A systems genomics approach identifies SIGLEC15 as a susceptibility factor in recurrent vulvovaginal candidiasis

Jaeger, M; Pinelli, M; Borghi, M; Constantini, C; Dindo, M; van Emst, L; Puccetti, M; Pariano, M; Ricaño-Ponce, I; Büll, C

Link to publication in University of Groningen/UMCG research database

FUNGAL INFECTIONS

A systems genomics approach identifies *SIGLEC15* as a susceptibility factor in recurrent vulvovaginal candidiasis

M. Jaeger[1,2]*, M. Pinelli[3,4]*, M. Borghi[5], C. Constantini[5], M. Dindo[5], L. van Emst[1], M. Puccetti[5], M. Pariano[5], I. Ricaño-Ponce[6], C. Büll[7], M. S. Gresnigt[1,2,8], X. Wang[1,9], J. Gutierrez Achury[6], C. W. M. Jacobs[1], N. Xu[10], M. Oosting[1], P. Arts[4], L. A. B. Joosten[1], F. L. van de Veerdonk[1], J. A. Veltman[4,11], J. ten Oever[1], B. J. Kullberg[1], M. Feng[10], G. J. Adema[7], C. Wijmenga[6], V. Kumar[6], J. Sobel[12], C. Gilissen[4], L. Romani[5], M. G. Netea[1,13]†

Candida vaginitis is a frequent clinical diagnosis with up to 8% of women experiencing recurrent vulvovaginal candidiasis (RVVC) globally. RVVC is characterized by at least three episodes per year. Most patients with RVVC lack known risk factors, suggesting a role for genetic risk factors in this condition. Through integration of genomic approaches and immunological studies in two independent cohorts of patients with RVVC and healthy individuals, we identified genes and cellular processes that contribute to the pathogenesis of RVVC, including cellular morphogenesis and metabolism, and cellular adhesion. We further identified *SIGLEC15*, a lectin expressed by various immune cells that binds sialic acid–containing structures, as a candidate gene involved in RVVC susceptibility. *Candida* stimulation induced *SIGLEC15* expression in human peripheral blood mononuclear cells (PBMCs) and a polymorphism in the *SIGLEC15* gene that was associated with RVVC in the patient cohorts led to an altered cytokine profile after PBMC stimulation. The same polymorphism led to an increase in *IL1B* and *NLRP3* expression after *Candida* stimulation in HeLa cells in vitro. Last, *Siglec15* expression was induced by *Candida* at the vaginal surface of mice, where in vivo silencing of *Siglec15* led to an increase in the fungal burden. *Siglec15* silencing was additionally accompanied by an increase in polymorphonuclear leukocytes during the course of infection. Identification of these pathways and cellular processes contributes to a better understanding of RVVC and may open new therapeutic avenues.

INTRODUCTION

The fungus *Candida* can cause both systemic and mucosal infections in humans. Although *Candida* is a commensal in most of the healthy population, dysregulation or defects in host defense mechanisms can lead to infection by *Candida* species (*1*, *2*). Systemic *Candida* infections occur almost exclusively in immunocompromised hosts and are characterized by sepsis and deep tissue invasion (*3*). Treatment remains difficult despite recent improvements in antifungal therapies (*4*). Noninvasive mucosal (oral, skin, and vaginal) infections are more prevalent than systemic infections, probably due to the presence of *Candida* spp. within the microbiome of healthy individuals. Although these infections rarely pose a life-threatening risk for immunocompetent patients, superficial candidiasis has a high burden of morbidity and can have a strong impact on quality of life.

Most women suffer at least once in their lifetime from vulvovaginal candidiasis (VVC), a mucosal vaginal infection with *Candida albicans* (*5*–*7*). An estimated proportion of 3 to 8% of these women will experience recurrent episodes of VVC (RVVC), characterized by at least three episodes of VVC per year (*5*, *7*–*12*). Recently, the global burden of RVVC has been estimated at 372 million women affected during their lifetime, demonstrative of the huge burden of disease (*13*). Although antifungal medication serves as an adequate remedy for single episodes of vaginal infection, long-term treatment for women with RVVC is challenging and often unsuccessful (*14*). Several risk factors that increase susceptibility to VVC have been reported, including diabetes mellitus, prolonged use of antibiotics, and oral contraceptives (*10*, *15*–*17*). However, the majority of women with RVVC do not have any of these known risk factors. Several studies have attempted to link specific *Candida* strains to disease incidence and severity, but no conclusive arguments for an association with RVVC have been provided (*18*, *19*). In a live challenge in vivo study, it was shown that VVC susceptibility is most likely dictated by biological, genetic, and environmental factors (*20*). This led us to hypothesize that the genetic background of the host is an important risk factor to be considered in RVVC pathophysiology.

A limited number of studies have examined a series of genetic polymorphisms in different pathways that contribute to increased

[1]Department of Internal Medicine and Radboud Centre for Infectious Diseases (RCI), Radboud University Medical Center, Nijmegen, 6525GA, Netherlands. [2]Radboud Institute for Molecular Life Sciences (RIMLS), Radboud University Nijmegen Medical Center (Radboudumc), Nijmegen, 6525GA, Netherlands. [3]Telethon Institute of Genetics and Medicine (TIGEM), Pozzuoli, Naples, 80078, Italy. [4]Department of Human Genetics, Donders Center for Neuroscience, Radboud University Medical Center, Nijmegen, 6525HR, Netherlands. [5]Department of Experimental Medicine, University of Perugia, Polo Unico Sant'Andrea delle Fratte, Perugia, 06123, Italy. [6]University of Groningen, University Medical Centre Groningen, Department of Genetics, Groningen, 9713GZ, Netherlands. [7]Department of Radiation Oncology, Radiotherapy & OncoImmunology Laboratory, Radboud Institute for Molecular Life Sciences, Radboud University Medical Center, Geert Grooteplein Zuid 32, Nijmegen, 6525GA, Netherlands. [8]Department of Microbial Pathogenicity Mechanisms, Leibniz Institute for Natural Product Research and Infection Biology–Hans Knöll Institute, Beutenbergstraße 11a, Jena, 07745, Germany. [9]College of Computer, Qinghai Normal University, 810008 Xining, China. [10]BGI-Shenzhen, Shenzhen 518083, China. [11]Department of Clinical Genetics, Maastricht University Medical Centre, Maastricht, 6229HX, Netherlands. [12]Infectious Diseases, Wayne State University School of Medicine, Detroit, MI 48201, USA. [13]Human Genomics Laboratory, Craiova University of Medicine and Pharmacy, Craiova, 200349, Romania.
*These authors contributed equally to this work as co-first authors.
†Corresponding author. Email: mihai.netea@radboudumc.nl

Jaeger *et al.*, *Sci. Transl. Med.* **11**, eaar3558 (2019)   12 June 2019

1 of 12

susceptibility to RVVC (*21–28*). However, these studies were based on only a few previously identified genes, and no comprehensive assessment of the host genetic component in RVVC has been performed so far. Genetic polymorphisms associated with the severe skin infections seen in chronic mucocutaneous candidiasis are not associated with RVVC infection (*29*).

The SIGLEC (sialic acid–binding immunoglobulin-like lectin) family proteins play an important role in the immune system because they can regulate immune cell function (*30*). Members of the SIGLEC family, including SIGLEC15, bind sialic acid sugars that frequently terminate glycan structures on the cell surface of different pathogens and modulate immune cell activation (*31*, *32*). Sialic acids have also been identified on the surface of *C. albicans* (*33*). In the present study, we applied a comprehensive genomic approach to assess both rare and common genetic variants that affect susceptibility to RVVC. We performed whole-exome sequencing (WES) and genome-wide association analysis of two independent cohorts of patients with RVVC and healthy controls to identify pathways and gene clusters that may influence host susceptibility to the disease. We used functional human in vitro and experimental murine in vivo infections to validate the identified candidate susceptibility genes, including *SIGLEC15*.

RESULTS
Patient cohorts
Two different cohorts of patients with RVVC were recruited for the WES study. The first cohort consisted of 75 patients with RVVC of Northern European genetic background [North European cohort (NEC)] with 98 healthy NEC individuals as a control population. The second cohort consisted of 85 patients with RVVC with Southern European genetic background [South European cohort (SEC)] and a control population of 77 healthy SEC individuals. Eight samples did not pass quality controls because they had either very uncommon frequencies of variants compared to their reference population (*34*) or both low median coverage together with an exceptionally high number of variants. In total, 170 NEC (75 cases and 95 controls) and 157 SEC (80 cases and 77 controls) samples were included in the analysis. In addition, isolated peripheral blood mononuclear cells (PBMCs) of two independent cohorts of 73 and 50 healthy control individuals of Northern European genetic background were subject to an in vitro *C. albicans* stimulation test of cytokine production capacity. A schematic overview of this study is shown in Fig. 1.

WES and analysis
WES read alignments covered an average of 85.6 and 84.1% of the target regions with more than 20-fold coverage in NEC and SEC, respectively. Variant calling identified 224,127 (NEC) and 312,018 (SEC) variant loci with a total of 8,385,054 (NEC) and 9,745,235 (SEC) alleles. We subsequently performed per-variant filtering to remove variants of low quality and low coverage (<20×). The final set of good-quality variants consisted of 109,014 (NEC) and 187,115 (SEC) loci with 3,678,608 (NEC) and 5,862,276 (SEC) alleles.

Common single-nucleotide polymorphisms (cSNPs) were defined as variants in Hardy-Weinberg equilibrium with a minor allele frequency (MAF) greater than 5%. We found a total of 2,979,231 alleles for 28,328 cSNPs (NEC) and 4,693,252 alleles for 45,819 cSNPs (SEC). Of these, 1471 NEC cSNPs (in 1026 genes) and 2152 SEC cSNPs (in 1370 genes) were associated with RVVC (Fisher's exact

test, $P < 0.05$) (fig. S1). Two hundred one genes contained RVVC-cSNPs (RcSNPs) in both cohorts, and 11 genes had the exact same allele associated with increased RVVC risk (table S1). This overlap of 201 genes between the 1026 RcSNPs genes in the NEC and 1370 RcSNPs genes in SEC was unlikely to occur by chance [hypergeometric test: $P < 10^{-9}$, odds ratio (OR) = 2.53; Fig. 2A]. Functional characterization of the overlapping genes revealed a strong enrichment [false discovery rate (FDR) < 0.05] in gene ontology (GO) annotations for cell adhesion and membrane protein functions (Fig. 2B and table S2), suggesting that these functions may underlie the biological mechanisms driving the disease association in both cohorts.

Functional rare variants (FRVs) were defined as those with an MAF lower than 5% and with a potential functional effect due to their predicted impact on protein function, high evolutionary nucleotide conservation (phyloP > 2), and population frequencies lower than 1%. In total, we identified 46,949 FRVs in 11,585 different genes (table S3). A set of 39 genes had FRVs in multiple cases of both cohorts and none present in control individuals (Table 1). We considered these 39 genes as candidate genes for highly penetrant mutations that predispose to RVVC. After performing a burden test, no gene reached FDR-corrected significance for an enrichment of rare variants in both cohorts (table S4).

Loci associated with *Candida*-stimulated cytokine release
We aimed to identify cSNPs that might influence the capacity of PBMCs to produce cytokines in response to stimulation with *C. albicans*. Because of their ability to alter a quantitative measure, we defined these cSNPs as cytokine-quantitative trait loci (cQTL). PBMCs of 73 NEC controls were sampled and exposed to either *C. albicans* yeast or hyphae for 24 hours or 7 days, and cytokines were subsequently measured. We selected the 52,811 SNPs in 10,938 different genes that were associated with a cytokine alteration after *C. albicans* stimulation (Mann-Whitney test, $P < 0.05$). A number of genes ($n = 1810$) harbored both cQTL and RcSNPs (hypergeometric test: $P = 1.3 \times 10^{-7}$, $OR = 2.06$; Fig. 2C). Most RcSNPs found in both cohorts were also cQTL. The overlap between RcSNPs and cQTL was still substantial after comparing the cQTL of NEC-RcSNPs and SEC-RcSNPs separately ($P = 1.9 \times 10^{-8}$ and $P = 5.1 \times 10^{-4}$, respectively). Thus, this result was not due to the potential bias that could be introduced by combining both cohorts in the meta-analysis. These results suggest a link between the host genetic makeup with regard to cytokine release in response to *C. albicans* stimulation and potential RVVC risk. GO functional characterization showed that these genes are mainly involved in processes of homophilic and cell-cell adhesion via plasma membrane proteins (Fig. 2D and table S5).

Identification of SIGLEC15 as a susceptibility factor in RVVC
By overlapping SNPs and their corresponding genes that were associated with RVVC and those that altered cytokine production after *C. albicans* stimulation, we obtained a signal in both datasets that derived from a polymorphism in the *SIGLEC15* gene (rs2919643) (fig. S1). This coding variant is in complete linkage disequilibrium with another noncoding variant in the same gene, and an overview of the genomic region is shown in fig. S2B and table S6 (*35*). None of the previously reported risk variants (*36*) for RVVC was detected in this study.

To identify the functional consequence of the rs2919643 polymorphism influencing RVVC susceptibility, we assessed cytokine production capacity in PBMCs isolated from individuals carrying

different alleles after stimulation with *C. albicans*. Monocyte-derived cytokines interleukin-6 (IL-6), tumor necrosis factor–α (TNFα), IL-1β, and IL-23 were unaffected by the risk variant, whereas the T cell cytokines IL-17A and IL-22 as well as interferon-γ (IFN-γ) were significantly increased in individuals carrying the disease-associated C allele (IL-17, $P = 0.01$; IL-22, $P = 0.02$; IFNγ, $P = 0.01$) (Fig. 3A). Validation of these findings in a second dataset in an independent cohort of 50 healthy individuals revealed similar results upon stimulation with *C. albicans* yeast (fig. S3). However, we did not observe the same response when cells were stimulated with *Staphylococcus aureus*, arguing for the specificity of this effect (fig. S4). The direct fungal killing capacity of PBMCs was not influenced by the different genotypes, whereas individuals carrying the C allele produced more reactive oxygen species (ROS) after *C. albicans* stimulation (Fig. 3B). SIGLECs are cellular receptors for sialic acid, making them candidate pattern recognition receptors; therefore, we aimed to validate their role in the pathogenesis of *Candida* vaginitis.

Induction of *SIGLEC15* expression by *C. albicans*

To investigate the potential of *C. albicans* to affect *SIGLEC15* expression, we stimulated PBMCs with the fungus for 4 hours. We consistently observed an increase in *SIGLEC15* expression (about 5.5× up-regulation), although this response was variable between donors ($n = 5$) (Fig. 3C). We found that *SIGLEC15* was more highly expressed in PBMCs and neutrophils compared to CD4+ or CD14+ cells. Moreover, next to its highest expression in myeloid cells, *SIGLEC15* expression was moderately high in a vaginal epithelial cell line and was inducible after stimulation with different *C. albicans* morphotypes (fig. S5).

Given our finding that SIGLEC15 likely plays a role in the immune response against
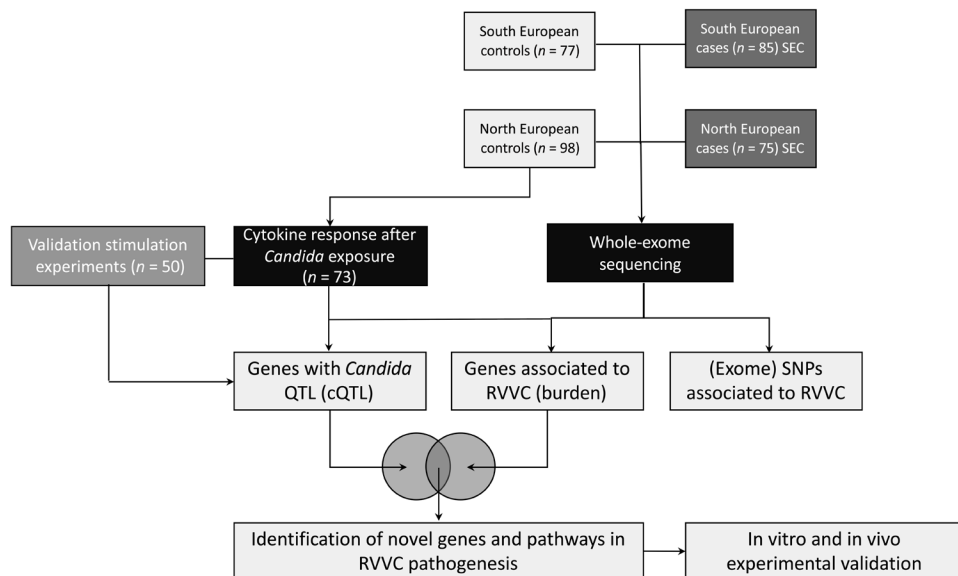
Fig. 1. Schematic overview of the study. We carried out WES on two independent cohorts of patients with RVVC and healthy controls to identify genes and pathways associated with RVVC. We also stimulated human PBMCs in vitro to determine genetic variants that influence *C. albicans*–induced cytokine release. By combining both approaches, we identified genes and pathways important in RVVC pathogenesis that we validated in vitro and in vivo.
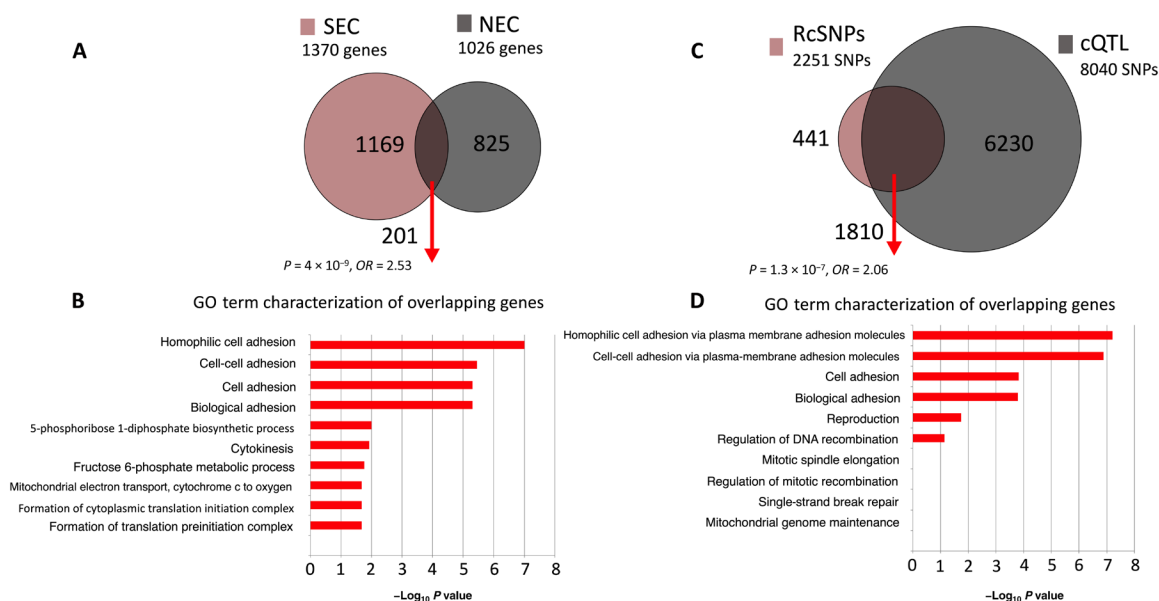
Fig. 2. Analysis of RVVC-related cSNPs (RcSNPs) and cSNPs in both RVVC cohorts. (**A**) Genes containing common RCCV-associated SNPs (RcSNPs) found in both the Southern European and Northern European cohorts (SEC and NEC, respectively) with (**B**) GO term characterization of the overlapping genes. (**C**) Overlap between SNPs that are associated with RVVC (RcSNPs) and SNPs that influenced cytokine production after stimulation (cQTL) in SEC and NEC. (**D**) GO term characterization of the overlapping genes.

Table 1. Genes with FRV only present in RVVC cases. Genes with at least two FRVs only present in candidiasis cases (both NEC and SEC) but not controls were tested using a burden test. This test compares the number of FRVs between cases and controls present within a single gene.

Gene	P value (FRV burden test)	FRVs in NEC	FRVs in SEC	All FRVs
IGFN1	0.002	3	5	8
ATP6AP1L	0.005	4	3	7
ADAMTS20	0.005	3	4	7
SALL1	0.011	4	2	6
YTHDC2	0.011	3	3	6
MMEL1	0.011	2	4	6
MAN2A2	0.011	2	4	6
EVC2	0.011	2	4	6
PKP3	0.011	2	4	6
SMG8	0.023	3	2	5
SLC45A1	0.023	3	2	5
SCAF8	0.023	3	2	5
OR10G3	0.023	3	2	5
SLC5A11	0.023	3	2	5
HSD17B4	0.023	3	2	5
CCKAR	0.023	3	2	5
OR4C16	0.023	2	3	5
PRPF40B	0.023	2	3	5
NFKB1	0.023	2	3	5
ABHD1	0.023	2	3	5
RASAL2	0.023	2	3	5
DACH2	0.023	2	3	5
AREL1	0.048	2	2	4
LANCL1	0.048	2	2	4
CHAMP1	0.048	2	2	4
AMPH	0.048	2	2	4
HACL1	0.048	2	2	4
FRK	0.048	2	2	4
CLCN6	0.048	2	2	4
PPEF2	0.048	2	2	4
RFT1	0.048	2	2	4
BBS5	0.048	2	2	4
SYNPO2	0.048	2	2	4
CYP4V2	0.048	2	2	4
CCDC90B	0.048	2	2	4
CLIP2	0.048	2	2	4
ABCG4	0.048	2	2	4
KIAA1407	0.048	2	2	4
DSG2	0.048	2	2	4

C. albicans, we were interested in the potential of *C. albicans* to induce other SIGLEC family members. We screened the gene expression of all SIGLEC family members in PBMCs and identified *SIGLEC15* as one of the SIGLEC family members most strongly induced by *C. albicans* stimulation. *SIGLEC15* was up-regulated both 4 and 24 hours after stimulation with *C. albicans* (4 hours, $P = 0.0092$; 24 hours, $P = 1.58 \times 10^{-9}$), although after multiple testing correction, only the 24-hour post-stimulation time point remained significant ($P = 3.8 \times 10^{-8}$) (Fig. 3D).

We next aimed to specifically silence *SIGLEC15* in human primary macrophages and PBMCs before stimulating the cells with *C. albicans*. *SIGLEC15*-silenced monocytes did not show reduced IL-6 or TNFα after 24 hours in three independent experiments (Fig. 3E). PGE$_2$ was slightly increased in *SIGLEC15* siRNA-silenced cells compared to control siRNA-treated cells (Fig. 3E). *SIGLEC15* silencing in PBMCs revealed no change in *Candida* killing capacity. We did however observe an increase of IL-1β, whereas TNFα, IL-6, and IL-1Ra remained unaffected (Fig. 3F). Gene silencing in primary cells is challenging because of reduced cell viability and low silencing efficiency; Fig. 3G shows the siRNA silencing efficiency in our experiments. Last, we observed slightly reduced *SIGLEC15* expression in PBMCs isolated from women with RVVC compared to age- and sex-matched PBMCs from healthy controls. This minor reduction was seen at baseline (nonstimulated PBMCs) and after stimulation with *Candida* yeast and hyphae (Fig. 3H). In a subsequent set of experiments, we transfected *SIGLEC15*-negative HeLa cells with plasmids containing either the wild-type or mutated (rs2919643) form of *SIGLEC15*. After stimulation with *Candida* for 24 hours, we observed a greater increase in *NLRP3* and *IL-1β* expression in cells carrying the mutation compared to wild-type cells (Fig. 3I).

Binding of SIGLEC15 to *C. albicans*

To investigate the interaction of SIGLEC15 with *C. albicans*, we performed binding assays using flow cytometry. Heat-killed *C. albicans* yeast was incubated with recombinant SIGLEC15, SIGLEC7, or remained untreated, and was then incubated with a secondary antibody labeled with Alexa Fluor 647. We observed a strong increase in the fluorescent signal in the recSIGLEC15-treated conditions, confirming the binding of the protein to *C. albicans* (Fig. 3J). Treating *C. albicans* with the secondary antibody alone also led to an increase in the signal due to some nonspecific binding of the antibody to *C. albicans* compared to the untreated condition. Within this setup, we used an antibody against β-glucan as positive control (Fig. 3K), because β-glucans are present in the cell wall of *C. albicans* and become highly exposed due to structural changes in the cell wall that occur upon heat killing. Performing the same experiments with a recombinant protein against SIGLEC7 did not result in an increased fluorescence signal underlining the specificity of SIGLEC15 to bind to *C. albicans* (Fig. 3L). To correct for the relatively harsh heat killing procedure, we also performed binding experiments with live *C. albicans* (Fig. 3M). In addition to *C. albicans*, we tested recombinant SIGLEC15 for its capacity to bind *Candida glabrata* and *Candida tropicalis*. Similar to *C. albicans*, SIGLEC15 showed strong binding to both strains (fig. S6).

To test whether the immune response against *Candida* is modulated by the presence or absence of sialic acids on the surface of the fungus, we pretreated *C. albicans* with a neuraminidase that cleaves sialic acid residues from glycans. *C. albicans*–induced ROS increased when the yeast was pretreated with neuraminidase compared to control (Fig. 4, A and C). When *Aspergillus fumigatus* conidia were treated with neuraminidase, no difference between treatment and control treatment was observed (Fig. 4, B and C), underlining the importance of sialic acids in the host response specifically against

Fig. 3. Functional assessment of SIGLEC15 in *Candida* immune response in vitro. (**A**) Cytokine production after stimulation with heat-killed *C. albicans* yeast for 24 hours or 7 days. TNFα, IL-6, IL-1β, IL-17A, IL-22, IFNγ, and IL-23 concentrations were measured in culture supernatant after 24 hours (IL-6, IL-1β, TNFα, and IL-23) or 7 days (IL-17A, IL-22, and IFNγ). PBMCs of individuals with different genotypes (TT, $n = 25$; CT + CC, $n = 9$) for the *SIGLEC15* variant rs2919643 were compared in a dominant model. (**B**) ROS production and *C. albicans* killing by PBMCs with different rs2919643 genotypes. AUC, area under curve. (**C**) RNA sequencing of *SIGLEC15* in PBMCs stimulated with either culture medium (red) or *C. albicans* (black) for 4 hours. (**D**) Gene expression of *SIGLEC* family members profiled in PBMCs of eight healthy donors after stimulation with *C. albicans* for 4 and 24 hours. (**E**) Cytokine production capacity in *SIGLEC15* small interfering RNA (siRNA)–treated human granulocyte-macrophage colony-stimulating factor–derived macrophages after *C. albicans* stimulation for 24 hours. TNFα, IL-6, and prostaglandin E₂ (PGE₂) concentrations were measured in culture supernatants of macrophages stimulated with *C. albicans* yeast for 24 hours after a transfection period of 24 hours with either *SIGLEC15* or control siRNA. (**F**) Cytokine production capacity of human PBMCs silenced for *SIGLEC15* for 24 hours and then stimulated for 24 hours. (**G**) Silencing efficiency in PBMCs at baseline and after stimulation with *Candida* for 24 hours. (**H**) Expression of *SIGLEC15* in healthy controls ($n = 4$) and patients ($n = 3$) suffering from RVVC at baseline and after stimulation with *Candida* (yeast and hyphae) for 7 days. (**I**) Expression of *NLRP3* and *IL1B* in HeLa cells transfected with wild-type (WT) or mutated *SIGLEC15* plasmids. (**J**) Binding of recombinant SIGLEC15 to heat-killed *C. albicans* assayed by mean fluorescent intensity (MFI). Ab, antibody. (**K**) Heat-killed *C. albicans* (blue) with secondary antibody [anti-mouse–phycoerythrin (PE); green] and an antibody against β-glucan (red). (**L**) Heat-killed *C. albicans* (blue) with secondary antibody (green) and recombinant SIGLEC7 (red). (**M**) Live *C. albicans* (blue), live *C. albicans* with secondary antibody (anti-IgG Alexa 647; green), and live *C. albicans* with recombinant SIGLEC15-FC and secondary antibody (red). Data are shown as means ± SEM. Mann-Whitney *U* test, *$P < 0.05$, **$P < 0.01$, and ***$P < 0.001$. ns, not significant. *Y* axes vary across plots.
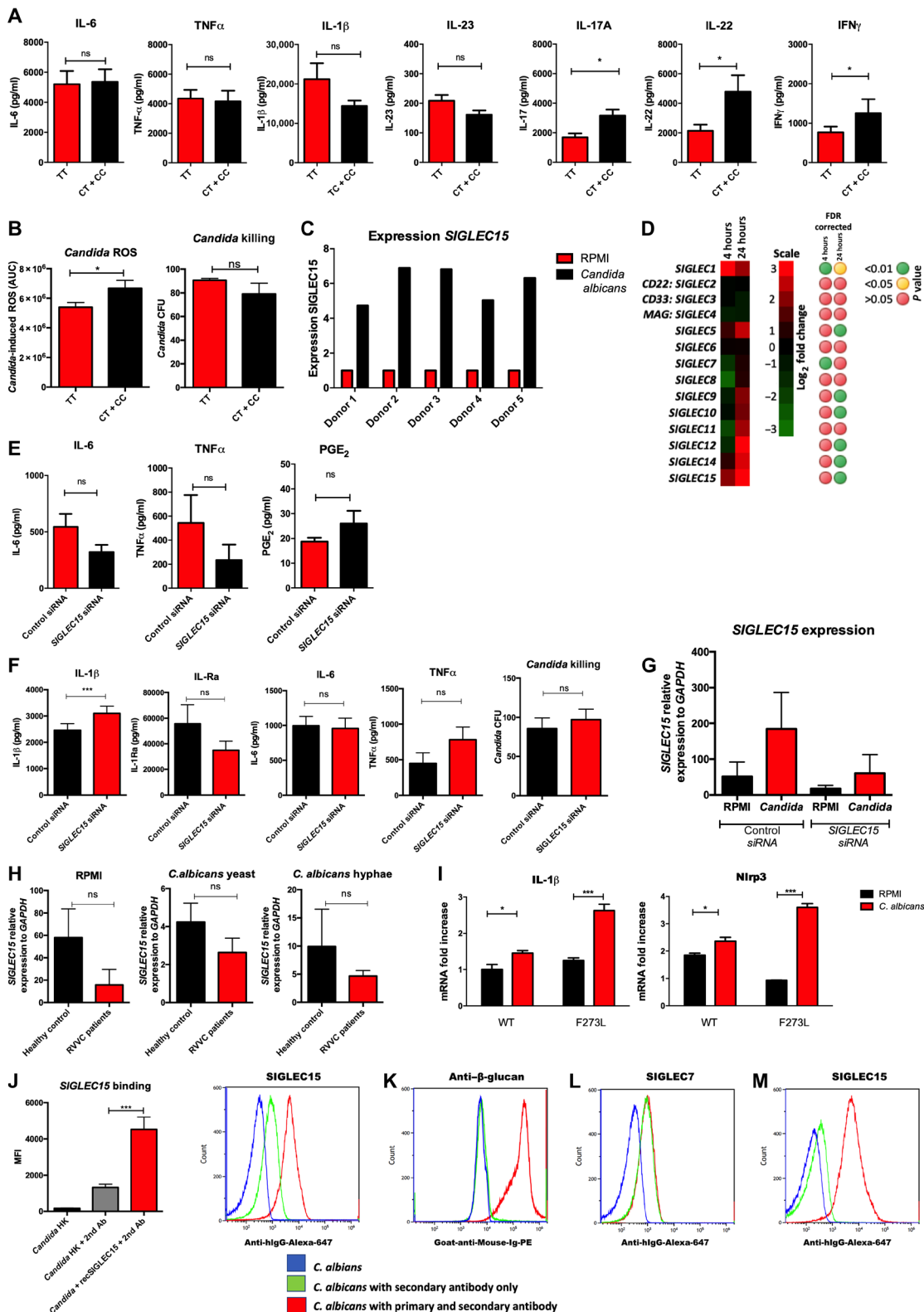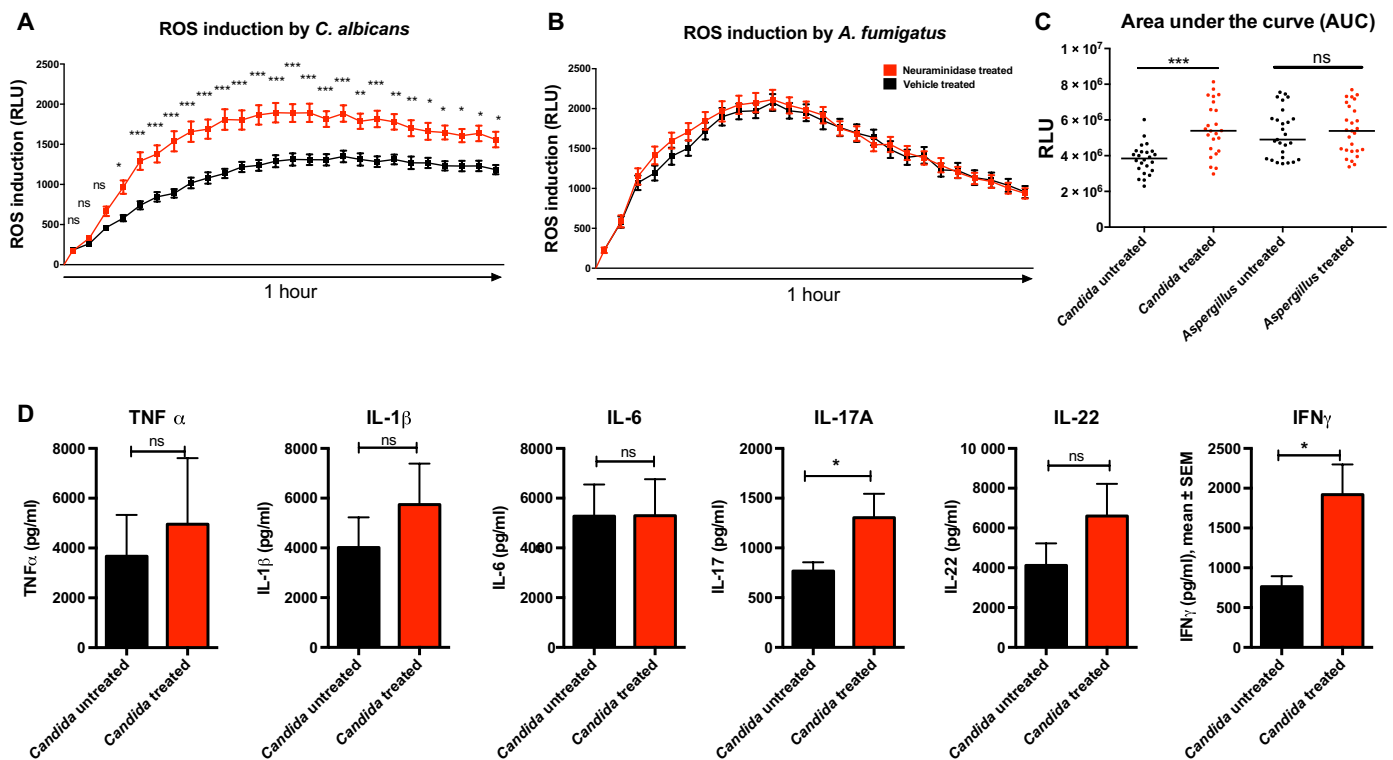
Fig. 4. ROS induced by neuramindase-treated fungi. ROS in PBMCs induced by (**A**) *C. albicans* (1×10^6/ml) and (**B**) *A. fumigatus* (1×10^7/ml) either treated or untreated with neuraminidase. Measurements of ROS induction in relative light units (RLU) per second were taken after the start of treatment for 1 hour in intervals of 2.23 min. (**C**) Area under the curve (AUC) of both experiments in (A) and (B). Data are shown as means ± SEM where *$P < 0.05$, **$P < 0.01$, and ***$P < 0.001$. ns, not significant. (**D**) Cytokine induction by PBMCs after 24 hours or 7 days of stimulation with *C. albicans* either treated or untreated with neuraminidase.

C. albicans. Given this observation, we stimulated PBMCs with either neuraminidase-treated or untreated *C. albicans* and measured the resulting cytokines. Again, minor differences between the treated and untreated *C. albicans* were seen after 24 hours, whereas significant differences were observed for the T cell–derived cytokines IL-17A ($P = 0.02$) and IFNγ ($P = 0.01$) after 7 days of stimulation (Fig. 4D). Together, these results suggest that SIGLEC15 is a receptor in the immune response against *C. albicans*, and patients with the risk allele display an altered cytokine profile in response to stimulation with *C. albicans*.

To investigate the involvement of Siglec15 in the mucosal in vivo immune response against *C. albicans*, we infected C57BL/6 mice intravaginally with *C. albicans*. Expression of *Siglec15* significantly increased during infection and peaked at day 3 after infection (3 days, $P = 0.0003$; 7 days, $P = 0.0025$) (Fig. 5, A and B). After 14 days, *Siglec15* mRNA expression was low but still detectable at the protein level, which might be due to its relatively slow protein degradation within the tissue.

Mice treated with siRNA against *Siglec15* showed significantly higher fungal burdens compared to mice treated with nontargeting siRNA at day 14 ($P = 0.0012$) (Fig. 5C). In line with this, the numbers of polymorphonuclear leukocytes (PMNs) were strongly increased in *Siglec15* siRNA-treated mice during infection, underlining the observation that Siglec15 plays an important role in the immune response against *C. albicans* [14 days post-infection (dpi), $P = 0.0017$] (Fig. 5, D and E). To strengthen these findings, we measured IL-1β and IL-1Ra in vaginal fluids of mice treated with *Siglec15* targeting siRNA or control siRNA. IL-1β was significantly increased in *Siglec15* siRNA-treated mice, whereas IL-1Ra decreased from day 7 to day 14 after infection [IL-1β: day 3 (d3)

$P = 0.0041$, d7 $P = 0.03$, and d14 $P = 0.0002$; IL-1Ra: d7 $P < 0.0001$ and d14 $P = 0.0001$] (Fig. 5F). Consistent with a pathogenic role in RVVC for NLRP3, a protein involved in the maturation of pro–IL-1β to mature and bioactive IL-1β, *Nlrp3* mRNA was decreased in Siglec15 siRNA-treated mice, whereas *Nlrc4* mRNA was significantly decreased throughout the infection (*Nlrp*: d3 $P < 0.0001$, d7 $P = 0.0008$, and d14 $P = 0.0001$; *Nlrc4* d3 $P < 0.0001$ and d14 $P = 0.0004$) (Fig. 5G). IL-6 and IL-10 were reduced in Siglec15 siRNA-treated mice compared to scrambled control-treated mice, whereas IL-17A and IL-22 were significantly increased on day 14 after infection (IL-6: d3 $P < 0.0001$, d7 $P < 0.0001$, and d14 $P = 0.001$; IL-10: d3 $P < 0.0001$, d7 $P < 0.0001$, and d14 $P = 0.0006$; IL-17A: d14 $P = 0.0006$; IL-22: d7 $P = 0.02$ and d14 $P < 0.0001$) (Fig. 5H).

IL-17A and IL-22 produced at the vaginal mucosal surface is not only produced by classical T cells but also by γδ T cells and group 3 innate lymphoid cells (ILC3). To address this, we performed flow cytometry from cells of naïve or infected vaginal tissue after 7 days of infection. The percentage of positive cells per condition is shown in fig. S7. After 7 days of infection, the number of either IL-17– or IL-22–positive cells increased for both γδ T cells and ILC3 cells.

DISCUSSION
The present study is a comprehensive genome-wide investigation on the role of genetic variation for the susceptibility to RVVC. Our pathway analysis of both rare and common genetic variants in the exome identified several major biological processes that are likely

Jaeger *et al.*, *Sci. Transl. Med.* **11**, eaar3558 (2019) 12 June 2019

6 of 12
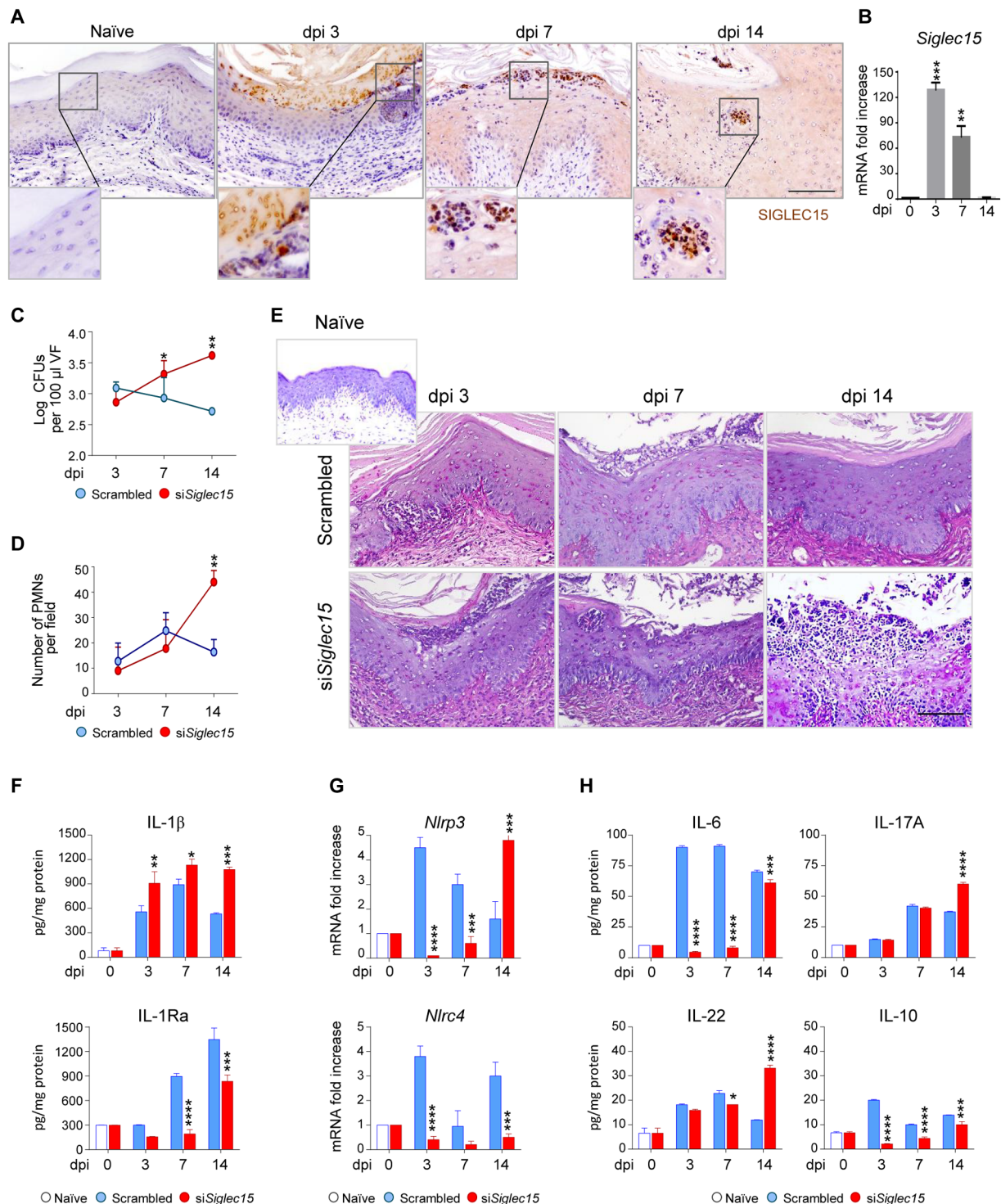
Fig. 5. SIGLEC15 activity in vaginal candidiasis. C57BL/6 mice ($n = 6$) were intravaginally inoculated with 5×10^6 *Candida* in 20 µl of phosphate-buffered saline (PBS). (**A**) Immunohistochemistry staining of vaginal sections with antibodies to SIGLEC15. (**B**) *Siglec15* expression by real-time reverse transcription polymerase chain reaction (RT-PCR) relative to uninfected mice. (**C** to **F**) Infected mice were intravaginally treated with unmodified siRNA against murine *Siglec15* (si*Siglec15*; 10 µg/kg) or equivalent doses of nonspecific duplex scrambled control siRNA. Intravaginal siRNA was given once the day before infection and every 2 days thereafter until sacrifice at 3, 7, or 14 dpi. One-way analysis of variance (ANOVA) with Bonferroni correction is shown. (C to E) Mice were assessed for (**C**) vaginal fungal burden [$\log_{10}$ colony-forming units (CFUs)/100 µl vaginal fluid (VF) ± SD], (D) PMN recruitment in vaginal fluid (×100 magnification), and (E) vaginal pathology as shown by periodic acid–Schiff staining of sections. (F) IL-1β/IL-1Ra concentrations (pg/mg, cytokine/total proteins for each sample) in vaginal fluid. (**G**) *Nlrp3*, *Nlrc4* gene expression in ex vivo murine vaginas by real-time RT-PCR. (**H**) IL-6, IL-17A, IL-22, and IL-10 concentrations (pg/mg, cytokine/total proteins for each sample) in vaginal fluid. Scale bars, 200 µm. Data are presented as means ± SD. Two-way ANOVA with Bonferroni post hoc test was used unless otherwise indicated. *$P < 0.05$, **$P < 0.001$, ***$P < 0.001$, and ****$P < 0.0001$.

Jaeger *et al.*, *Sci. Transl. Med.* **11**, eaar3558 (2019)   12 June 2019

7 of 12

dysregulated in patients with RVVC. Among these, cell morphogenesis, cell adhesion, and cellular metabolism showed the strongest association with disease susceptibility. In addition, a complementary approach using genetic and functional immune data identified a number of susceptibility genes in immune system, among them *SIGLEC15*, a recognition receptor for *C. albicans*. Both in vitro and in vivo validation studies supported an important role of SIGLEC15 for RVVC and also suggested that immune hyperresponsiveness of individuals bearing the risk genetic variants is responsible for the clinical symptoms, as supported by earlier studies (6, 20, 37, 38).

There are limitations to our study, one of these being the relatively small sample size of the cohorts. One reason for this was the strict inclusion criteria for RVVC in our study: only patients with both clinical and microbiologically confirmed candidiasis were included. Although this approach reduced the number of patients in the cohorts, we believe that it has improved the likelihood of a positive result by limiting the "background noise" in the genetic signal. We also used additional approaches to compensate for the small cohorts: In the two independent patient cohorts, we used several independent types of analyses to assess both rare and common polymorphisms. We combined data with a functional assessment of immune function in a third cohort of individuals and in human primary cells and murine experimental studies to confirm the important role of SIGLEC15 as a receptor involved in host defense against *C. albicans*. Future approaches will need to further investigate the exact mechanistic cause of the association. Another limitation of our study is the partial coverage of the entire genome obtained by WES, thus providing limited information about noncoding regions. Last, only a limited number of pathways have been functionally validated here, and future studies should expand the functional assays.

The Siglec family of receptors recognizes sialylated structures (30, 39), and this is presumably the ligand recognized in the fungal cell wall as well. Sialic acids have been previously identified on the surface of *C. albicans*, and treatment of the fungus with neuraminidase that cleaves sialic acid from glycans changed the induction of inflammatory mediators (33). SIGLEC15 has been shown to be present on several immune cells, such as macrophages and dendritic cells (40). *SIGLEC15* has also been shown to be constitutively expressed in osteoclasts, and mice lacking this receptor are prone to develop osteopetrosis due to imbalances in osteoclast differentiation processes (41, 42).

The polymorphism we identified in this study (rs2919643) at position 273 within the *SIGLEC15* gene causes a missense mutation from phenylalanine to leucine (F273L). A change from a large aromatic amino acid into a medium-size hydrophobic amino acid is likely to lead to structural consequences. A lysine residue is located at amino acid position 274 within the transmembrane domain, which is critical for signal transduction through the DAP12-Syk pathway (43). The amino acid change might influence signaling upon ligand binding to SIGLEC15, and this needs to be investigated in future studies. The functional consequence of the risk allele in this polymorphism is an increase in T cell cytokines such as IL-17A and IL-22: Through increased neutrophil recruitment and release of inflammatory mediators, this contributes to the pathophysiology of the disease. The current concept supported by a wealth of experimental studies is that hyperinflammation, rather than immunodeficiency, is the basis of the pathological effects of RVVC (6, 44, 45, 20, 37). Moreover, in transfection experiments using wild-type and mutated SIGLEC15 plasmids, we found the expression of IL-1β and NLRP3 to be increased in cells transfected with the mutant form of the receptor. IL-1β is known

to stimulate T cells to produce IL-17 and IL-22 and has previously been shown to be highly up-regulated in patients with RVVC (37).

In addition to the immunological genes and pathways identified here, one conclusion of our analyses of rare and common genetic variants is that not only immune-related mechanisms but also biological processes related to cell morphology and adhesion are important in RVVC susceptibility. Adhesion and colonization of epithelium by *Candida* is well known to be important for mucosal infections (46), and the specific genes and pathways that influence RVVC pathophysiology need to be further studied in the future. Furthermore, one interesting observation was the impact of genetic variants in genes involved in fatty acid oxidation in susceptibility to RVVC. Cellular metabolism has been recently identified as one of the main processes determining the function of immune cells, with glycolysis being strongly up-regulated during induction of inflammation, whereas fatty acid oxidation is crucial for the maintenance of the balance between inflammatory and regulatory T cells (47). Fatty acid oxidation is also essential for the function of regulatory T cells (48) and for induction of long-term CD8$^+$ T cell memory (49), mechanisms that may have important roles in maintaining a balanced mucosal antifungal response. A thorough investigation of the impact of potential dysregulation in immune cell metabolism in patients with RVVC is warranted, and future experiments should aim to dissect these pathways and the individual genes and their interactions in more detail.

In conclusion, combining genomic data with functional immune measurements provided a systems-based approach that allowed us to identify important pathophysiological pathways and potential therapeutic targets in patients with RVVC. Validating these findings in additional independent studies and approaching these targets for therapy are needed to fulfill the need for diagnostic and therapeutic methods for RVVC.

MATERIALS AND METHODS
Study design
The aim of the present study was to identify genes and pathways that are dysregulated in RVVC. We recruited two independent cohorts of patients with RVVC and healthy controls (NEC, n = 75 patients and n = 98 controls; SEC, n = 85 patients and n = 77 controls). Both cohorts consisted of patients that were positively culture proven in the given medical center at that time point. All groups underwent WES. At the same time, we functionally assessed the control group for their capacity to respond to *Candida* stimulation in vitro. Genetic variants that we found associated with the disease were overlapped with variants that led to a changed cytokine profile after stimulation with *C. albicans*, leading to the identification of pathways and genetic variants. Using different in vitro approaches such as cytokine measurements [enzyme-linked immunosorbent assay (ELISA)], flow cytometry, and gene silencing (siRNA), we evaluated the role of SIGLEC15 in the host immune response against *Candida*. All in vitro stimulation experiments were performed at least three times with blood from healthy volunteers and, if stated, with blood from patients with RVVC. Moreover, we designed a vector containing the genetic variant and tested it in an isogenic background. Last, we evaluated specific gene silencing in mice that were intravaginally infected with *Candida*. Healthy female mice aged 8 to 10 weeks were used throughout the experiments.

Patient and control cohorts
The inclusion of the North European RVVC cohort and controls took place from January 2009 to December 2013 at the Radboud University

Medical Center (Radboudumc) (Nijmegen, Netherlands) and Wayne State University School of Medicine (Detroit, USA). The South European RVVC cohort and controls were recruited at the Santa Maria della Misericordia Hospital (Perugia, Italy), as previously described (*37*). Inclusion was approved by the institutional review boards of all centers. All patients gave written informed consent, and the study was done in accordance with the Declaration of Helsinki. Included patients were diagnosed with at least three documented episodes of VVC in a year with the cause of disease microbiologically confirmed to be *C. albicans*. Recruitment took place during episodes of acute vaginitis or, if asymptomatic, while receiving maintenance fluconazole therapy. All controls were of European origin, asymptomatic, and without a history of *Candida* infections. All subjects gave written informed consent.

Whole-exome sequencing
WES of both cohorts was carried out at the Beijing Genome Institute (Beijing, China). Two different enrichment kits were used during the enrichment step: Agilent SureSelect version 2 (NEC) and version 4 (SEC). The genomic DNA samples were randomly fragmented by Covaris, and adapters were ligated to both ends of the resulting fragments. The adaptor-ligated templates were purified by the Agencourt AMPure SPRI beads and fragments with an insert size of about 176 base pairs (bp) were excised. Extracted DNA was amplified, purified, and hybridized to the SureSelect Biotinylated RNA Library (BAITS) for enrichment. The enriched libraries were loaded on a HiSeq 2000 platform for the high-throughput sequencing. Raw image files were processed by Illumina base-calling software 1.7 for base calling with default parameters, and the sequences of each individual were generated as 90-bp pair-end reads.

Bioinformatics
Sequence reads were aligned to the human genome (hg19) using BWA (0.7.0) (*50*). Potential PCR duplicates were removed using Picard version 1.84 (http://broadinstitute.github.oi/picard), and reads were locally realigned using the genome analysis toolkit (GATK) "IndelRealigner" tool. The base quality was recalibrated using the GATK "BaseRecalibration." Both single-nucleotide variants and small insertion/deletions were called using the GATK "UnifiedGenotyper" on the four samples simultaneously (*51*). All steps were performed in accordance with the GATK version 2.4.9 "Best Practices" (https://software.broadinstitute.org/gatk/) (*52, 53*). We filtered out all variants with a "quality by depth" lower than 60× and with less than 20× coverage. Variants were annotated using a custom-made pipeline described earlier (*54*). Annotations include RefSeq gene annotation, amino acid consequences, variants and frequencies from dbSNP135, variants and frequencies from the Radboudumc Human Genetics database, and disease information from Online Mendelian Inheritance in Man (OMIM) and The Human Gene Mutation Database (HGMD) (*55*). Samples with non-normal sequencing coverage or non-normal number of called variants were excluded. To calculate the coverage and the number of variants, we used in-house scripts and bedtools version 2.17.0 (*56*). Variants were not considered in further analyses if they were either marked as failed by the variant calling software [GATK; (*51*)], had a quality score lower than 50, a coverage depth greater than 200× or lower than 20×, or a variant call that failed in more than 20 samples (either reference or alternative).

A subset of FRVs was selected for the higher probability to have a highly penetrant and direct effect on RVVC risk (rare-variant common-phenotype hypothesis). FRVs were those variants that were predicted to affect the protein sequence, were rare within the study cohort (MAF < 5%), and had a low MAF (<1%) in any of the available public population databases [1000 Genomes Project (www.internationalgenome.org) and HGMD database (www.hgmd.cf.ac.uk/ac)]. Furthermore, their phyloP score had to be higher than two, meaning that the sequence has a high evolutionary conservation. A second subset of cSNPs was selected because of their potential low-penetrant effect (common-variant common-phenotype hypothesis). These variants had to be biallelic, frequent in the study sample (MAF of ≥5%), and in Hardy-Weinberg equilibrium.

SNP analysis
cSNP analyses were performed using the Fisher's exact test function implemented in PLINK software (version 1.07) to detect cSNPs with imbalanced number of alleles between cases and controls in the NEC and SEC cohorts. Subsequently, the overlap between the cSNPs and genes between the two cohorts was calculated. The rare-variant analysis focused on the genes enriched in FRVs in cases and being absent in controls. Ideally, candidate genes had multiple FRVs present in cases in both cohorts and none in any controls. We also explored a statistical association test for rare variants by recursively performing the Fisher's exact test for every gene, counting the FRVs between cases and controls and either within or without the gene, but without correcting for multiple testing. In theory, FRVs might act as a protecting factor by inactivating genes in controls; however, we only considered the more likely mechanism of gene inactivation in cases as predisposing factors.

cQTL analysis was performed by calculating the difference between pre- and post-stimulation concentrations of cytokines using the Mann-Whitney *U* test. All sample stimulation pairs that were positive for the negative RPMI stimulation were removed from the analysis. The cytokines used for the cQTL analysis were as follows: TNFα, IL-1β, and IL-6 (all after 24 hours) and IL-17A, IL-22, and IFNγ (after 7 days).

The functional characterization of gene sets was performed with the topGO package (R code can be downloaded from Bioconductor at www.bioconductor.org/packages/GEM/). In every analysis, the reference set contained all genes with at least one variant in the corresponding cohort. The statistical significance of overlapping gene sets was calculated with the hypergeometric test implemented in the R function ("phyper"). All analyses were performed with R version 3.2.5. When specified, the multiple test correction was performed by the calculation of FDR using the Benjamini-Hochberg method.

Power analysis
We performed a power analysis to estimate the smallest effect size detectable by statistical tests within this study. A power of 80% and a baseline α of 0.05 were considered. Alpha values were further corrected for the number of variants with the Bonferroni method. The statistical power of the common variant analysis was estimated with the procedure described by Purcell *et al.* (*57*), whereas that of rare variants was performed with the standard procedure for repetitions of two-by-two Fisher's exact tests. The power analysis was corrected with the Bonferroni method. The two cSNPs analyses were performed with allelic association tests on 28,328 (NEC) and 45,819 (SEC) variants, resulting in corrected alpha values of 1.7×10^{-6} (NEC) and 1.1×10^{-6} (SEC). With sample sizes of 170 (NEC, 75 cases and 95 controls) and 157 (SEC, 80 cases and 77 controls), minimum relative risks of 2.5 (NEC) and 2.65 (SEC) are most likely to be detectable. The rare variants burden test was performed on 48,308 variants, resulting in a corrected alpha value of 4×10^{-6}, that, with

Jaeger *et al.*, *Sci. Transl. Med.* **11**, eaar3558 (2019)   12 June 2019

9 of 12

the sample sizes of 327 (150 cases and 172 controls), is likely to detect the minimum relative risks of about 10 (2% versus 20%).

Isolation and stimulation of human PBMCs

Seventy-three healthy controls participated in an in vitro stimulation assay. EDTA venous blood was collected from these volunteers, and PBMCs were subsequently isolated as described earlier (58). Briefly, venous blood was drawn into 10-ml EDTA tubes (Monoject) from the cubital vein of healthy volunteers (with different genotypes for the variant in *SIGLEC15*). The PBMC fraction was obtained by density centrifugation of blood diluted 1:1 in PBS over Ficoll-Paque (Pharmacia Biotech). Afterward, cells were washed three times in PBS and resuspended in RPMI 1640 (Dutch modified) supplemented with gentamicin (50 mg/liter), 2 mM L-glutamine, and 1 mM pyruvate. Cells were counted in a Coulter Counter ZH (Beckman Coulter) and adjusted to 5×10^6 cells/ml. Red blood cells were lysed before counting using Zap-Oglobin reagent (Beckman Coulter). Subsequently, 100 μl of cells (5×10^6) were plated in round-bottom 96-wells plates (Costar, Corning) together with different stimuli (*C. albicans* yeast and hyphae) to a final volume of 200 μl. After 24 hours or 7 days (with 10% human serum) of incubation at 37°C and 5% CO_2, the plates were centrifuged at 1400g for 8 min, and the supernatant was collected and stored at −20°C until different cytokine assays were performed.

Fungal killing assays

Human PBMCs from 33 healthy volunteers were plated in 96-well flat bottom plates (5×10^5 cells per well). Live *Candida* (1×10^6 cells/ml) was added for 24 hours. After incubation, the well contents were serial diluted in sterile water and plated on Sabouraud agar for counting of CFUs after 24 hours of incubation.

Human PBMC *SIGLEC15* siRNA experiments

PBMCs derived from healthy volunteers (three independent experiments with at least two donors) were isolated as described before (58). For macrophage differentiation, 10% serum in the presence of 5 ng/ml was added for 6 days and refreshed on day 4. Adherent cells were counted, and 1.5×10^5 serum-derived macrophages were plated into 96-well flat bottom plates. After 2 hours of adherence, cells were subsequently transfected with either 25 nM *SIGLEC15* siRNA (on target) or scrambled (nontarget) control siRNA (SmartPool, Thermo Scientific) for 48 hours at 37°C. After centrifugation, the supernatant was discarded and the pellet was resuspended in 200 μl of stimuli containing either *C. albicans* yeast (1×10^6 cells/ml; UC820, heat killed) or RPMI 1640 for 24 hours. Subsequently, the supernatants were collected for cytokine analysis using ELISA for IL-6 and TNFα (Sanquin Research and R&D Systems). Silencing efficiency and *SIGLEC15* expression was tested using quantitative PCR and the following primer pairs:

Gene	Sequences
h*SIGLEC15*	Sense: CGCGGATCGTCAACATCTC
	αSense: GTTCGGCGGTCACTAGGTG
h*GAPDH*	Sense: AGGGGAGATTCAGTGTGGTG
	αSense: CGACCACTTTGTCAAGCTCA

Mice

Female C57BL/6 mice, 8 to 10 weeks of age, were purchased from Charles River and then bred under specific pathogen–free conditions at the Animal Facility of the University of Perugia, Perugia, Italy. Murine experiments were performed according to the Italian Approved Animal Welfare Authorization 360/2015-PR and Legislative decree 26/2014 regarding the animal license obtained by the Italian Ministry of Health lasting for 5 years (2015 to 2020). All efforts were made to minimize suffering.

Fungal strains and vaginal infection

In the VVC model, mice were injected subcutaneously with β-estradiol 17-valerate (0.1 mg/100 μl) (Sigma Chemical Co.) dissolved in sesame oil (Sigma-Aldrich) per mouse 48 hours before vaginal infection. Estrogen administration continued weekly until completion of the study to maintain pseudoestrus. The estrogen-treated mice were inoculated intravaginally with 20 μl of a PBS suspension of 5×10^6 viable *C. albicans* yeast from early stationary phase cultures (18 hours of culture at 36°C in Sabouraud dextrose agar with chloramphenicol plates, BD Diagnostics). The time course of infection was monitored in individual mice by culturing 100 μl of serially diluted (1:10) vaginal lavages on Sabouraud dextrose agar with chloramphenicol plates. Vaginal lavages were conducted using 100 μl of sterile PBS. CFUs were enumerated after incubation at 37°C for 24 hours and expressed as $\log_{10}$ CFUs/100 μl of lavage fluid. In addition, mice were assessed for PMN recruitment in vaginal fluids by cytospin preparations of the lavage fluids, which were stained with May-Grünwald-Giemsa and observed with a BX51 microscope equipped with a high-resolution DP71 camera (Olympus). Images were acquired with a 40× objective and processed using analySIS image software.

siRNA design and delivery

siRNA against murine *Siglec15* (MMC.RNAI.N001101038.1) was purchased from Integrated DNA Technologies (IDT) (TEMA). Each mouse was intravaginally treated with unmodified siRNA (10 μg/kg) or equivalent doses of scrambled control siRNA duplex in a volume of 20 μl of duplex buffer (IDT, TEMA). Intravaginal siRNA was given once on the day before infection and every 2 days after the infection until sacrifice at 3, 7, or 14 dpi. Assessment of SIGLEC15 silencing was performed by real-time RT-PCR and electrophoretic separation of amplified fragments. Amplification efficiencies were normalized against β-actin (fig. S8).

Statistical analysis

When comparing two treatment groups, or groups with different genotypes, we used Student's *t* tests or the Mann-Whitney *U* test depending on the distribution of the data. ANOVA with Bonferroni's adjustment was used to determine the statistical significance for multiple groups. For naïve versus infected mice at different days after infection, a one-way ANOVA was used. A two-way ANOVA with Bonferroni post hoc test was used to compare infected mice at different days after infection treated with siRNA or scrambled control. In vivo experiments used six mice per group. Data were analyzed by GraphPad Prism version 4.03 program (GraphPad Software). Data are presented as means + SEM for all bar graphs. $*P < 0.05$, $**P < 0.001$, $***P < 0.001$, and $****P < 0.0001$.

SUPPLEMENTARY MATERIALS

Jaeger *et al.*, *Sci. Transl. Med.* **11**, eaar3558 (2019)   12 June 2019

10 of 12

Fig. S2. Genomic region of *SIGLEC15* and surrounding genetic variants in linkage disequilibrium with rs2919643.

Fig. S3. Validation of rs2919643 in an independent cohort.

Fig. S4. Genotype stratification after stimulation with *S. aureus*.

Fig. S5. SIGLEC15 expression in different cell types.

Fig. S6. Binding of SIGLEC15 to different *Candida* strains.

Fig. S7. Flow cytometry of mouse cells after *Candida* stimulation.

Fig. S8. Silencing efficiency in mouse siRNA experiments.

Table S1. RVVC-associated cSNPs with concordant effects in both cohorts.

Table S2. GO characterization of genes that have cSNPs associated with RVVC in both cohorts.

Table S3. FRVs.

Table S4. Rare variants in both cohorts after multiple testing correction.

Table S5. GO characterization of genes associated with both RVVC risk and QTL.

Table S6. Proxies for rs2919643 in *SIGLEC15* and their degree of linkage disequilibrium.

REFERENCES AND NOTES

1. H. Wisplinghoff, T. Bischoff, S. M. Tallent, H. Seifert, R. P. Wenzel, M. B. Edmond, Nosocomial bloodstream infections in US hospitals: Analysis of 24,179 cases from a prospective nationwide surveillance study. *Clin. Infect. Dis.* **39**, 309–317 (2004).

2. V. J. Fraser, M. Jones, J. Dunkel, S. Storfer, G. Medoff, W. C. Dunagan, Candidemia in a tertiary care hospital: Epidemiology, risk factors, and predictors of mortality. *Clin. Infect. Dis.* **15**, 414–421 (1992).

3. I. Das, P. Nightingale, M. Patel, P. Jumaa, Epidemiology, clinical characteristics, and outcome of candidemia: Experience in a tertiary referral center in the UK. *Int. J. Infect. Dis.* **15**, e759–e763 (2011).

4. B. J. Kullberg, M. C. Arendrup, Invasive candidiasis. *N. Engl. J. Med.* **373**, 1445–1456 (2015).

5. B. Foxman, R. Muraglia, J. P. Dietz, J. D. Sobel, J. Wagner, Prevalence of recurrent vulvovaginal candidiasis in 5 European countries and the United States: Results from an internet panel survey. *J. Low. Genit. Tract Dis.* **17**, 340–345 (2013).

6. B. M. Peters, J. Yano, M. C. Noverr, P. L. Fidel Jr., *Candida* vaginitis: When opportunism knocks, the host responds. *PLOS Pathog.* **10**, e1003965 (2014).

7. J. D. Sobel, Pathogenesis and treatment of recurrent vulvovaginal candidiasis. *Clin. Infect. Dis.* **14** (suppl. 1), S148–S153 (1992).

8. P. L. Fidel Jr., History and update on host defense against vaginal candidiasis. *Am. J. Reprod. Immunol.* **57**, 2–12 (2007).

9. J. D. Sobel, S. Faro, R. W. Force, B. Foxman, W. J. Ledger, P. R. Nyirjesy, B. D. Reed, P. R. Summers, Vulvovaginal candidiasis: Epidemiologic, diagnostic, and therapeutic considerations. *Am. J. Obstet. Gynecol.* **178**, 203–211 (1998).

10. J. D. Sobel, Vulvovaginal candidosis. *Lancet* **369**, 1961–1971 (2007).

11. B. Foxman, R. Barlow, H. D'Arcy, B. Gillespie, J. D. Sobel, *Candida vaginitis*: Self-reported incidence and associated costs. *Sex. Transm. Dis.* **27**, 230–235 (2000).

12. J. D. Sobel, Pathogenesis and epidemiology of vulvovaginal candidiasis. *Ann. N. Y. Acad. Sci.* **544**, 547–557 (1988).

13. D. W. Denning, M. Kneale, J. D. Sobel, R. Rautemaa-Richardson, Global burden of recurrent vulvovaginal candidiasis: A systematic review. *Lancet Infect. Dis.* **18**, e339–e347 (2018).

14. J. D. Sobel, H. C. Wiesenfeld, M. Martens, P. Danna, T. M. Hooton, A. Rompalo, M. Sperling, C. Livengood III, B. Horowitz, J. Von Thron, L. Edwards, H. Panzer, T.-C. Chu, Maintenance fluconazole therapy for recurrent vulvovaginal candidiasis. *N. Engl. J. Med.* **351**, 876–883 (2004).

15. J. D. Sobel, Epidemiology and pathogenesis of recurrent vulvovaginal candidiasis. *Am. J. Obstet. Gynecol.* **152**, 924–935 (1985).

16. H. L. Kent, Epidemiology of vaginitis. *Am. J. Obstet. Gynecol.* **165**, 1168–1176 (1991).

17. J. D. Sobel, Candidal vulvovaginitis. *Clin. Obstet. Gynecol.* **36**, 153–165 (1993).

18. L. S. Gunther, H. P. Martins, F. Gimenes, A. L. Abreu, M. E. Consolaro, T. I. Svidzinski, Prevalence of *Candida albicans* and non-*albicans* isolates from vaginal secretions: Comparative evaluation of colonization, vaginal candidiasis and recurrent vaginal candidiasis in diabetic and non-diabetic women. *Sao Paulo Med. J.* **132**, 116–120 (2014).

19. I. Amouri, H. Sellami, S. Abbes, I. Hadrich, N. Mahfoudh, H. Makni, A. Ayadi, Microsatellite analysis of *Candida* isolates from recurrent vulvovaginal candidiasis. *J. Med. Microbiol.* **61**, 1091–1096 (2012).

20. P. L. Fidel Jr., M. Barousse, T. Espinosa, M. Ficarra, J. Sturtevant, D. H. Martin, A. J. Quayle, K. Dunlap, An intravaginal live *Candida* challenge in humans leads to new hypotheses for the immunopathogenesis of vulvovaginal candidiasis. *Infect. Immun.* **72**, 2939–2946 (2004).

21. B. Ferwerda, G. Ferwerda, T. S. Plantinga, J. A. Willment, A. B. van Spriel, H. Venselaar, C. C. Elbers, M. D. Johnson, A. Cambi, C. Huysamen, L. Jacobs, T. Jansen, K. Verheijen, L. Masthoff, S. A. Morré, G. Vriend, D. L. Williams, J. R. Perfect, L. A. B. Joosten, C. Wijmenga, J. W. M. van der Meer, G. J. Adema, B. J. Kullberg, G. D. Brown, M. G. Netea, Human dectin-1 deficiency and mucocutaneous fungal infections. *N. Engl. J. Med.* **361**, 1760–1767 (2009).

22. A. Lev-Sagie, D. Prus, I. M. Linhares, Y. Lavy, W. J. Ledger, S. S. Witkin, Polymorphism in a gene coding for the inflammasome component NALP3 and recurrent vulvovaginal candidiasis in women with vulvar vestibulitis syndrome. *Am. J. Obstet. Gynecol.* **200**, 303.e1–303.e6 (2009).

23. O. Babula, G. Lazdane, J. Kroica, W. J. Ledger, S. S. Witkin, Relation between recurrent vulvovaginal candidiasis, vaginal concentrations of mannose-binding lectin, and a mannose-binding lectin gene polymorphism in Latvian women. *Clin. Infect. Dis.* **37**, 733–737 (2003).

24. G. G. G. Donders, O. Babula, G. Bellen, I. M. Linhares, S. S. Witkin, Mannose-binding lectin gene polymorphism and resistance to therapy in women with recurrent vulvovaginal candidiasis. *BJOG* **115**, 1225–1231 (2008).

25. F. Liu, Q. Liao, Z. Liu, Mannose-binding lectin and vulvovaginal candidiasis. *Int. J. Gynaecol. Obstet.* **92**, 43–47 (2006).

26. P. C. Giraldo, O. Babula, A. K. Goncalves, I. M. Linhares, R. L. Amaral, W. J. Ledger, S. S. Witkin, Mannose-binding lectin gene polymorphism, vulvovaginal candidiasis, and bacterial vaginosis. *Obstet. Gynecol.* **109**, 1123–1128 (2007).

27. M. D. K. H. Wojitani, L. M. de Aguiar, E. C. Baracat, I. M. Linhares, Association between mannose-binding lectin and interleukin-1 receptor antagonist gene polymorphisms and recurrent vulvovaginal candidiasis. *Arch. Gynecol. Obstet.* **285**, 149–153 (2012).

28. O. Babula, G. Lazdāne, J. Kroica, I. M. Linhares, W. J. Ledger, S. S. Witkin, Frequency of interleukin-4 (IL-4) -589 gene polymorphism and vaginal concentrations of IL-4, nitric oxide, and mannose-binding lectin in women with recurrent vulvovaginal candidiasis. *Clin. Infect. Dis.* **40**, 1258–1262 (2005).

29. S. P. Smeekens, F. L. van de Veerdonk, B. J. Kullberg, M. G. Netea, Genetic susceptibility to *Candida* infections. *EMBO Mol. Med.* **5**, 805–813 (2013).

30. M. S. Macauley, P. R. Crocker, J. C. Paulson, Siglec-mediated regulation of immune cell function in disease. *Nat. Rev. Immunol.* **14**, 653–666 (2014).

31. C. Büll, T. Heise, G. J. Adema, T. J. Boltje, Sialic acid mimetics to target the sialic acid-Siglec axis. *Trends Biochem. Sci.* **41**, 519–531 (2016).

32. Y.-C. Chang, V. Nizet, The interplay between Siglecs and sialylated pathogens. *Glycobiology* **24**, 818–825 (2014).

33. R. M. A. Soares, R. M. de A Soares, D. S. Alviano, J. Angluster, C. S. Alviano, L. R. Travassos, Identification of sialic acids on the cell surface of *Candida albicans*. *Biochim. Biophys. Acta* **1474**, 262–268 (2000).

34. V. Heinrich, T. Kamphans, J. Stange, D. Parkhomchuk, J. Hecht, T. Dickhaus, P. N. Robinson, P. M. Krawitz, Estimating exome genotyping accuracy by comparing to data from large scale sequencing projects. *Genome Med.* **5**, 69 (2013).

35. M. J. Machiela, S. J. Chanock, LDlink: A web-based application for exploring population-specific haplotype structure and linking correlated alleles of possible functional variants. *Bioinformatics* **31**, 3555–3557 (2015).

36. M. Jaeger, T. S. Plantinga, L. A. Joosten, B.-J. Kullberg, M. G. Netea, Genetic basis for recurrent vulvo-vaginal candidiasis. *Curr. Infect. Dis. Rep.* **15**, 136–142 (2013).

37. M. Jaeger, A. Carvalho, C. Cunha, T. S. Plantinga, F. van de Veerdonk, M. Puccetti, C. Galosi, L. A. B. Joosten, B. Dupont, B. J. Kullberg, J. D. Sobel, L. Romani, M. G. Netea, Association of a variable number tandem repeat in the *NLRP3* gene in women with susceptibility to RVVC. *Eur. J. Clin. Microbiol. Infect. Dis.* **35**, 797–801 (2016).

38. M. M. Barousse, T. Espinosa, K. Dunlap, P. L. Fidel Jr., Vaginal epithelial cell anti-*Candida albicans* activity is associated with protection against symptomatic vaginal candidiasis. *Infect. Immun.* **73**, 7765–7767 (2005).

39. P. R. Crocker, S. Mucklow, V. Bouckson, A. McWilliam, A. C. Willis, S. Gordon, G. Milon, S. Kelm, P. Bradfield, Sialoadhesin, a macrophage sialic acid binding receptor for haemopoietic cells with 17 immunoglobulin-like domains. *EMBO J.* **13**, 4490–4503 (1994).

40. T. Angata, Y. Tabuchi, K. Nakamura, M. Nakamura, Siglec-15: An immune system Siglec conserved throughout vertebrate evolution. *Glycobiology* **17**, 838–846 (2007).

41. Y. Hiruma, E. Tsuda, N. Maeda, A. Okada, N. Kabasawa, M. Miyamoto, H. Hattori, C. Fukuda, Impaired osteoclast differentiation and function and mild osteopetrosis development in Siglec-15-deficient mice. *Bone* **53**, 87–93 (2013).

42. Y. Kameda, M. Takahata, M. Komatsu, S. Mikuni, S. Hatakeyama, T. Shimizu, T. Angata, M. Kinjo, A. Minami, N. Iwasaki, Siglec-15 regulates osteoclast differentiation by modulating RANKL-induced phosphatidylinositol 3-kinase/Akt and Erk pathways in association with signaling Adaptor DAP12. *J. Bone Miner. Res.* **28**, 2463–2475 (2013).

43. R. Takamiya, K. Ohtsubo, S. Takamatsu, N. Taniguchi, T. Angata, The interaction between Siglec-15 and tumor-associated sialyl-Tn antigen enhances TGF-beta secretion from monocytes/macrophages through the DAP12-Syk pathway. *Glycobiology* **23**, 178–187 (2013).

44. J. Yano, E. Lilly, M. Barousse, P. L. Fidel Jr., Epithelial cell-derived S100 calcium-binding proteins as key mediators in the hallmark acute neutrophil response during *Candida* vaginitis. *Infect. Immun.* **78**, 5126–5137 (2010).

45. M. Borghi, A. De Luca, M. Puccetti, M. Jaeger, A. Mencacci, V. Oikonomou, M. Pariano, C. Garlanda, S. Moretti, A. Bartoli, J. Sobel, F. L. van de Veerdonk, C. A. Dinarello, M. G. Netea, L. Romani, Pathogenic NLRP3 inflammasome activity during *Candida* infection is negatively regulated by IL-22 via activation of NLRC4 and IL-1Ra. *Cell Host Microbe* **18**, 198–209 (2015).

Jaeger *et al.*, *Sci. Transl. Med.* **11**, eaar3558 (2019)   12 June 2019

11 of 12

46. J. R. Naglik, D. L. Moyes, B. Wächtler, B. Hube, *Candida albicans* interactions with epithelial cells and mucosal immunity. *Microbes Infect.* **13**, 963–976 (2011).

47. L. A. J. O'Neill, R. J. Kishton, J. Rathmell, A guide to immunometabolism for immunologists. *Nat. Rev. Immunol.* **16**, 553–565 (2016).

48. V. A. Gerriets, R. J. Kishton, A. G. Nichols, A. N. Macintyre, M. Inoue, O. Ilkayeva, P. S. Winter, X. Liu, B. Priyadharshini, M. E. Slawinska, L. Haeberli, C. Huck, L. A. Turka, K. C. Wood, L. P. Hale, P. A. Smith, M. A. Schneider, N. J. MacIver, J. W. Locasale, C. B. Newgard, M. L. Shinohara, J. C. Rathmell, Metabolic programming and PDHK1 control CD4⁺ T cell subsets and inflammation. *J. Clin. Invest.* **125**, 194–207 (2015).

49. G. J. W. van der Windt, D. O'Sullivan, B. Everts, S. C.-C. Huang, M. D. Buck, J. D. Curtis, C.-H. Chang, A. M. Smith, T. Ai, B. Faubert, R. G. Jones, E. J. Pearce, E. L. Pearce, CD8 memory T cells have a bioenergetic advantage that underlies their rapid recall ability. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 14336–14341 (2013).

50. H. Li, R. Durbin, Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).

51. A. McKenna, M. Hanna, E. Banks, A. Sivachenko, K. Cibulskis, A. Kernytsky, K. Garimella, D. Altshuler, S. Gabriel, M. Daly, M. A. DePristo, The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* **20**, 1297–1303 (2010).

52. M. A. DePristo, E. Banks, R. Poplin, K. V. Garimella, J. R. Maguire, C. Hartl, A. A. Philippakis, G. del Angel, M. A. Rivas, M. Hanna, A. McKenna, T. J. Fennell, A. M. Kernytsky, A. Y. Sivachenko, K. Cibulskis, S. B. Gabriel, D. Altshuler, M. J. Daly, A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat. Genet.* **43**, 491–498 (2011).

53. G. A. Van der Auwera, M. O. Carneiro, C. Hartl, R. Poplin, G. Del Angel, A. Levy-Moonshine, T. Jordan, K. Shakir, D. Roazen, J. Thibault, E. Banks, K. V. Garimella, D. Altshuler, S. Gabriel, M. A. DePristo, From FastQ data to high confidence variant calls: The Genome Analysis Toolkit best practices pipeline. *Curr. Protoc. Bioinformatics* **43**, 11.10.11–11.10.33 (2013).

54. C. Gilissen, H. H. Arts, A. Hoischen, L. Spruijt, D. A. Mans, P. Arts, B. van Lier, M. Steehouwer, J. van Reeuwijk, S. G. Kant, R. Roepman, N. V. A. M. Knoers, J. A. Veltman, H. G. Brunner, Exome sequencing identifies *WDR35* variants involved in Sensenbrenner syndrome. *Am. J. Hum. Genet.* **87**, 418–423 (2010).

55. P. D. Stenson, E. V. Ball, M. Mort, A. D. Phillips, J. A. Shiel, N. S. Thomas, S. Abeysinghe, M. Krawczak, D. N. Cooper, Human Gene Mutation Database (HGMD): 2003 update. *Hum. Mutat.* **21**, 577–581 (2003).

56. A. R. Quinlan, I. M. Hall, BEDTools: A flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841–842 (2010).

57. S. Purcell, S. S. Cherny, P. C. Sham, Genetic power calculator: Design of linkage and association genetic mapping studies of complex traits. *Bioinformatics* **19**, 149–150 (2003).

58. M. G. Netea, N. A. R. Gow, C. A. Munro, S. Bates, C. Collins, G. Ferwerda, R. P. Hobson, G. Bertram, H. B. Hughes, T. Jansen, L. Jacobs, E. T. Buurman, K. Gijzen, D. L. Williams, R. Torensma, A. McKinnon, D. M. MacCallum, F. C. Odds, J. W. M. Van der Meer, A. J. P. Brown, B. J. Kullberg, Immune sensing of *Candida albicans* requires cooperative recognition of mannans and glucans by lectin and Toll-like receptors. *J. Clin. Invest.* **116**, 1642–1650 (2006).

Citation: M. Jaeger, M. Pinelli, M. Borghi, C. Constantini, M. Dindo, L. van Emst, M. Puccetti, M. Pariano, I. Ricaño-Ponce, C. Büll, M. S. Gresnigt, X. Wang, J. Gutierrez Achury, C. W. M. Jacobs, N. Xu, M. Oosting, P. Arts, L. A. B. Joosten, F. L. van de Veerdonk, J. A. Veltman, J. ten Oever, B. J. Kullberg, M. Feng, G. J. Adema, C. Wijmenga, V. Kumar, J. Sobel, C. Gilissen, L. Romani, M. G. Netea, A systems genomics approach identifies *SIGLEC15* as a susceptibility factor in recurrent vulvovaginal candidiasis. *Sci. Transl. Med.* **11**, eaar3558 (2019).

Jaeger *et al.*, *Sci. Transl. Med.* **11**, eaar3558 (2019)   12 June 2019

12 of 12

Science Translational Medicine

A systems genomics approach identifies *SIGLEC15* as a susceptibility factor in recurrent vulvovaginal candidiasis

M. Jaeger, M. Pinelli, M. Borghi, C. Constantini, M. Dindo, L. van Emst, M. Puccetti, M. Pariano, I. Ricaño-Ponce, C. Büll, M. S. Gresnigt, X. Wang, J. Gutierrez Achury, C. W. M. Jacobs, N. Xu, M. Oosting, P. Arts, L. A. B. Joosten, F. L. van de Veerdonk, J. A. Veltman, J. ten Oever, B. J. Kullberg, M. Feng, G. J. Adema, C. Wijmenga, V. Kumar, J. Sobel, C. Gilissen, L. Romani and M. G. Netea

A genetic fungal risk factor

Although a number of women experience recurrent vulvovaginal candidiasis (RVVC), it is not well understood why they are prone to frequent infections. In addition, antifungals used to treat isolated infections can be inadequate in the recurrent setting, heightening the need to understand the etiology of RVVC. Jaeger *et al.* found that a polymorphism in lectin *SIGLEC15* associated with recurrent candidiasis in two independent patient cohorts. Stimulation with *Candida albicans* increased *SIGLEC15* expression and altered T cell cytokine production in human peripheral blood mononuclear cells and in a mouse model of RVVC. This work suggests that the *SIGLEC15* variant leads to an altered immune response in patients with RVVC.

ARTICLE TOOLS	http://stm.sciencemag.org/content/11/496/eaar3558
SUPPLEMENTARY MATERIALS	http://stm.sciencemag.org/content/suppl/2019/06/10/11.496.eaar3558.DC1
RELATED CONTENT	http://stm.sciencemag.org/content/scitransmed/10/429/eaag2616.full http://stm.sciencemag.org/content/scitransmed/11/482/eaat0852.full http://stm.sciencemag.org/content/scitransmed/8/322/322ra10.full
REFERENCES	This article cites 58 articles, 6 of which you can access for free http://stm.sciencemag.org/content/11/496/eaar3558#BIBL
PERMISSIONS	http://www.sciencemag.org/help/reprints-and-permissions

Use of this article is subject to the Terms of Service