Regular Article

blood

TRANSPLANTATION

Validation of single nucleotide polymorphisms in invasive aspergillosis following hematopoietic cell transplantation

Cynthia E. Fisher,[1,2] Tobias M. Hohl,[3] Wenhong Fan,[4] Barry E. Storer,[5] David M. Levine,[6] Lu Ping Zhao,[4] Paul J. Martin,[1,5] Edus H. Warren,[5] Michael Boeckh,[1,2] and John A. Hansen[1,5]

[1]Department of Medicine, University of Washington, Seattle, WA; [2]Vaccine and Infectious Disease Division, Fred Hutchinson Cancer Research Center, Seattle, WA; [3]Infectious Disease Service, Department of Medicine, and Immunology Program, Memorial Sloan Kettering Cancer Center, New York, NY; [4]Division of Public Health Sciences, and [5]Clinical Research Division, Fred Hutchinson Cancer Research Center, Seattle, WA; and [6]Department of Biostatistics, University of Washington, Seattle, WA

Key Points

- Two SNPs in *PTX3* and *CLEC7a* previously associated with development of proven or probable invasive aspergillosis were validated.
- Thirteen SNPs in 9 genes were associated at $P \leq .05$ with development of IA using a different genetic model than the original study.

Invasive aspergillosis (IA) is a significant cause of morbidity and mortality following allogeneic hematopoietic cell transplantation (HCT). Previous studies have reported an association between IA development and single nucleotide polymorphisms (SNPs), but many SNPs have not been replicated in a separate cohort. The presence of a positive serum galactomannan assay (SGM+) has also been associated with a worse prognosis in patients with IA, and genetic determinants in this subset of patients have not been systematically studied. The study cohort included 2609 HCT recipients and their donor pairs: 483 with proven/probable IA (183 SGM+) and 2126 with no IA by standard criteria. Of 25 SNPs previously published, we analyzed 20 in 14 genes that passed quality control. Samples were genotyped via microarray, and SNPs that could not be genotyped were imputed. The primary aim was to replicate SNPs associated with proven/probable IA at 2 years; secondary goals were to explore the associations using an end point of SGM+ IA or proven/probable IA using a different genetic model or time to IA (3 months vs 2 years) compared with the original study. Two SNPs in 2 genes (*PTX3*, *CLEC7a*) were replicated. Thirteen SNPs in 9 genes had an association at $P \leq .05$ using the secondary aims (*PTX3*, *CLEC7a*, *CD209*, *CXCL10*, *TLR6*, *S100B*, *IFNG*, *PLG*, *TNFR1*), with hazard ratios ranging from 1.2 to 3.29. Underlying genetic differences can influence development of IA following HCT. Identification of genetic predispositions to IA could have important implications in donor screening, risk stratification of recipients, monitoring, and prophylaxis. (*Blood*. 2017;129(19):2693-2701)

Introduction

Invasive aspergillosis (IA) remains a significant cause of infectious mortality in hematopoietic cell transplantation (HCT) recipients, despite improvements in transplantation practices and the introduction of mold-active azoles in the last decade. In contemporary patient cohorts, IA is associated with mortality rates that range from 19% to 60%.[1-5]

HCT recipients have increased susceptibility to IA as a result of profound immune defects from their underlying disease or the HCT process.[1-3,6] Despite similar risk factors, only approximately 3% to 15% of patients who undergo allogeneic HCT develop IA, indicating that other factors are important beyond universal *Aspergillus* exposure. The most robust identified risk factors relate to immune function in HCT recipients, and genetic variation within genes that regulate *Aspergillus*-specific responses or general immune reactivity are likely to contribute to IA susceptibility. In the last decade, more than 40 single nucleotide polymorphisms (SNPs) have been reported to influence the risk of *Aspergillus*-related diseases, and more than 20 have been associated with IAin patients with hematologic malignancies or after HCT (Table 1).[7-18] However, the majority of these studies were not validated independently, and sometimes the results of a study conflict with those of other studies. For example, although associations with Dectin-1/CLEC7A and IA have been reported in some studies (Table 1),[8,9] Chai et al[19] found no association in their study. Additionally, the studies examined small patient cohorts and, at most, several SNPs simultaneously. Thus, published data are insufficient to estimate the pretransplantation risk for individual patients and to guide the clinical management of those at highest risk. Identifying patients at risk for IA through pretransplantation genetic analysis could improve patient care through targeted prophylaxis, surveillance, and treatment strategies.

The dominant role of the fungal biomarker galactomannan (GM) in the contemporary diagnosis of IA is another possible important factor in genetic analyses. The current IA classification system (proven, probable, possible, and no IA) is based on the 2008 revised European Organisation for Research and Treatment of Cancer/Invasive Fungal Infections Cooperative Group and the National Institute of Allergy and Infectious Diseases Mycoses Study Group (EORTC/MSG) criteria.[20] With the introduction of the GM assay, tissue-based and therefore proven diagnoses have declined sharply. Additionally, these

Submitted 13 October 2016; accepted 27 February 2017. Prepublished online as *Blood* First Edition paper, 7 March 2017; DOI 10.1182/blood-2016-10-743294.

The online version of this article contains a data supplement.

The publication costs of this article were defrayed in part by page charge payment. Therefore, and solely to indicate this fact, this article is hereby marked "advertisement" in accordance with 18 USC section 1734.

Table 1. SNPs previously associated with risk of IA in hematopoietic stem cell recipients

Gene	Reference	Population	Genome	Group	N*	SNP	Alleles†	MAF	Statistic	95% CI
AGER	7	W	PT	HCT	223 (41)	rs2070600	T/A‡	0.07	aHR, 2.0	1.0-4.1
			DNR						aHR 2.0	1.0-3.8
CLEC7A	8	W	DNR	HCT	205 (39)	rs16910526	G/T	0.04§	aHR, 2.5	1.0-6.5
			DNR + PT						aHR, 3.9	1.5-10.0
	9	W	PT	HCT and HM	182 (57)	rs7309123	C/G	0.28	OR, 4.9	1.5-15.9
						rs3901533	G/T	0.46	OR, 5.6	1.4-22.8
CD209	9	W	PT	HCT and HM	182 (57)	rs4804800	A/G	0.22	OR, 2.8	1.3-6.0
						rs11465384	C/T	0.03	OR, 2.7	1.2-5.9
						rs7248637	A/G	0.23	OR, 2.4	1.1-5.2
						rs7252229	G/C	0.20	OR, 2.1	1.0-4.2
PTX3	10	W	DNR	HCT	330 (107)	rs2305619‖	A/G	0.44§	aHR, 2.9¶	1.7-5.0
						rs3816527	A/C	0.29§	aHR, 2.6¶	1.5-4.5
S100B	7	W	DNR	HCT	223 (41)	rs9722	C/T	0.18	aHR, 3.2	1.6-6.2
IFNG	11	W	PT#	HCT	139 (81)	rs2069705	C/T	0.48§	NR	$P = .01$
IFNG+	12	W	DNR	HCT	108 (44)	rs2430561	T/A	0.39	OR, 6.1	1.1-35.5
TLR4						rs4986790**	A/G	0.03		
TLR4	13	W	DNR	HCT	366 (103)	rs4986791††	C/T	0.04	HR, 8.0‡‡	2.5-25.3
		W							HR, 2.1[a]	1.0-4.5
TLR1	14	W	PT	HCT	127 (22)	rs5743611	G/C	0.08	OR, 1.2	1.0-1.5
TLR1+	14	W	PT	HCT	127 (22)	rs4833095	A/G	0.35	OR, 1.2	1.0-1.5
TLR6						rs5743810	C/T	0.33		
TLR3	15	W	DNR	HCT	189 (42)	rs3775296	C/A	0.18§	aHR, 2.4	1.3-4.6
TLR5	16	W	PT	HCT	171 (41)	rs5744168	C/T	0.05	OR, 3.3	1.2-9.0
CXCL10	11	W	PT	HCT	139 (81)	rs1554013	C/T	0.30§	OR, 2.2	1.2-3.8
						rs3921	C/G	0.31§	OR, 2.6	1.4-5.0
						rs4257674	A/G	0.31§	OR, 2.8	1.6-5.2
PLG	17	W	PT	HCT	220 (83)	rs4252125	G/A	0.25	aHR, 5.6	1.9-16.5
									aHR, 3.0	1.5-6.1
TNFR1	18	W	PT	HCT and HM	144 (77)	rs4149570	G/T	0.31§	OR, 0.3	$P = .018$
						rs767455	A/G	0.30§	OR, 2.2	$P = .033$

aHR, adjusted hazard ratio; CI, confidence interval; DNR, donor; HCT, hematopoietic cell transplantation; HM, hematologic malignancy; NR, not reported; OR, odds ratio; PT, patient; W, white.

*Total patients (no. of IA cases). Note: We did not include haplotype studies (Seo 2005, Sainz 2008).

†Major/minor.

‡Alleles are designated as reported by Cunha et al.[7] In the National Center for Biotechnology Information Single Nucleotide Polymorphism Database, the alleles are A/G, and in our data set, the corresponding alleles are T/C.

§1000 genomes.

‖rs1840680 in strong linkage disequilibrium.

¶From discovery study. Confirmation study: 2.1 (1.2-3.8) and 1.9 (0.9-3.0).

#Patients with 100% donor chimerism.

**In strong linkage disequilibrium with rs4986791.

††From supplemental Table S3 in Bochud 2008,[13] rs4986791 is in strong linkage disequilibrium with rs4986790.

‡‡Discovery study.

[a]Validation study.

classifications only inconsistently correlate with outcomes; conversely, several studies have shown that dissemination or a positive serum GM (SGM+) portends worse outcomes.[1,21] Furthermore, genetic studies that included patients prior to the widespread use of GM may have missed IA cases, given that the case definition would have relied heavily on less sensitive, culture-based methods, primarily from biopsy specimens.

In this study, we address several key limitations of prior studies by independently validating SNPs that were previously reported to be associated with IA using a large (>2500 patients) cohort of HCT donors and recipients. To minimize diagnostic bias, we retrospectively performed both bronchoalveolar lavage (BAL) fluid and SGM testing on banked patient samples when available. Retrospective testing was performed on samples from patients who met possible IA criteria but were diagnosed prior to the advent of routine GM assay testing. We also analyzed not only the proven and probable IA categories, but further refined "probable IA" to examine if there are different or stronger associations in those with a positive serum GM assay versus all patients with probable and proven IA.

Methods

Literature search

We performed a comprehensive PubMed search using the keywords "aspergillosis" and "polymorphism" to identify studies published through November 2015 that reported an association between a genetic polymorphism and IA at a significance of $P \leq .05$ in HCT recipients. Studies on SNPs in non-HCT populations and those that reported associations with other *Aspergillus*-related conditions (eg, allergic bronchopulmonary aspergillosis, colonization, etc) were excluded. Reported associations with nonsingle nucleotide polymorphism genetic variants, such as variable number tandem repeats, genetic deletions, insertions, or microsatellites were excluded because of limitations of our SNP genotyping array.

Study cohort

Patients who received an allogeneic HCT at the Fred Hutchinson Cancer Research Center (FHCRC) and Seattle Cancer Care Alliance between 1990 and 2011 were eligible. All included patients received T-cell–replete grafts. All recipient and donor samples were collected prior to HCT per approved research

Table 2. Patient and transplantation characteristics by IA designation

Characteristic	Total (n = 2609)	No IA (n = 2126)	Proven/probable (n = 483)	Proven only (n = 194)	SGM+ only (n = 183)
Time to IA diagnosis, median (IQR), d	N/A	N/A	88 (38.5-204.5)	86 (43.3-182)	66 (20.5-122)
Median age (IQR), y	43.9 (32.1-53.1)	43.2 (31.6-52.4)	48.1 (34.7-55.5)	43.7 (31.7-52.0)	49.4 (36.6-56.9)
Diagnosis					
Acute leukemia	966 (37.0)	773 (36.4)	193 (40.0)	78 (40.2)	76 (41.5)
CML	675 (25.9)	589 (27.7)	86 (17.8)	39 (20.1)	33 (18.0)
CLL	75 (2.9)	57 (2.7)	18 (3.7)	5 (2.6)	7 (3.8)
MDS	471 (18.1)	372 (17.5)	99 (20.5)	40 (20.6)	35 (19.1)
MM	143 (5.5)	111 (5.2)	32 (6.6)	12 (6.2)	15 (8.2)
Malignant lymphoma	279 (10.7)	224 (10.5)	55 (11.4)	20 (10.3)	17 (9.3)
Sex match					
M-M	893 (34.2)	722 (34.0)	171 (35.4)	72 (37.1)	67 (36.6)
M-F	611 (23.4)	497 (23.4)	114 (23.6)	38 (19.6)	47 (25.7)
F-F	582 (22.3)	432 (20.3)	91 (18.8)	38 (19.6)	29 (15.8)
F-M	523 (20.0)	475 (22.3)	107 (22.2)	46 (23.7)	40 (21.9)
Stem cell source					
BM	1397 (53.5)	1165 (54.8)	232 (48.0)	125 (64.4)	89 (48.6)
Peripheral	1205 (46.2)	954 (44.9)	251 (52.0)	69 (35.6)	94 (51.4)
Donor type					
Matched related	1188 (45.5)	998 (46.9)	190 (39.3)	83 (42.8)	70 (38.3)
Matched unrelated	957 (36.7)	775 (36.5)	182 (37.7)	69 (35.6)	70 (38.3)
Mismatched	464 (17.8)	353 (16.6)	111 (23.0)	42 (21.6)	43 (23.5)
Myeloablative	2125 (82.7)	1769 (83.2)	388 (80.3)	172 (88.7)	154 (84.2)
Disease risk					
Low	621 (23.8)	538 (25.3)	83 (17.2)	29 (14.9)	39 (21.3)
Intermediate	679 (26.0)	576 (27.1)	103 (21.3)	32 (16.5)	42 (23.0)
High	1154 (44.2)	889 (41.8)	265 (54.9)	126 (64.9)	88 (48.1)
TBI dose					
None or ≤1200	2260 (77.8)	1572 (73.9)	394 (81.5)	132 (68.0)	162 (88.5)
>1200	726 (27.8)	554 (26.1)	89 (18.4)	62 (32.0)	21 (11.5)
CMV serostatus					
R+/D+	606 (23.2)	489 (23.0)	117 (24.2)	48 (24.7)	50 (27.3)
R+/D−	670 (25.7)	523 (24.6)	147 (30.4)	52 (26.8)	57 (31.1)
R−/D+	354 (13.6)	298 (14.0)	56 (11.6)	25 (12.9)	24 (13.1)
R−/D−	979 (37.5)	816 (38.4)	163 (33.7)	69 (35.6)	52 (28.4)

BM, bone marrow; CLL, chronic lymphocytic leukemia; CML, chronic myeloid leukemia; D+, donor seropositive; D−, donor seronegative; F, female; IQR, interquartile range; M, male; MDS, myelodysplastic syndrome; MM, multiple myeloma; N/A, not applicable; R+, recipient seropositive; R−, recipient seronegative; TBI, total body irradiation.

protocols. The FHCRC maintains an extensive electronic patient record database, and baseline pretransplantation demographic information, including age, sex, XY karyotype, ABO blood group, and race were prospectively collected through this clinical research database. Given that IA designations are not captured in this database and represent a combination of clinical, radiographic, and laboratory findings, IA phenotypes were based on retrospective chart review of all patients. Because only a small percentage of patients were nonwhite (9%), only white patients were included in the analysis. Genotyping for HLA-A, B, C, DRB1, and DQB1 was performed to determine patient-donor HLA matching. The study and use of these samples was approved by the FHCRC Institutional Review Board.

IA classifications were per the 2008 revised EORTC/MSG criteria.[20] The primary adjudication of IA phenotypes was performed by author C.E.F. on the basis of chart review, including review of images, reading of images by a radiologist, and review of laboratory/microbiologic information. In cases that were not clear, authors T.M.H. and M.B. also evaluated the data, and a consensus was reached through discussion. Patients with possible IA (typically because of the presence of clinical and radiographic criteria but lacking a microbiologic criterion) were not included. Other reasons for exclusion were: if patients had possible, probable, or proven IA pretransplantation, or another invasive mold infection either pre- or posttransplantation (n = 235); if patients received mold-active therapy or prophylaxis for their transplantation (voriconazole, amphotericin B deoxycholate or its derivatives, or posaconazole; n = 308), or if patients had insufficient follow-up information (n = 150).

Laboratory information

Preparation, genotyping, and imputation of samples. All recipient and donor samples were collected before HCT according to research protocols approved by the FHCRC Institutional Review Board. Samples were prepared and genotyped using 3 different genotyping platforms, as described previously.[22] The genotypes of the candidate SNPs were determined separately for each platform. Candidate SNPs not genotyped on the array were imputed using the 1000 Genomes Project phase 1 SNPs as a reference panel and the software IMPUTE version 2 (http://mathgen.stats.ox.ac.uk/impute/impute_v2.html). For imputed SNPs, the posterior probability of the most probable genotype is calculated as the probability of observing an unobserved genotype at the imputed locus, given all of the observed genotypes in the flanking region. The imputed SNP genotype in a sample was retained only if the posterior probability of the most probable genotype exceeded 0.8; otherwise, it was treated as missing. On each platform, an imputed SNP was considered passing if the percent of missing genotypes was <10%, the average posterior probability of the most probable genotype for all samples on that platform was >0.9, and the Hardy-Weinberg equilibrium P value was $<10^{-8}$. Five SNPs did not pass these quality control measures on the Affymetrix 5.0 platform (Thermo Fisher Scientific), and the samples genotyped on that platform were excluded from the analysis (supplemental Table 1, available on the *Blood* Web site). Supplemental Table 1 details whether each SNP was genotyped or imputed for each cohort.

SGM testing

FHCRC maintains an extensive specimen biorepository dating back to the early 1990s, with banking of excess serum and plasma samples, and in some cases, BAL fluid. For patients who were initially considered as having "possible IA" and who did not have a BAL and/or SGM assay performed clinically, we attempted to locate samples to test retrospectively. Serum and BAL fluid that was

Table 3. Replication of SNPs associated with development of proven/probable IA (n = 439) following hematopoietic stem cell transplantation by 2 years

Gene	SNP	Genome	Alleles	MAF	Model	HR	95% CI	*P*
PTX3	rs2305619	DNR	A/G*	0.47	Recessive	1.33	1.09-1.64	.005
CLEC7A	rs16910526	DNR + PT	G/T†	0.08	Dominant + wild type‡	1.49	1.11-2.02	.009

HR, hazard ratio.
*Minor/major for PTX3 and TLR4 to keep our model consistent with the original articles, given that the opposite alleles were more frequent in our study.
†Major/minor.
‡Dominant in donor, wild type in recipient, with wild type in both as comparison.

collected within 7 days of the possible IA diagnosis were eligible for testing. The GM testing was performed on serum and BAL fluid by a certified technician using the Bio-Rad Platelia *Aspergillus* EIA (Bio-Rad, Redmond, WA), per package insert. An index of ≥0.5 was considered positive for both serum and BAL.[23]

Statistical analysis

We analyzed the association with 2 major phenotypes: proven/probable IA, according to EORTC/MSG criteria, and a novel phenotype based on SGM. Genetic associations with IA were evaluated using Cox regression, based on time from HCT to infection. Death without infection was a competing risk for all analyses. All analyses were adjusted for donor relationship and for the first 4 principal components of genome-wide association study SNPs to account for possible population stratification. We did not adjust for any baseline risk factors for IA, because these factors are not plausibly related to genotype and would not be true confounders. Despite this, given the large time period of this study, we did perform an analysis controlling for year of transplantation, and this did not change our results (data not shown). We also did not adjust for post-HCT factors that may influence the risk of infection, such as neutropenia or steroid use. First, because adjusting for factors potentially in the causal pathway for infection represents a study of mechanism, this was beyond the scope of this investigation. Second, when we analyzed the relationship between neutropenia and steroid use (using graft-versus-host disease as a proxy) and each SNP, no association was found with any of the SNPs. For each candidate SNP locus, we assessed allelic and genotypic (recessive and dominant) association, as previously described.[22,24] We analyzed the relationship of the different SNPs and development of IA during 2 time periods following HCT: within 3 months and 2 years. These end points were chosen on the basis of prior studies that looked at similar time frames and also to investigate if there is a different relationship for those who develop early IA versus those who develop IA later.

A 2-sided $P \leq .05$ was used as the threshold of significance when the SNP was strictly replicated (eg, same genotype [patient or donor] and genetic model) using the standard phenotype of proven/probable IA at 2 years. When any of these factors differed from the original article (eg, SGM+ IA phenotype, IA at 3 months, different genetic model), we considered it an exploratory analysis. Although the SNPs analyzed were still previously reported in some manner and have biologic plausibility, if the analysis differed from that originally reported, we reported the results and displayed the *P* values separately if the *P* value was ≤.05.

Results

Candidate SNP selection

Previously published studies of 25 SNPs in 14 genes met study inclusion criteria, as described in the Methods section. Table 1 outlines these SNPs and the reported degree of association with IA susceptibility. Of the 25 SNPs, 20 passed quality control by genotyping or imputation on all 3 platforms used in this study. Of the 5 that did not pass quality control, all problems occurred with the first platform (Affymetrix), and included 1 SNP with low posterior probability (*TNFR1*, rs767455), 1 SNP with low Hardy-Weinberg equilibrium (*TLR1*, rs4833095), and 3 SNPs with low call rates (*CLEC7A*, rs3901533; *CD209*, rs4804800 and rs7252229). Three SNPs had a minor allele frequency (MAF) of <0.05 in the Affymetrix cohort (*TLR5*, rs5744168, MAF, 0.03; *AGER*, rs2070600, MAF, 0.04; and *CD209*, rs11465384, MAF, 0.04), although the other 2 cohorts had an MAF of ≥0.05 for these SNPs, and these were included in the final analysis (supplemental Table 1).

Study cohort

Overall, 483 patients met criteria for either proven (n = 194) or probable (n = 289) IA ("proven/probable"), and 2126 HCT recipients had no evidence of IA during follow-up ("no IA"). Additionally, we separately analyzed the subset of patients who had positive SGM ("SGM+"), either performed prospectively as part of their diagnostic

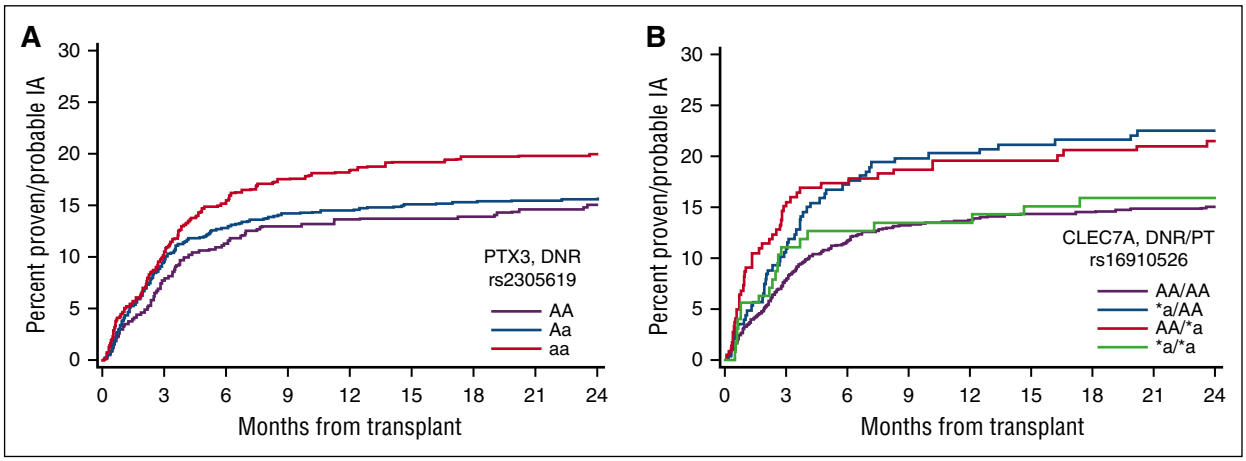

Figure 1. Cumulative incidence curves of replicated SNPs in *PTX3* and *CLEC7A* and development of IA in the 24 months following hematopoietic cell transplantation. (A) *PTX3*; (B) *CLEC7A*. AA, homozygous for major allele; Aa, heterozygous; aa, homozygous for minor allele; *a, Aa, or aa (dominant for minor allele); DNR/PT, donor and patient genotypes.

Table 4. Discovery of different models of SNPs associated with IA development within 3 months and 2 years after hematopoietic stem cell transplantation; cases defined as either proven/probable or SGM+

Gene	SNP	Allele*	MAF	Phenotype	Genome	Time after transplantation	Model	HR/OR	95% CI	P
PTX3	rs2305619	A/G	0.47	Proven/probable	DNR	2 y	Allelic	1.20	1.04-1.37	.009
				SGM+	DNR	3 mo	Allelic	1.42	1.08-1.86	.01
							Recessive	2.44	1.33-4.42	.004
						2 y	Allelic	1.30	1.04-1.641	.022
							Recessive	1.54	1.02-2.34	.039
PTX3	rs3816527	C/A	0.42	SGM+	DNR	3 mo	Allelic	1.41	1.07-1.87	.015
							Recessive	3.29	1.53-7.08	.002
						2 y	Allelic	1.30	1.03-1.61	.024
							Recessive	1.82	1.11-2.94	.017
CLEC7A	rs16910526	A/C	0.08	Proven/probable	DNR	2 y	Allelic	1.30	1.03-1.65	.03
					PT	3 mo	Allelic	1.69	1.29-2.22	.0001
							Dominant	1.77	1.29-2.42	.0003
						2 y	Allelic	1.29	1.03-1.62	.03
							Dominant	1.30	1.01-1.68	.04
CLEC7A	rs7309123	C/G	0.43	Proven/probable	DNR	2 y	Recessive	1.29	1.02-1.63	.030
CD209	rs7248637	G/A†	0.10	SGM+	DNR	3 mo	Allelic	0.56	0.32-0.97	.042
					PT	3 mo	Allelic	0.54	0.30-0.97	.037
							Dominant	0.50	0.27-0.93	.028
						2 y	Allelic	0.65	0.42-0.99	.046
							Dominant	0.56	0.35-0.91	.018
CXCL10	rs1554013	C/T	0.45	SGM+	PT	3 mo	Allelic	1.38	1.05-1.81	.020
							Recessive	1.59	1.04-2.42	.033
CXCL10	rs3921	C/G	0.45	SGM+	PT	3 mo	Allelic	1.35	1.03-1.77	.031
							Recessive	1.55	1.01-2.36	.044
CXCL10	rs4257674	A/G	0.45	SGM+	PT	3 mo	Allelic	1.35	1.03-1.77	.028
							Recessive	1.56	1.02-2.38	.041
TLR6	rs5743810	C/T	0.40	Proven/probable	DNR	2 y	Dominant	1.38	1.12-1.71	.003
				SGM+	DNR	2 y	Dominant	1.41	1.01-1.98	.046
S100B	rs9722	C/T	0.11	SGM+	DNR	3 mo	Allelic	1.66	1.17-2.34	.004
							Dominant	1.79	1.19-2.69	.005
						2 y	Allelic	1.47	1.09-1.98	.011
							Dominant	1.54	1.09-2.17	.013
IFNG	rs2069705	T/C‡	0.34	SGM+	DNR	2 y	Dominant	0.73	0.58-0.99	.045
PLG	rs4252125	G/A	0.28	Proven/probable	DNR	3 mo	Recessive	1.57	1.06-2.33	.026
						2 y	Recessive	1.39	1.02-1.90	.037
TNFR1	rs4149570	C/A	0.40	SGM+	DNR	3 mo	Recessive	0.53	0.27-1.00	.05

CI, confidence interval; DNR, donor; HR, hazard ratio; OR, odds ratio; PT, patient.
*Major/minor allele.
†Original study had major/minor alleles A/G.
‡Original study had major/minor alleles C/T.

workup or retrospectively in patients for whom we had stored serum. Eighty additional probable IA cases (17%) were identified through retrospective serum and BAL GM testing, and in 58 patients who were proven/probable by a different method and did not have an SGM performed at the time, we confirmed a positive SGM. Table 2 summarizes clinical characteristics of these groups.

Validation of previous SNPs

Using strict replication criteria (same genotype, genetic model, proven/probable IA end point at 2 years only) we validated 2 SNPs in 2 genes: rs1840680 in *PTX3* and rs7309123 in *CLECL7a* (Table 3). Figure 1 illustrates the cumulative incidence of IA development following HCT in the validated SNPs.

Exploratory analysis of SNPs

In addition to using the strict replication criteria above, we also explored associations using the same SNPs but with deviations in the initial model, including different genotypes (donor vs recipient), genetic models (allelic, recessive, dominant), time to IA development (3 months vs 2 years), and end point (SGM+). Overall, 13 different SNPs in 9 genes had an association at $P \leq .05$ with at least 1 of the models (Table 4): 5 SNPs in 4 genes in the proven/probable group, and 10 SNPs in 7 genes in the SGM+ group. Associations in 2 SNPs (rs2305619 in PTX3 and rs5743801 in *TLR6*) were found with both end points. Cumulative incidence curves for these SNPs are shown in Figure 2.

Several SNPs had a low MAF in our population. In order to investigate how this might affect our ability to successfully validate these SNPs, we performed an analysis of the minimal detectable hazard ratio for a range of MAF with a power of 80% and a type 1 error rate of 0.05 (supplemental Figure 1). For the available sample size and number of cases in our study, a low MAF, such as 0.05 for rs4986790 in *TLR4*, requires a very large association (hazard ratio of >12) to reliably identify the relationship.

Discussion

Using a large cohort of HCT recipients with donor-recipient–paired DNA samples, we examined 20 candidate SNPs in 14 genes that had well-documented associations with IA susceptibility following HCT in

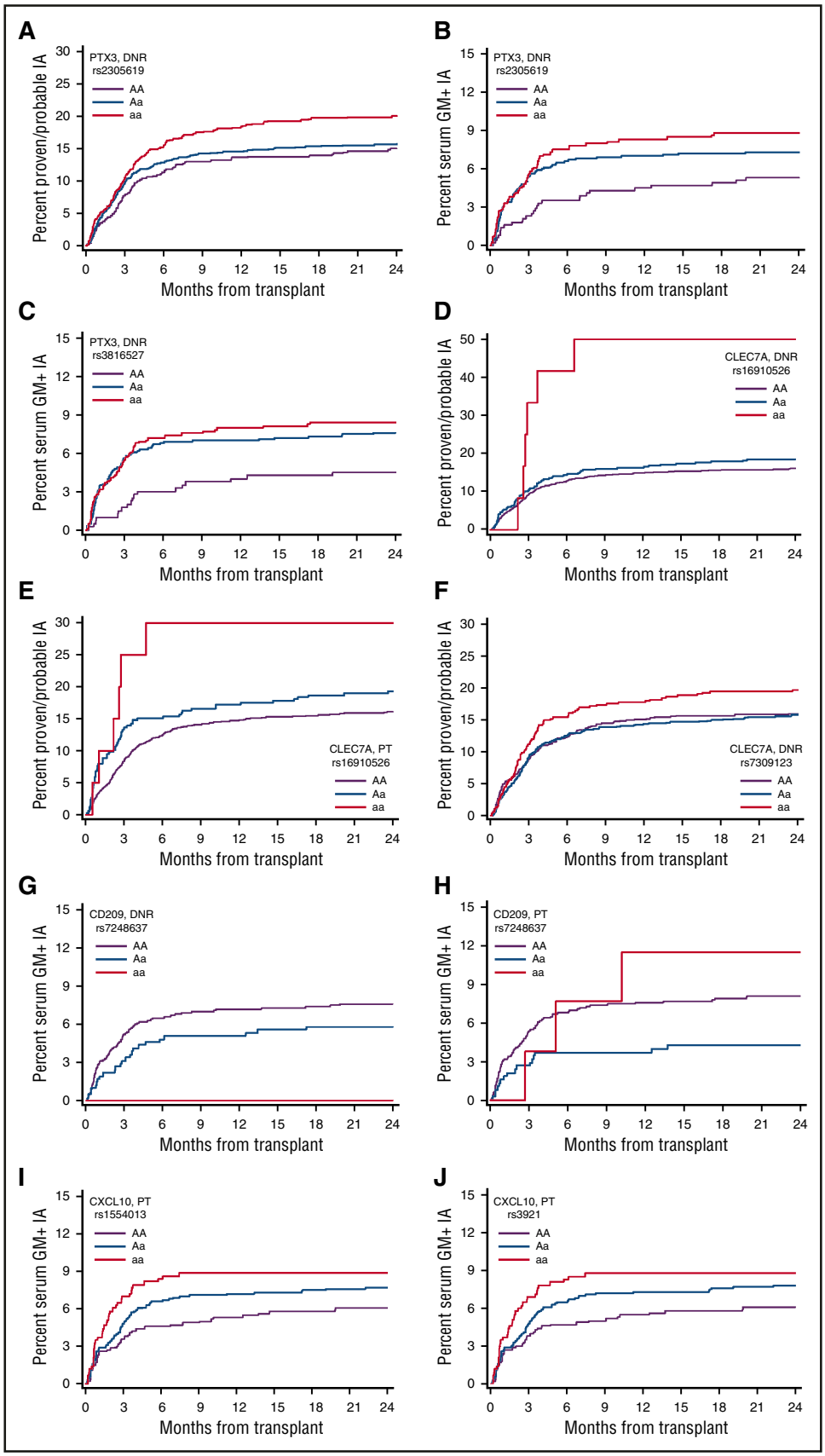

Figure 2. Cumulative incidence curves of SNPs identified in the discovery study and development of IA in the 24 months following hematopoietic cell transplantation, using endpoints of proven/probable IA or serum GM+ IA. (A,D,E,F,L,P) IA; (B,C,G-K,M-O,Q) SGM+.

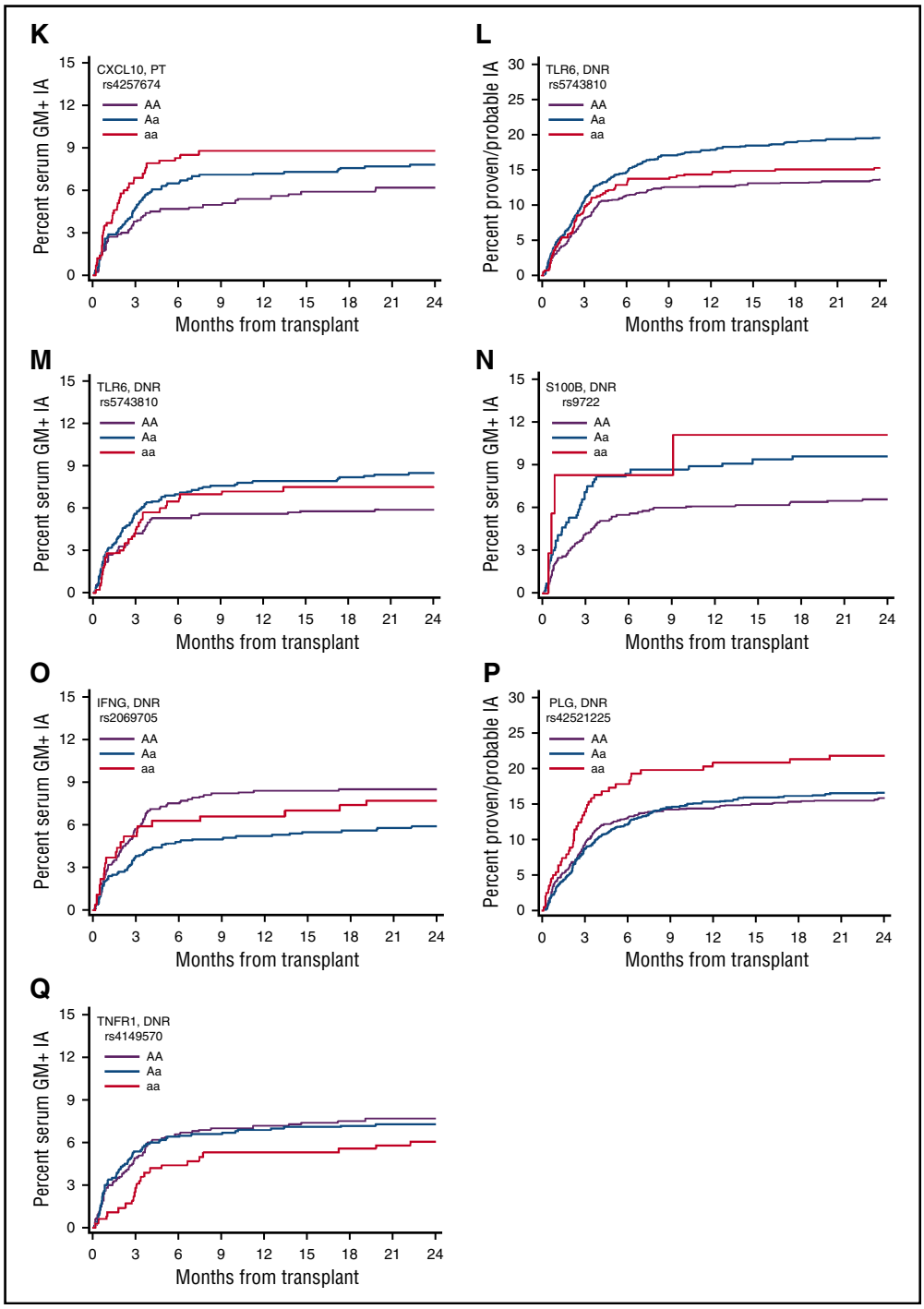

Figure 2. (Continued).

previous studies. All of the SNPs examined are on genes that are important in immune function and therefore have biologic plausibility in IA susceptibility. Using the EORTC/MSG diagnostic categories of proven and probable to define our IA phenotype, we strictly replicated 2 SNPs from 2 genes. We also examined the association using not only the traditional EORTC/MSG case definitions of proven or probable, but also a novel phenotype based on a positive SGM assay. In modern cohorts, the diagnosis of IA increasingly relies on GM, and a positive SGM has been associated with worse prognosis, raising the possibility that patients with fungal angioinvasion, as judged by a positive SGM assay, may have more profound immune defects.[21] However, because this is a novel phenotype in genetic analysis, we did not consider analysis of this end point as a strict replication. Using the SGM+ end point, we did find several associations at $P \leq .05$, including SNPs in *PTX3*, *CD209*, *CXCL10*, *TLR6*, *TNFR1*, IFNG, and *S100B*. Additionally, using the proven/probable end point, but with genetic models that differed from the original studies, we also found associations in *PTX3*, *CLEC7a*, *TLR6*, and *PLG*. It is important to note that our statistical model purposely did not include adjustments for any baseline risk factors for IA, given that these factors are not plausibly related to

genotype and would not be true confounders. We also did not adjust for post-HCT factors that may influence the risk of infection, such as neutropenia or steroid use, because adjusting for factors potentially in the causal pathway for infection represents a study of mechanism and was beyond the scope of this investigation.

Among the SNPs that were either strictly replicated or had an association in this study, a number occur in genes that have well-defined roles in host defense against aspergillosis, either in human cells challenged with *A. fumigatus* or in murine pulmonary challenge models. Pentraxin-3 is soluble collectin that binds to conidia and facilitates their uptake by myeloid cells via complement receptor 3 and Fc γ receptor IIA–dependent mechanisms.[25] Pentraxin-3–deficient mice are susceptible to pulmonary challenge with high doses of *A. fumigatus* conidia in the absence of exogenous immune suppression.[26,27] CD209/DC-SIGN is a type II C–type lectin receptor that mediates the in vitro binding and internalization of *A. fumigatus* conidia by human alveolar macrophages. This process can be inhibited by inclusion of *A. fumigatus*–derived GMs in the assay.[28] Dectin-1/CLEC7A, a well-characterized type II C–type lectin receptor with an intracellular immunoreceptor tyrosine–based activation motif, binds *A. fumigatus* β-glucan on germinating conidia and initiates a signal cascade that culminates in alveolar macrophage nuclear factor κB–dependent cytokine responses.[29-31] Dectin-1–deficient mice have variable susceptibility to *A. fumigatus* challenge in the absence of exogenous immune suppression,[32,33] in part due to molecular redundancy in *A. fumigatus* recognition by host cells. Humans with a Mendelian defect in Dectin-1/CLEC7A signaling (a stop codon polymorphism: Y238X) do not spontaneously develop IA,[34] although this study and others have defined a clear IA risk in the context of immune damage associated with HCT.[8] CARD9 is a signal transducer that integrates signals from Dectin-1 and other CLRs (eg, Dectin-2/CLEC6A) that are implicated in *Aspergillus* recognition.[35,36] Mendelian defects in *CARD9* can lead to the spontaneous development of IA in humans,[37] consistent with the model that multiple recognition events of *A. fumigatus* polysaccharide ligands by host receptors converge on CARD9 and individually contribute to innate defense against aspergillosis. The associations of other SNPs in this study (eg, in the *CXCL10* promoter) provide a rationale for further investigations into the biologic role of associated genes in cultured human cells and animal models of disease.

A key strength of this study was the number of cases and controls and the presence of both donor and recipient samples for genetic testing. The aggressive standard diagnostic workup throughout the study period at our institution, with routine radiographic studies, lung fluid sampling, and serum and BAL GM testing in the era of GM assay availability, coupled with our ability to test banked serum and BAL samples retrospectively for the presence of GM, reduced the diagnostic bias over time and allowed for inclusion of more cases. We also manually reviewed all patient charts, including imaging, clinical history, and microbiologic results, in order to minimize phenotype misclassification bias.

This study also had several limitations and/or unanticipated results. Although the main goal of this analysis was to replicate previously reported associations, our analysis broadened the search for associations by considering recipient or donor genomes, 2 IA phenotypes, 2 time points, and 3 genetic models. The 2-sided significance level of .05 or less that we used should be considered the minimal threshold of significance. This threshold does not account for the multiplicity of SNPs examined nor the multiple association analyses performed for each SNP when the genetic model or end point deviated in any way from that in the primary article. The appropriate significance threshold to use in these cases is debatable, given that there is a higher pretest probability of these SNPs than in a true discovery study. The significance threshold of .0001, representing a Bonferroni correction for the 20 approximately independent SNPs × 24 genetic/end point combinations, would constitute the most stringent value. However, because this will likely be the largest IA genome-wide association study in this population, we presented all SNPs that had a $P \leq .05$ in Table 4, under the premise that the results will provide some guidance for ranking candidate SNPs in future association studies. Although we were able to replicate 2 SNPs found in genes with strong biologic plausibility, we were unable to validate all of the SNPs tested. Some SNPs could not be validated because of quality control (eg, call rate, posterior probability, etc). However, an unanticipated result was our inability to validate the *TLR4* SNP rs4986790. This SNP was reported as part of a haplotype (S4) with *TLR4* SNP rs4986791 (which is in strong linkage disequilibrium with rs4986790) by our group as significantly associated with IA (adjusted hazard ratio, 2.49; 95% confidence interval, 1.15-5.41; $P = .02$ in the validation study for all donors),[13] although the SNP rs4986791 was also found to be significant by itself (reported in their supplemental Table S3[13] and in our Table 1). The low MAF of this SNP, 0.05 in our study, could explain our inability to replicate results of the previous study. As supplemental Figure 1 illustrates, even with our large cohort, with 80% power, our minimum detectable hazard ratio approached 13, which is much larger than the hazard ratio of 2.49 reported in the previous study. Additionally, we removed all of the patients that were included in both our cohort and the former cohort, which decreased our statistical power further. Therefore, this lack of validation may be a statistical issue and not a biologic one. Of note, de Boer et al[12] similarly did not validate this SNP alone (only in combination with interferon γ). Last, the majority (>80%) of recipients in our study received myeloablative conditioning, and so the generalizability to other HCT settings is unclear.

This study is the largest to date, to our knowledge, to validate SNPs previously shown to confer an increased risk of IA in HCT recipients. Our findings suggest that the development of IA may be significantly influenced by genetic variation in genes important to immune function. Identifying patients who are at risk for IA pretransplantation through genetic analysis, or, in future studies, identifying those with IA who are at risk for poor outcomes, could improve patient care through targeted prophylaxis, surveillance, and treatment (ie, combination vs monotherapy) strategies, or even influence donor selection. This study constitutes a critical milestone toward designing such strategies. Future studies in large cohorts using a discovery approach may provide further insight into the genetic influences that underlie IA development.

Acknowledgments

The authors thank Zachary Stednick for database service; Terry Stevens-Ayers, Jessica N. Yi, Elsa Garnace, and Bradley Edmison for laboratory and repository assistance; Jenna Gravley for technical assistance; Anajane Smith for HLA typing; Sonia Goyal and Jessica N. Yi for medical chart assistance; Hu Xie for the creation of supplemental Figure 1; and Gary Schoch for data management.

This work was supported by grants from the National Institutes of Health (NIH) National Heart, Lung, and Blood Institute (HL087690 and HL105914 [J.A.H., P.J.M., E.H.W., M.B., B.E.S., D.M.L., and W.F.], K24HL093294 [M.B.]), National Cancer Institute (CA18029 and CA15704 [Fred Hutchinson Cancer Research Center]), NIH core grant P30 CA008748 [Memorial Sloan Kettering Cancer Center]), National Institute of Allergy and Infectious Diseases (T32 AI007044-37 [C.E.F.]),

and by Joel Meyer's Endowment Scholarship (C.E.F.), the Robert Sinskey Foundation (C.E.F., T.M.H.), and a Burroughs Wellcome Fund Investigator in the Pathogenesis of Infectious Diseases Award (T.M.H.).

Authorship

Contribution: C.E.F. was responsible for literature search, study design, data collection, data analysis and interpretation, and writing; M.B., J.A.H., and T.M.H. were responsible for literature search, study design, data interpretation, and writing; P.J.M. and E.H.W. were responsible for study design, data interpretation, and writing; and D.M.L., W.F., L.P.Z., and B.E.S. were responsible for primary statistical analysis and interpretation, study design, and writing.

Conflict-of-interest disclosure: The authors declare no competing financial interests.

Correspondence: John A. Hansen, Fred Hutchinson Cancer Research Center, 1100 Fairview Ave N, D2-100, PO Box 19024, Seattle, WA 98109; e-mail: jhansen@fredhutch.org.

References

1. Upton A, Kirby KA, Carpenter P, Boeckh M, Marr KA. Invasive aspergillosis following hematopoietic cell transplantation: outcomes and prognostic factors associated with mortality. *Clin Infect Dis*. 2007;44(4):531-540.
2. Baddley JW, Andes DR, Marr KA, et al. Factors associated with mortality in transplant patients with invasive aspergillosis. *Clin Infect Dis*. 2010; 50(12):1559-1567.
3. Steinbach WJ, Marr KA, Anaissie EJ, et al. Clinical epidemiology of 960 patients with invasive aspergillosis from the PATH Alliance registry. *J Infect*. 2012;65(5):453-464.
4. Maertens JA, Raad II, Marr KA, et al. Isavuconazole versus voriconazole for primary treatment of invasive mould disease caused by Aspergillus and other filamentous fungi (SECURE): a phase 3, randomised-controlled, non-inferiority trial. *Lancet*. 2016;387(10020): 760-769.
5. Marr KA, Schlamm HT, Herbrecht R, et al. Combination antifungal therapy for invasive aspergillosis: a randomized trial. *Ann Intern Med*. 2015;162(2):81-89.
6. Gerson SL, Talbot GH, Hurwitz S, Strom BL, Lusk EJ, Cassileth PA. Prolonged granulocytopenia: the major risk factor for invasive pulmonary aspergillosis in patients with acute leukemia. *Ann Intern Med*. 1984;100(3):345-351.
7. Cunha C, Giovannini G, Pierini A, et al. Genetically-determined hyperfunction of the S100B/RAGE axis is a risk factor for aspergillosis in stem cell transplant recipients. *PLoS One*. 2011;6(11):e27962.
8. Cunha C, Di Ianni M, Bozza S, et al. Dectin-1 Y238X polymorphism associates with susceptibility to invasive aspergillosis in hematopoietic transplantation through impairment of both recipient- and donor-dependent mechanisms of antifungal immunity. *Blood*. 2010; 116(24):5394-5402.
9. Sainz J, Lupiáñez CB, Segura-Catena J, et al. Dectin-1 and DC-SIGN polymorphisms associated with invasive pulmonary Aspergillosis infection. *PLoS One*. 2012;7(2):e32273.
10. Cunha C, Aversa F, Lacerda JF, et al. Genetic PTX3 deficiency and aspergillosis in stem-cell transplantation. *N Engl J Med*. 2014;370(5): 421-432.
11. Mezger M, Steffens M, Beyer M, et al. Polymorphisms in the chemokine (C-X-C motif) ligand 10 are associated with invasive aspergillosis after allogeneic stem-cell transplantation and influence CXCL10 expression in monocyte-derived dendritic cells. *Blood*. 2008; 111(2):534-536.
12. de Boer MG, Jolink H, Halkes CJ, et al. Influence of polymorphisms in innate immunity genes on susceptibility to invasive aspergillosis after stem cell transplantation. *PLoS One*. 2011;6(4): e18403.
13. Bochud PY, Chien JW, Marr KA, et al. Toll-like receptor 4 polymorphisms and aspergillosis in stem-cell transplantation. *N Engl J Med*. 2008; 359(17):1766-1777.
14. Kesh S, Mensah NY, Peterlongo P, et al. TLR1 and TLR6 polymorphisms are associated with susceptibility to invasive aspergillosis after allogeneic stem cell transplantation. *Ann N Y Acad Sci*. 2005;1062:95-103.
15. Carvalho A, De Luca A, Bozza S, et al. TLR3 essentially promotes protective class I-restricted memory $CD8^+$ T-cell responses to Aspergillus fumigatus in hematopoietic transplanted patients. *Blood*. 2012;119(4):967-977.
16. Grube M, Loeffler J, Mezger M, et al. TLR5 stop codon polymorphism is associated with invasive aspergillosis after allogeneic stem cell transplantation. *Med Mycol*. 2013;51(8):818-825.
17. Zaas AK, Liao G, Chien JW, et al. Plasminogen alleles influence susceptibility to invasive aspergillosis. *PLoS Genet*. 2008;4(6):e1000101.
18. Sainz J, Salas-Alvarado I, López-Fernández E, et al. TNFR1 mRNA expression level and TNFR1 gene polymorphisms are predictive markers for susceptibility to develop invasive pulmonary aspergillosis. *Int J Immunopathol Pharmacol*. 2010;23(2):423-436.
19. Chai LY, de Boer MG, van der Velden WJ, et al. The Y238X stop codon polymorphism in the human β-glucan receptor dectin-1 and susceptibility to invasive aspergillosis. *J Infect Dis*. 2011;203(5):736-743.
20. De Pauw B, Walsh TJ, Donnelly JP, et al; European Organization for Research and Treatment of Cancer/Invasive Fungal Infections Cooperative Group; National Institute of Allergy and Infectious Diseases Mycoses Study Group (EORTC/MSG) Consensus Group. Revised definitions of invasive fungal disease from the European Organization for Research and Treatment of Cancer/Invasive Fungal Infections Cooperative Group and the National Institute of Allergy and Infectious Diseases Mycoses Study Group (EORTC/MSG) Consensus Group. *Clin Infect Dis*. 2008;46(12):1813-1821.
21. Fisher CE, Stevens AM, Leisenring W, Pergam SA, Boeckh M, Hohl TM. The serum galactomannan index predicts mortality in hematopoietic stem cell transplant recipients with invasive aspergillosis. *Clin Infect Dis*. 2013;57(7): 1001-1004.
22. Inamoto Y, Martin PJ, Flowers ME, et al. Genetic risk factors for sclerotic graft-versus-host disease. *Blood*. 2016;128(11):1516-1524.
23. Patterson TF, Thompson GR III, Denning DW, et al. Practice guidelines for the diagnosis and management of aspergillosis: 2016 update by the Infectious Diseases Society of America. *Clin Infect Dis*. 2016;63(4):e1-e60.
24. Chien JW, Zhang XC, Fan W, et al. Evaluation of published single nucleotide polymorphisms associated with acute GVHD. *Blood*. 2012; 119(22):5311-5319.
25. Moalli F, Doni A, Deban L, et al. Role of complement and Fcgamma receptors in the protective activity of the long pentraxin PTX3 against Aspergillus fumigatus. *Blood*. 2010; 116(24):5170-5180.
26. Jaillon S, Peri G, Delneste Y, et al. The humoral pattern recognition receptor PTX3 is stored in neutrophil granules and localizes in extracellular traps. *J Exp Med*. 2007;204(4):793-804.
27. Garlanda C, Hirsch E, Bozza S, et al. Non-redundant role of the long pentraxin PTX3 in anti-fungal innate immune response. *Nature*. 2002; 420(6912):182-186.
28. Serrano-Gómez D, Domínguez-Soto A, Ancochea J, Jimenez-Heffernan JA, Leal JA, Corbí AL. Dendritic cell-specific intercellular adhesion molecule 3-grabbing nonintegrin mediates binding and internalization of Aspergillus fumigatus conidia by dendritic cells and macrophages. *J Immunol*. 2004;173(9):5635-5643.
29. Hohl TM, Van Epps HL, Rivera A, et al. Aspergillus fumigatus triggers inflammatory responses by stage-specific beta-glucan display. *PLoS Pathog*. 2005;1(3):e30.
30. Brown GD. Dectin-1: a signalling non-TLR pattern-recognition receptor. *Nat Rev Immunol*. 2006;6(1):33-43.
31. Brown GD, Gordon S. Immune recognition. A new receptor for beta-glucans. *Nature*. 2001; 413(6851):36-37.
32. Jhingran A, Mar KB, Kumasaka DK, et al. Tracing conidial fate and measuring host cell antifungal activity using a reporter of microbial viability in the lung. *Cell Reports*. 2012;2(6):1762-1773.
33. Werner JL, Metz AE, Horn D, et al. Requisite role for the dectin-1 beta-glucan receptor in pulmonary defense against Aspergillus fumigatus. *J Immunol*. 2009;182(8):4938-4946.
34. Ferwerda B, Ferwerda G, Plantinga TS, et al. Human dectin-1 deficiency and mucocutaneous fungal infections. *N Engl J Med*. 2009;361(18): 1760-1767.
35. Loures FV, Röhm M, Lee CK, et al. Recognition of Aspergillus fumigatus hyphae by human plasmacytoid dendritic cells is mediated by dectin-2 and results in formation of extracellular traps. *PLoS Pathog*. 2015;11(2):e1004643.
36. Jhingran A, Kasahara S, Shepardson KM, et al. Compartment-specific and sequential role of MyD88 and CARD9 in chemokine induction and innate defense during respiratory fungal infection. *PLoS Pathog*. 2015;11(1):e1004589.
37. Rieber N, Gazendam RP, Freeman AF, et al. Extrapulmonary Aspergillus infection in patients with CARD9 deficiency. *JCI Insight*. 2016;1(17): e89890.