuals for transplantation. As no definitive conclusion of non-HLA genes important in GvHD predisposition has been made to date, the data presented in this study was aimed at exemplifying the magnitude of overall genetic difference in non-HLA genes between any 2 siblings undergoing HSCT.

The average rate for any genetic marker across the genome to be IBD is 50% for siblings with the same parents. This figure is valid only on a genome-wide level and is dependent on recombination rates in any specific chromosomal region. In the present study the observed similarities between siblings were well above 50% for all the chromosomes. This results from the fact that as no parental genotypes were available, the definition for similarity used in this study was not based on IBD, but rather on IBS. We therefore counted similar-looking patient and donor genotypes as identical regardless of the fact that they may have originated from different, but similar-looking parental chromosomes. Thus, homozygosity of a parent at any study locus increases the similarity rate observed in siblings.

Petersdorf *et al.* have recently shown that patients who have received a hematopoietic stem cell graft from an HLA haplotype mismatched donor suffer statistically significantly more often from severe acute GvHD than those patients who have received a graft of HLA haplotype match, even though all patient-donor pairs in both groups were matched for HLA genes (*HLA-A, -B, -C, -DRB1, -DQB1*).[13] They concluded that the difference in GvHD risk results therefore from additional genes mapping within the MHC region and differing between patients and donors with different HLA haplotypes. Furthermore they proposed that, if replicable, the finding would warrant efforts to identify haplotype-matched donors in donor selection in order to decrease the GvHD risk. Our findings broaden the significance of results by Petersdorf *et al*. Even though we did not aim at finding the risk genes or SNPs for the GvHD predisposition, and our sample material was limited in size, we could show that the length of fragments siblings share IBS varies significantly. We see no particular reason to restrict the location of currently still unidentified but important histocompatibility antigens within the MHC region itself. We are confident that within the extended region of >130 Mb containing >1,300 genes there are several good candidate genes with transplantation related immunological events.

Also Baron *et al*. recently reported interesting results on genome-wide significance of donor genotypes and how donor gene expression profiles predict the incidence of GvHD in patients.[14] They could show that expression profiles of certain donor genes in CD4[+] and CD8[+] T cells have a dominant influence on the incidence of both acute GvHD and chronic GvHD. Interestingly, these profiles persisted long-term post-transplantation in the patients indicating that genetic factors rather than environmental ones, like new patient host, regulate the level of transcription. Five of the genes (*CD24, HLA-DRB3, STK38, VEGF* and *HDAC2*) differentially expressed in GvHD positive and negative patients mapped to chromosome 6. Unfortunately, our SNP coverage for these genes was rather poor and only one SNP mapped within these genes (rs200774 in STK38). This SNP did not confer any predisposition to acute or chronic GvHD in our data set.

Although Baron *et al.* correlated the expression profiles of donor genes (pre- and post-transplantation) with GvHD risk, it could well be hypothesized that not only the donor or patient expression profiles alone but rather the combination of them would be critical in disease predisposition. It is possible that shifting expression profiles of immune related genes would not be optimal for a patient's immune system reconstituting after HSCT. We have shown that some transplantation pairs share significantly longer haplotypes in common than others do. It could be hypothesized that those pairs sharing longer fragments IBD would also share similar expression profiles for a greater number of transplantation related genes and would therefore predispose the patient's immune system to more original-like expression profiles. To support this hypothesis, it has been reported that patients who have received a renal graft from a donor with differential expression level of transforming growth factor β suffer more often from acute rejection.[15] This hypothesis of potential harmful effects for recipient's immune system after shifting (genome-wide) expression profiles after transplantations remains to be evaluated in future studies with systematic comparisons of patient and donor expression profiles for transplantation related genes.

A mean gene length with introns for non-pseudogenes present on human chromosome 6 is estimated to be ~45000 bp.[16] This figure can be used to estimate the sufficiency of 3,962 markers used in the present study to cover the chromosome 6. On average, this means that in the present study there is only one SNP per gene (distance between two neighboring SNPs in this study: mean 43075 bp, minimum 17 bp, maximum 3355565 bp). However, on average, it has been estimated that there is 0.0111 recombination events per Mb per meiosis on chromosome 6 and only 0.0044 recombination events per Mb per meiosis across the 7.5Mb extended HLA region[5] resulting on average only <2 recombination events per chromosome 6 per meiosis. Thus, most likely we have not missed any (double) recombinations with our marker set and consecutive stretches of IBS observed represent very accurately the real IBD situation and the location and length of IBD regions.

An explanation for single SNPs differing in genotype between patient and donor and locating within the relative long consecutive segment of otherwise similar genotypes (single *yellow stripes* in the middle of long *blue segments* in Figure 1) is problematic. We figured out three possible explanations: i) errors in genotyping, ii) new mutation and iii) double-recombination. Although the genotyping accuracy seemed to be rather high in this study (identical twins had only 27 out of all 55,720 SNPs tested (0.048%) on all chromosomes *differing* from each other and 1.11% of all 261,492 genotypes (3,962 SNPs x 33 recipient-donor pairs) on chromosome 6 failed) most likely some of these inconsistent SNPs result from genotyping error in either patient or sibling donor. The mutation rate for humans is estimated to be some 2.5×10^{-8} per generation and per nucleotide locus resulting thus <100 new mutation loci where a child's genome differ from

parental ones.[17] This could explain some of the observed differences. In theory, double-recombination, where a very short segment of DNA factually originates from different parental chromosome in the middle of an otherwise long segment of DNA inherited IBD could also explain the phenomenon observed. All these explanations are relevant in theory, but the number of these *inconsistent* SNPs (~30 SNPs on chromosome 6 among all 33 sibling pairs) is so limited that the biological meaning of these in relation to transplantation histocompatibility is doubtful.

It is of note that more than half of the SNPs genotyped were similar even between 2 randomly chosen, unrelated donors. Although the Finns have regularly been reported to be more homogenous than many other populations, the similarity rate of SNPs between unrelated persons in this study is actually comparable with the other published data. In Phase II HapMap, 3,893 SNPs (out of 3,962 SNPs, 98.3%) studied in this study were also genotyped. The IBS rates for 60 unrelated individuals with European derived ancestry (CEPH samples) were between 52% and 63%.[18]

Although the sample material in the present study was small for a proper case-control association study, we wanted to analyze the association of all the SNPs on chromosome 6 with acute and chronic GvHD in order to test whether promising candidates for further studies could be revealed. We did observe several nominally statistically significant SNPs predisposing to or protecting from GvHD. However, because of the small sample size, it is not possible to conclude the real significance of these results.

We have shown that HLA matched sibling pairs undergoing HSCT vary considerably, when the overall similarity of genetic markers on chromosome 6 is considered. Some pairs showed similarity for more than 90% of the markers, while others were similar only for 65%. The classical HLA matching in HSCT has been shown to actually extend substantially longer chromosomal fragments than the HLA region itself. As the GvHD predisposition in HLA matched transplantation cases is dictated mainly by genetic markers outside the HLA region and as no definite identification of these variants have been made to date, we believe that the data provided in this report would be useful when overall predisposition to GvHD is considered.

The huge potential of new whole genome SNP association and sequencing approaches in histocompatibility testings should be considered in the future.

Authorship and Disclosures

References

1. Dickinson AM. Risk assessment in haematopoietic stem cell transplantation: pre-transplant patient and donor factors: non-HLA genetics. Best Pract Res Clin Haematol 2007; 20:189-207.

2. Hambach L, Spierings E, Goulmy E. Risk assessment in haematopoietic stem cell transplantation: minor histocompatibility antigens. Best Pract Res Clin Haematol 2007;20:171-87.

3. Gagnon A, Beise J, Vaupel JW. Genome-wide identity-by-descent sharing among CEPH siblings. Genet Epidemiol 2005;29:215-24.

4. Gabriel SB, Schaffner SF, Nguyen H, Moore JM, Roy J, Blumenstiel B, et al. The structure of haplotype blocks in the human genome. Science 2002; 296:2225-9.

5. de Bakker PI, McVean G, Sabeti PC, Miretti MM, Green T, Marchini J, et al. A high-resolution HLA and SNP haplotype map for disease association studies in the extended human MHC. Nat Genet 2006;38:1166-72.

6. Kikuchi T, Naruse TK, Onizuka M, Li S, Kimura T, Oka A, et al. Mapping of susceptibility and protective loci for acute GVHD in unrelated HLA-matched bone marrow transplantation donors and recipients using 155 microsatellite markers on chromosome 22. Immunogenetics 2007;59:99-108.

7. Malkki M, Single R, Carrington M, Thomson G, Petersdorf E. MHC microsatellite diversity and linkage disequilibrium among common HLA-A, HLA-B, DRB1 haplotypes: implications for unrelated donor hematopoietic transplantation and disease association studies. Tissue Antigens 2005;66:114-24.

8. Przepiorka D, Weisdorf D, Martin P, Klingemann HG, Beatty P, Hows J, et al. 1994 Consensus Conference on Acute GVHD Grading. Bone Marrow Transplant 1995;15:825-8.

9. Nikolsky Y, Ekins S, Nikolskaya T, Bugrim A. A novel method for generation of signature networks as biomarkers from complex high throughput data. Toxicol Lett 2005;158:20-9.

10. Peltonen L, Jalanko A, Varilo T. Molecular genetics of the Finnish disease heritage. Hum Mol Genet 1999; 8:1913-23.

11. Varilo T, Peltonen L. Isolates and their potential use in complex gene mapping efforts. Curr Opin Genet Dev 2004;14:316-23.

12. Service S, DeYoung J, Karayiorgou M, Roos JL, Pretorious H, Bedoya G, et al. Magnitude and distribution of linkage disequilibrium in population isolates and implications for genome-wide association studies. Nat Genet 2006;38:556-60.

13. Petersdorf EW, Malkki M, Gooley TA, Martin PJ, Guo Z. MHC haplotype matching for unrelated hematopoietic cell transplantation. PLoS Med 2007;4:e8.

14. Baron C, Somogyi R, Greller LD, Rineau V, Wilkinson P, Cho CR, et al. Prediction of graft-versus-host disease in humans by donor gene-expression profiling. PLoS Med 2007;4:e23.

15. Lacha J, Hribova P, Kotsch K, Brabcova I, Bartosova K, Volk HD, et al. Effect of cytokines and chemokines (TGF-β, TNF-β, IL-6, IL-10, MCP-1, RANTES) gene polymorphisms in kidney recipients on post-transplantation outcome: influence of donor-recipient match. Transplant Proc 2005;37:764-6.

16. Mungall AJ, Palmer SA, Sims SK, Edwards CA, Ashurst JL, Wilming L, et al. The DNA sequence and analysis of human chromosome 6. Nature 2003;425:805-11.

17. Nachman MW, Crowell SL. Estimate of the mutation rate per nucleotide in humans. Genetics 2000;156:297-304.

18. Frazer KA, Ballinger DG, Cox DR, Hinds DA, Stuve LL, Gibbs RA, et al. A second generation human haplotype map of over 3.1 million SNPs. Nature 2007;449:851-61.