The network interplay of interferon and toll-like receptor signaling pathways in the anti-
Candida immune response


aDepartment of Immunology, Institute of Biomedical Sciences, University of São Paulo, São Paulo, SP, Brazil.
bDepartment of Clinical and Toxicological Analyses, School of Pharmaceutical Sciences, University of São Paulo, São Paulo, SP, Brazil.
cInformation Systems, School of Arts, Sciences and Humanities, University of Sao Paulo, São Paulo, SP, Brazil.
dDepartment of Biostatistics and the Data Science Initiative at Brown University, Providence, RI, USA.
eDepartment of Pediatrics, University of Washington School of Medicine, and Seattle Children's Research Institute, Seattle, WA, USA.
fOsteoarthritis Research Program, Division of Orthopedic Surgery, Schroeder Arthritis Institute, UHN; Data Science Discovery Centre, Krembil Research Institute, UHN; Departments of Medical Biophysics and Computer Science, University of Toronto, Toronto, Canada.
gInstitute of Neuroimmunology, Slovak Academy of Sciences, Bratislava, Slovakia.
hNetwork of Immunity in Infection, Malignancy, and Autoimmunity (NIIMA), Universal Scientific Education and Research Network (USERN), São Paulo, SP, Brazil.

Correspondence to:
Ranieri Coelho Salgado
Department of Immunology
Institute of Biomedical Sciences - University of São Paulo
Lineu Prestes Avenue,1730, São Paulo, Brazil
Email: ranierics@usp.br
Phone: +55 11 30917435

Otavio Cabral-Marques, MSc, PhD
Department of Immunology
Institute of Biomedical Sciences - University of São Paulo
Lineu Prestes Avenue,1730, São Paulo, Brazil
Email: otavio.cmarques@usp.br
Phone: +55 11 974642022
ABSTRACT

Fungal infections represent a major global health problem that affects over a billion people and kills more than 1.5 million individuals annually. Here we employed an integrative approach to unravel the landscape of the human immune responses to Candida spp. (C. albicans and C. auris) by performing a meta-analysis of microarray, bulk, and single-cell RNA-sequencing (RNASeq) of blood transcriptome data. We identified that C. albicans activates a network interplay of signaling molecules commonly involved in both toll-like receptor (TLR) and interferon (IFN) signaling cascades. These molecules form a highly interconnected interferome network, which contains an immune overlap with the anti-viral responses. scRNAseq data confirmed that genes commonly identified by the three transcriptomic methods present a consistent upregulation pattern across innate immune and adaptive cells (CD4+, CD8+, and CD19+ lymphocytes). Thus, our results shed new lights on the molecular basis of immune response to Candida spp.
INTRODUCTION

Fungal infections, including the emergence of new fungal pathogens highly resistant to antifungal drugs, represent a major global health issue\(^1\)-\(^5\). Fungal infections affect over a billion people worldwide and kill more than 1.5 million individuals annually. This mortality rate is similar to that of tuberculosis and 3-fold higher than that of malaria\(^6\),\(^7\). Invasive candidiasis (IC) is the most common fungal disease, affecting approximately 250,000 people annually and causing more than 50,000 deaths\(^8\),\(^9\). The increasing number of patients with human immunodeficiency virus (HIV) infection, malignancies, inborn errors of immunity (IEI), autoimmune diseases (receiving immunosuppressive treatment), and hematopoietic stem cell or organ transplant recipients contributes to this high frequency of individuals susceptible to life-threatening fungal pathogens\(^10\),\(^11\). The severity of fungal infections in healthy subjects ranges from asymptomatic or mild mucocutaneous manifestations to life-threatening systemic infections. These data reinforce the critical role of interactions between the microbial and host immune system in the outcome of infection\(^7\),\(^12\), which needs to be further investigated to find new therapies to reduce morbidity and mortality caused by Candida infections\(^13\),\(^14\).

Linear and mechanistic approaches have elegantly demonstrated that the anti-fungal immune response involves the appropriate recognition of pathogen-associated molecular patterns (PAMPs) by different pattern recognition receptors (PRRs) expressed on the cell membrane such as C-type lectin receptors (CLR: dectin-1, dectin-2, and CD209), scavenger receptors (CD36), and toll-like receptors (TLR: e.g., TLR2 and 4). Intracellular PRRs including RIG-I-like receptors (RLR: melanoma differentiation-associated protein 5 or MDA5), TLRs (e.g., TLR3 and TLR9), and NOD-like receptors (NLR: nucleotide-binding
oligomerization domain-containing protein or NOD1/2, NOD-, LRR- and pyrin domain-containing 3 or NLRP3) are also relevant and expressed by antigen-presenting cells and phagocytes, which bind to well-known ligands\textsuperscript{15-17}. Activation of PRRs induces several signaling events such as the canonical Nuclear factor (NF)-κB pathway\textsuperscript{18} that trigger effector anti-fungal mechanisms such as phagocytosis, production of reactive oxygen species (ROS)\textsuperscript{19}, degranulation, and neutrophil extracellular traps (NETs)\textsuperscript{20,21}. Simultaneously, PRRs promote the production of key inflammatory cytokines such as tumor necrosis factor (TNF)-α, Interleukin (IL)-1β, IL-6, IL-17, type I Interferons (IFNs [IFN-α/β]), and the IL-12/IFN-γ axis\textsuperscript{17,22-24}, which shape and instruct immune cells\textsuperscript{25}.

However, the landscape of anti-fungal molecules in a holistic and integrative way remains to be provided. To reach this goal, we performed a meta-analysis of blood transcriptome data of microarray, bulk, and single-cell RNA-sequencing (scRNAseq) to unravel the landscape of the human immune responses to Candida spp. (C. albicans and C. auris). This integrative approach revealed a previous unnoted network interplay of type 1 interferon and toll-like receptor signaling in the anti-candida immune response.
RESULTS

_C. albicans_ activates signaling molecules commonly involved in both toll-like receptor and interferon signaling cascades.

We surveyed published RNAseq datasets and found a total of 8 datasets related to the human immune response to _Candida spp.,_ being 5 of microarray, 2 bulk RNAseq, and one scRNAseq (further details in Methods section). We explored the scRNAseq by performing over representation analysis (ORA) of differentially expressed genes (DEGs) from innate immune (monocytes, natural killer, and plasmacytoid dendritic cells) and adaptive cells (CD4+, CD8+, and CD19+ lymphocytes), which were assigned to clusters as previously described (Fig. 1a). TCD4+ cells were most prevalent in the peripheral blood mononuclear cells (PBMCs). We found similar distribution pattern when assigning these cell types to clusters in resting and after _C. albicans_ activation (Fig. 1b-c). A total of 6722 DEGs (Suppl. Table 1) were present in these clusters when comparing _C. albicans_-activated to resting cells. Enriched pathways associated with the immune response to _C. albicans_ are shown in Fig. 1d while all enriched categories are present in Suppl. Table 2. Among them are 72 and 99 DEGs belonging to TLR and IFN (both type I and type II) signaling cascades, suggesting an interplay previously noted only in response to lipopolysaccharide (LPS)\(^26\). Across these DEGs there were 7 in common between both pathways. We also searched in literature and found more 55 DEGs involved in both TLR and IFN signaling cascades (Suppl. Table 3); totaling, thus 62 associated DEGs, six of them illustrated in Fig. 1e-f.

Modular gene co-expression analysis reveals an interplay of TLR- and IFN-associated genes
We next performed modular gene co-expression analysis\textsuperscript{27}, to better understand the interplay between TLR and IFN signaling cascades. For this analysis, we used the microarray dataset from Smeekens et al. (GSE42606)\textsuperscript{28}, the unique public dataset available containing more than 15 samples per group (30 resting and 24 \textit{C. albicans}-activated samples), which is required to obtain biologically meaningful modular networks\textsuperscript{29}. Modular gene co-expression analysis using CEMiT tool\textsuperscript{30} identified thirteen enriched co-expression modules from the total expressed genes by PBMCs (which contain lymphocyte subpopulations, monocytes, and dendritic cells). Among these modules, 12 were significantly enriched (9 downregulated and 3 upregulated) in response to \textit{C. albicans} infection (Fig. 2a). Of note, modules M1 and M2 indicate gene co-expression and upregulation of IFN and interleukin signaling with TLR cascades (Fig. 2b-e).

Based on the results obtained by the modular co-expression analysis, we dissected the significantly enriched pathways from differentially expressed genes (DEGs) induced by \textit{C. albicans}\textsuperscript{28}. In agreement with the topological results obtained using CEMiT tool, ORA of DEGs using the ClusterProfiler tool\textsuperscript{31} pinpointed different clusters related to the activation of the TLR and IFN signaling cascades (Suppl. Fig 1a-b). IFN signaling was the most enriched pathway modulated by \textit{C. albicans}. The relationship between the 30 most enriched pathways and their associated genes is shown in a network view (Suppl. Fig 1c) while the entire list of all enriched pathways is summarized in Suppl. Table 4. \textit{C. albicans} activation significantly enriched several TLR signaling events such as TLR4, TLR3, TLR7/8, and TLR9 as well as MyD88/TIR-domain-containing adapter-inducing interferon-\textbeta\ (TRIF)/TIR Domain Containing Adaptor Protein (TIRAP) cascades, and TRAF6-mediated NF-\kappaB activation. ORA also indicated that \textit{C. albicans} activates chemokine (G protein-coupled receptors [GPCR]
ligand binding) and cytokine signaling pathways (IL-10, IL-3 and IL4), IFN-α/β signaling, Interferon-stimulated gene 15 (ISG15) antiviral mechanism, TNF Receptor Associated Factor 3 (TRAF3)-dependent IRF activation, DExD/H-Box Helicase 58 (DDX58)/Interferon Induced with Helicase C Domain 1 (IFIH1)-mediated induction of IFN-α/β, and regulation of type I and II IFN among the IFN signaling events (Suppl. Fig 1a and c; Suppl. Table 4).

C. albicans infection activates common TLR- and IFN-associated genes in peripheral blood leukocytes

We further investigated which DEGs and signaling pathways are consistently activated by C. albicans in peripheral blood leukocytes such as PBMCs (datasets GSE42606 and GSE154911) and peripheral white blood cells (WBCs, datasets GSE65088 and GSE114180) throughout all publicly available datasets. WBCs contain PBMCs (lymphocytes 20-45% and monocytes 2-10%) and granulocytes (neutrophils: 50-70%; basophils: 0-1%; and eosinophils: 1-5%)\(^{32}\). A meta-analysis of WBCs and PBMCs gene expression datasets using P-value combination method revealed among the meta-significant genes 44 commonly activated DEGs (40 upregulated and 4 downregulated) (Fig 3a, Suppl. Table 5). According to the cell population, these DEGs form well-defined hierarchical clusters, i.e., PBMCs datasets present a closer expression pattern among them as well as the WBCs datasets, when we compare both regulation and significance (Fig 3b). Enrichment analyses using EnrichR of these 44 genes revealed 87 significantly affected pathways (Suppl. Table 6), including TLR and IFN-α/β signaling pathways (Fig. 3c). Furthermore, these 44 DEGs also enrich other related interleukin signaling pathways such as JAK-STAT, IL-12, IL-17, IL-23, TNF, and chemokines (GPCR ligand binding) and PRRs, including RIG-I like receptor and NOD
signaling. Multi-study Factor Analysis of eligible datasets (WBCs: GSE65088 and PBMCs: GSE42606; those with minimal number of samples required for this analyses) identified two common latent factors with high loadings, while specific latent factors showed low loadings across these studies. Thus, strengthening the biological relevance of these 44 common genes (Fig. 3d, Suppl. Table 7).

*C. albicans* activates common TLR and IFN signaling pathways across different layers of immunity

Subsequently, we added monocyte-derived dendritic cells (moDCs) datasets (GSE77969, E-MTAB-135, E-MTAB-751) into our integrative analysis. moDCs are known to be essential players of anti-fungal immunity, bridging the immune system's innate and adaptive arms. We searched for genes commonly regulated by *C. albicans* in transcriptomes of WBCs, PBMCs, and moDCs, in resting or *C. albicans* activation conditions. Intersection analyses performed according to cell population identified 123, 223, and 57 common DEGs among WBCs, PBMCs, and moDCs datasets, respectively (Fig. 4a-c). However, only 2 common DEGs were present across all seven datasets (Fig. 4d), which by themselves do not significantly enrich signaling pathways. We then asked if DEGs from each dataset enrich common signaling biological processes among all studies. Gene Ontology (GO) analysis using ClusterProfiler analysis revealed 173 common biological processes (Suppl. Table 8). We found several molecules/pathways essential for the anti-fungal immune response. Among them, there is a cluster of IFN-γ, and NF-kB signaling, and a previously described overlap with the immune response to virus (IFN-α/β) (Fig. 4e). Additional ORA of DEGs involved in this cluster found significant enrichment of signaling cascades of single TLRs (TLR2, TLR3, TLR4,
TLR5, TLR9, TLR9, and TLR10), TLR heterodimers (TLR1/TLR2, TLR2/TLR6, TLR7/8), and TLR adapter molecules (MyD88/TIRAP, TRAF6, TRIF) as well as several interleukin signaling pathways such as IL-1, IL-4/IL-13, IL-6, IL-10, IL-17, and IFN-α/β (Fig. 4f). We found 1096 DEGs (Suppl. Table 10) affecting common biological processes among WBCs, PBMCs and moDCs. Fig. 4g shows the interactome obtained from some of these 1096 DEGs and enriched signaling cascades (Suppl. Table 11), thus highlighting the association of TLR- and IFN-signaling cascades, which were consistently enriched during our analyses. These 1096 DEGs also enrich other PRR and Interleukin signaling pathways such as CLRs (dectin-1), NLRs (NOD1/2), pro- (IL-1, IL6, IL-17, IL-12), anti-inflammatory (IL-10), and T helper 2 (IL-4 and IL-13) cytokines. This immunological balance between a pro-and anti-inflammatory event is crucial for the proper control of fungal infections while maintaining immune homeostasis35,36.

**C. albicans infection increases the correlation between TLR- and IFN-associated genes**

After verifying TLR and type I and II IFN signaling cascades’ consistency, we assessed the degree of association between these two variables during the immune response to *C. albicans*. Due to minimum sample size requirement37, we selected TLR and IFN-associated genes present in the PBMCs transcriptome data from Smeekens et al. (GSE42606)28. This dataset contains 45 and 14 TLR- and IFN-associated DEGs modulated by *C. albicans* when compared to the resting group. *C. albicans* infection increased mainly positive correlations between TLR- and IFN-associated DEGs (Fig. 5a-b). We performed Canonical Correlation Analysis (CCA) to further assess the association’s strength between TLR and IFN DEGs. CCA is a generic parametric model used to quantify relationships between two groups of
interrelated and interdependent variables\(^{38}\). This approach unveiled a pair of canonical variates (x-CV1 and y-CV1) highlighting the strong association between most of TLR - and IFN -associated DEGs in both resting and \(C.\) \textit{albicans}-infected PBMCs (Fig. 5c), although, they are able to stratify these conditions (Fig. 5d-e).

The multidrug-resistant \(C.\) \textit{auris} also induces the interplay between TLR and IFN signaling pathways

We asked if only \(C.\) \textit{albicans} induces the association between TLR- and IFN-associated genes or also the multidrug-resistant \(C.\) \textit{auris}\(^{39,40}\). We used the unique publicly available dataset analyzing the immune response to \(C.\) \textit{auris} and \(C.\) \textit{albicans} (GSE154911). Similar to \(C.\) \textit{albicans} activation, ORA of DEGs induced by \(C.\) \textit{auris} included TLR and IFN signaling cascades among the 30 most enriched pathways (Fig. 6a-b). \(C.\) \textit{albicans} and \(C.\) \textit{auris} similarly modulated DEGs' levels involved in TLR signaling, including NF-κB1, NF-κB2, JUN, and DUSP4, as well as IFN signaling such as IRFs, GBP1, SOCS1, ISG20, TRIM, and IFIT3 (Fig. 6c). When we compared the DEGs induced by \(C.\) \textit{auris} with those enriching common pathways among all datasets (1096 DEGs, Suppl. Table 10) assessing the immune response to \(C.\) \textit{albicans}, we identified 237 common DEGs (Fig. 6d). ORA of these common DEGs indicates that the interplay between TLR and IFN signaling cascades is a consistent immunologic feature in response to these two Candida species (Fig. 6e).

Inborn errors of immunity (IEI) corroborate the interplay between TLR and IFN signaling cascades
Finally, we aimed to evaluate the potential clinical and translational relevance of the TLR- and IFN-associated genes and molecular pathways consistently modulated by Candida. Therefore, we searched for IEI associated genes that are known to increase human susceptibility to fungal infections. So far, mutations in 100 genes have been associated with IEI that cause increased susceptibility to candidiasis and often other clinical manifestations. We compared them with the 1096 genes (Suppl. Table 10, i.e., those enriching the common biological processes activated by Candida (Fig. 4e). These 1096 genes encode molecules present in different cellular compartments such as extracellular regions, organelles, and nuclei and those forming macromolecular complexes. Together, they form a highly interconnected physical protein-protein interaction network (Fig. 7a), which contains several hubs (Fig. 7b), here defined as those having more than or equal to 200 interaction partners. Of note, 34 genes associated with IEI are also present across the studies. Meanwhile, although 66 genes associated with IEI are not identified in the datasets, these genes are highly connected with the other DEGs in this network. Furthermore, the 1096 DEGs mostly contain Type I and II IFN-associated genes, being 868 in total (Fig. 7c, Suppl. Table 12).

The 34 IEI associated genes present in this network underly 7 groups of IEI including congenital defects of phagocytes, defects of intrinsic and innate immunity, predominantly antibody deficiencies, and diseases of immune dysregulation, as defined by the International Union of Immunological Societies Expert Committee (IUIS) (Suppl. Fig. 2a). Notably, among the hubs are STAT1, STAT3, NFKBIA (IκBa), and NFKB1, which are well known to be associated with TLR and IFN signaling pathways. ORA of these 34 genes indicates that in addition to dectin-1 and NLRs signaling, they mostly enrich both type I and II IFN and
several TLR signaling pathways (TLR1/2, TLR2/6, MyD88, and TRAF6-mediated NF-κB activation) (Suppl. Fig. 2b-c).

**Common TLR-and IFN-associated DEGs and signaling pathways across microarray, bulk, and single-cell RNA-seq datasets**

Finally, we revisited the scRNAseq data and found that 11 TLR- and 23 IFN-associated DEGs are also among those WBCs, PBMCs and moDCs DEGs identified by microarray and bulk RNAseq datasets (Suppl. Table 13). Thus, indicating that the network interplay of TLR- and IFN-associated DEGs are not a particular feature of a specific leukocyte cell population since *C. albicans* systemically activated this network throughout the different innate (monocytes, natural killer, and plasmacytoid dendritic cells) and adaptive cells (CD4+, CD8+, and CD19+ lymphocytes) identified by the scRNAseq dataset. **Fig. 8a-b** illustrate these 34 common genes across the leukocytes subpopulations and those present in the WBCs, PBMCs, and moDCs datasets (**Fig. 3a-c**). Hierarchical clustering of common enriched pathways across the cell subpopulations identified by scRNAseq showed a similar up-regulation pattern of TLR- and IFN-associated signaling pathways, forming clusters (**Fig. 8c**), as seen by microarray and bulk RNAseq.

**DISCUSSION**

The association between PRR activation and cytokine production by immune cells is important for an adequate immune response against pathogens and has been abundantly investigated by linear approaches or strategies designed to identify the anti-fungal
transcriptomic signature\textsuperscript{12,24,39,50}. For instance, several immunologic molecules and pathways such as those triggered by TLR and IFN, which induce the generation of T cell subpopulations (e.g., T helper 1 [Th1], Th17, and T regulatory [T reg] cells) have been successfully characterized by individual studies and mechanistic approaches\textsuperscript{17}. However, until now, there was no system immunology study to holistically understand the anti-fungal immune responses. Our approach integrates dispersed transcriptomic datasets that investigated the immune response against \textit{C. albicans} and \textit{C. auris}, indicating that the anti-candida immune responses are marked by a previously uncharacterized intricated interferome chain, interconnecting PRR (e.g., CLRs, TLRs, and NLRs) and interleukin (e.g., IFN, TNF, and IL-10) cascades. This immune network is hallmarked by dynamic and consistent crosstalks between the network interplay of TLR and IFN signaling pathways. Besides, we show that there is a consistent overlap between the antiviral and antifungal immune responses, which supports the previously reported pivotal role of IFN type I in the immune response against \textit{C. albicans}\textsuperscript{24}. Notably, this immunologic overlap might be not restricted to viral infection. For instance, studies need to be performed to investigate its extension to the anti-\textit{Mycobacteria tuberculosis} immune responses. The host’s protection against this intracellular bacterium relies not only on an IFN-\textgreek{gamma} centered phenomenon but also requires the synergism of type I IFNs and other cytokines such as IL-17\textsuperscript{51}. Thus, our results strongly suggest that the immune system employs a multitude of molecules, working as a “social” network, in which cells effectively collaborate and communicate to maintain immunologic homeostasis\textsuperscript{52,53}.

Our integrative and systems immunology approach provides a transcriptomic landscape of the anti-candida defenses that will contribute to better understand the host
immunological dynamics initiated against these fungal pathogens. Since we currently are confronted with increasing numbers of invasive fungal infections, in part due to the emergence of anti-fungal drug resistance, it is imperative to address this global health problem\textsuperscript{2,40,54}. Diseases caused by viruses and bacteria have been recognized as important public health issues for centuries, while fungal infections have historically been neglected\textsuperscript{55}. Numerous transcriptomic, epigenomic, and proteomic data are available that investigate the immune response against viruses and bacteria using public databases\textsuperscript{56–58}. However, most investigations addressing the anti-fungal immune response employ linear and mechanistic approaches\textsuperscript{15} and there are very limited numbers of publicly available transcriptomic datasets of human immune responses to fungal infections. Most of these studies focus on the transcriptomic response of human immune cells activated with \textit{C. albicans} but only one dataset explores the immune response to \textit{C. auris}, one of the most critical emerging fungal pathogen\textsuperscript{39}. The integrative approach we employed was designed to obtain a comprehensive understanding of the anti-fungal immune response.

The results of our study provide complementary arguments for linear and mechanistic strategies that confirm a dynamic interplay between TLR and type I and II IFN-associated molecules\textsuperscript{59}. It has been suggested that different TLRs synergistically activate immune cells to, for instance, induce the expression of several proinflammatory molecules through the cooperation of NF-\(\kappa\)B, IRF, STAT, MAPK, ITAM, and PI3K signaling pathways\textsuperscript{60–62}. On the one hand, TLR-induced NF-\(\kappa\)B signaling promotes the production of several key cytokines including IFNs that activate STAT1-mediated signaling pathway\textsuperscript{63}. On the other hand, IFN-\(\gamma\) increases the expression of genes encoding TLRs\textsuperscript{64–67}. IFNs also potentialize TLR-induced gene transcription by creating a primed chromatin environment by histone acetylation that
allows sustained occupancy of transcription factors STAT1 and IRF-1 at promtors and enhancers at the *TNF, IL-6*, and *IL12B* loci. Thus, our phenomenological study confirms these previously reported mechanistic studies and provides new insights into the molecular network of TLR and IFN signaling pathways in anti-Candida immune response. These networks also need to be investigated in other mycoses (paracoccidioidomycosis, histoplasmosis, and cryptococcosis) and other neglected diseases (Dengue, Zika, leishmaniasis, and Chagas disease) occurring in developing countries. The TLR and IFN interactome involve more complex events than previously thought, demanding further bottom-up and top-down systems immunology investigations.

Our conclusions are based on the integration of publicly available human transcriptomes that identified common DEGs, as well as biological processes and signaling pathways consistently modulated across several leukocyte subpopulations in response to fungal pathogens (*C. albicans* and *C. auris*). Among these DEGs, we highlight those involved in IFN-α/β (e.g., ISGs, IRFs, SOCS, and GBP1), TLR3,4,7/8,9, and TRAF-mediated NF-κB signaling cascades. The correlation levels of DEGs involved in these signature clusters increased upon stimulation with *C. albicans*. Of note, among the consistently identified DEGs are those previously associated with IEI that increase the host susceptibility to fungal infections such as those causing chronic mucocutaneous candidiasis. Besides, these DEGs are also involved with immunological pathways related to the development of IEI phenocopies such as those targeted by anti-IL-17 or anti-IL17RA autoantibodies that result in increased susceptibility to Candida spp. infections. Because the outcome of fungal infections depends primarily on the host immune response, it is most relevant to review those IEIs that predispose to Candida infections. IEIs represent an essential research field that...
investigates natural human susceptibility models to infection, often revealing the non-redundant role of genes involved in immunologic homeostasis\textsuperscript{73-75}. Of the 416 molecular defined IEI recently summarized by the expert committee of the IUIS, more than 20 syndromes were recognized to be associated with susceptibility to fungal infections\textsuperscript{42}. This list of genes associated with increased risk of fungal infections includes genes regulating signaling via the IL-2 receptor, via NF-kB activation, IFN induced signaling, activation of STATs, and TLR signaling. Thus, these observations support the relevance of the interactome and interplay events characterized by our analysis increasing the understanding of consistent immunologic pathways essential for the immune response to Candida infections.

All in all, our work provides a systems immunology view of the interactome of antifungal molecules, revealing a consistent network interplay between TLR and IFN signaling pathways in response to \textit{C. albicans} and \textit{C. auris}. This study also indicates new biomarkers and provides novel insights into the systemic immunological mechanism against fungal infections. Future works dissecting this interplay will pave the way for new immunotherapy approaches to reduce the high mortality rate caused by fungal infections. Also, our study indicates that the exploration of functional genomic approaches by applying systems immunology methods to investigate IEI will provide new opportunities to further understand the immune system \textit{in natura}.

\textbf{ONLINE METHODS}

\textbf{METHODS}

\textbf{Datasets and curation}
We performed an integrative analysis by searching on NCBI GEO database\textsuperscript{76} and ArrayExpress database\textsuperscript{77}, to identify publicly available gene expression data of infection by *C. albicans* and *C. auris* in whole blood, PBMCs, and moDCs. This search comprised studies published between March 2010 and July 2020. Since transcriptome datasets from patients with candidiasis were not publicly available we consider as criteria for inclusion: (1) gene expression data of whole blood, PBMCs, and moDCs of in vitro infection with *C. albicans*; (2) studies composed of at least 2 samples per group; (3) the inclusion of control groups for comparison; (4) all gene expression analysis platforms were considered; and (5) only studies that have provided the transcriptome data were included for the integrative analyses. Our exclusion criteria were (1) non-human samples, (2) treatment before molecular genetic analysis, and (3) review studies. RNAseq and MicroArray studies were included in our integrative analysis, five studies were retrieved from the NCBI GEO database\textsuperscript{76} (GSE65088 and GSE114180, GSE42606, GSE154911, and GSE77969) and two from ArrayExpress database\textsuperscript{77} (E-MTAB-135, E-MTAB-751). Also, a single-cell RNAseq study was included\textsuperscript{78}.

**Single-cell RNASeq analysis**

We obtained the Seurat Object containing the scRNAseq data from De Vries et al. (2020), which was deposited in the single-cell eQTLGen Consortium database (https://eqtlgen.org/candida.html). We followed the default Seurat pipeline (https://satijalab.org/seurat/articles/pbmc3k_tutorial.html) as previously described by Stuart et al.\textsuperscript{78} to perform differential expression analysis and data visualization (UMAP, dotplot, and heatmap).
Differential Expression Analysis of bulk RNAseq and microarray data

To characterize the immunological signature from the global transcriptional profiles in infection by C. albicans, read counts of each RNAseq study were transformed (log2 count per million), and NetworkAnalyst 3.0 webtool (https://www.networkanalyst.ca/) was used to perform differential expression analysis, applying DESeq2 pipeline. The microarray studies were analyzed through GEO2R web application, available at http://www.ncbi.nlm.nih.gov/geo/geo2r/, using limma-voom pipeline. To select the up and downregulated genes between C. albicans infection and the normal group, we used the statistical cutoffs of log2 fold-change > 1 (upregulated) or < -1 (downregulated) and adjusted p-value < 0.05.

Analysis of Gene Co-Expression Modules

We selected the dataset GSE42606 to analyze the gene co-expression modules with the R-package CEMiTool 1.12.2 using default parameters.

Enrichment Analysis and Data Visualization

We used the differentially expressed genes (DEGs) to identify enriched ontology terms. The pathways and the biological processes were identified through an Over-representation Analysis (ORA) and EnrichR, and the significant enriched immunological terms were generated according to adjusted p-value < 0.05. The Upset and Venn Graphs demonstrating the intersections and comparisons between common DEGs among the datasets were generated through the webtool Intervene and Bioinformatics & Evolutionary Genomics.
We plotted the set of genes shared between the dataset in bubble-based heat maps, applying One minus cosine similarity through the webtool Morpheus (https://software.broadinstitute.org/morpheus/). We used ClusterProfiler to obtain dot plots of enriched terms associated with *C. albicans* and *C. auris* infections. ClusterProfiler and ORA, were performed on R software version 4.0.2 (https://www.r-project.org/index.html), through the packages DOSE, enrichplot, reactomePA, and clusterprofiler. The GOplot was plotted using the R packages unikn, circlize, and GOplot. The statistical graphs were constructed using the functionalities of the ggplot2 package. We represented the shared DEGs between different fungal infections (*C. albicans* and *C. auris*) through circular heatmaps, using the R packages circlize and ComplexHeatmap.

**Correlation Analysis**

We used the GSE42606 dataset to perform the correlation analysis between genes associated with TLRs as well as Type I and II IFN signaling cascades. The correlation matrices were generated with the webtool Intervene (https://intervene.readthedocs.io/en/latest/index.html), using Pearson coefficient. The Canonical Correlation Analysis (CCA) was applied to investigate patterns of association between IFN and TLR genes from the same dataset. The CCA was performed on R software version 4.0.2 (https://www.r-project.org/index.html), through the packages CCA, and whitening. Principal Component Analysis (PCA) analysis was built using R functions prcomp and princomp, through factoextra package.
**Molecular Network**

Networks of related pathways to fungal infection immune responses and physical protein-protein interaction (PPI) networks of DEGs found across all datasets were annotated, analyzed and visualized using NAViGaTOR 3.0,14. Node color represents Gene Ontology cellular component as per legend. DEGs were used as input into Integrated Interactions Database (IID version 2020-05; http://ophid.utoronto.ca/iid)\textsuperscript{91,92} to identify direct physical protein interactions. Networks were exported in SVG file format, and finalized in Adobe Illustrator 2021.

**Multi-study Factor Analysis (MSFA)**

MSFA is a generalized version of factor analysis that allows for the joint analysis of multiple studies. MSFA estimates shared factors common to all studies, as well as factors specific to individual studies. Estimation of parameters for the MSFA model can be computed using either a frequentist or a Bayesian approach. Compared with the frequentist analysis, the Bayesian offers two major advantages: 1- it provides a better defined factors, and 2- it chooses the dimension of the common and study-specific factors through a practical and useful approach. We adopt the Bayesian multi-study, for the inferential analysis to identify common and study-specific factors\textsuperscript{25,33,93} shared by GSE65088 and GSE42606. The Bayesian MSFA considers all data at once in an integrated approach, estimating parameters by maximum-likelihood analysis\textsuperscript{94}.

**Interferome Analysis**
The identification of interferome genes was performed with Interferome V2.01 (http://www.interferome.org/interferome/home.jspx).

**Single-cell RNA-seq differential expression analysis**

Seurat package was used to obtain the DEGs between the different cell types under the conditions of infection by *C. albicans* and resting cells. Enrichment of DEGs by cell group and by total DEGs was done according to the described for ClusterProfiler package.

**Code Availability**

R codes used in this work are available at https://github.com/ranieri131/SalgadoRC_CANDIDA_IMMUNE_RESPONSE_2021

**Acknowledgments**

We acknowledge the São Paulo Research Foundation (FAPESP grants 2018/18886-9, 2020/01688-0, and 2020/07069-0 to OCM) for financial support. RCS is financed by the national council for scientific and technological development (CNPQ). Computational analysis was supported by FAPESP and partially by the grants from Ontario Research Fund (#34876), Natural Sciences Research Council (NSERC #203475), Canada Foundation for Innovation (CFI #29272, #225404, #33536), and IBM. This study was financed in part by the coordination for the improvement of higher education personnel – Brazil (CAPES) – finance code 001.

**Author Contributions**
RCS, PPF, OCM co-wrote the manuscript; RCS, DLMF, TTF, PPF, NOSC, VLGC, LFS, OCM provided scientific insights; RCS, DLMF, AHCM, SMSN, KTA, CASP, GCB, DRP, ISF, RV, IJ, and OCM performed bioinformatics analyses; RCS, DLMF, and OCM conceived and designed the study; RCS, TTF, LFS, PPF, NOSC, VLGC, HDO, LFS, IJ, ACN, and OCM revised and edited the final manuscript; ACN and OCM supervised the project.

Competing interest statement

The authors declare no competing financial and/or non-financial interests concerning the work described.

Data availability

The published transcriptome datasets can be found in the GEO and Array Express databases (IDs. GSE65088, GSE114180, GSE42606, GSE154911, GSE77969, E-MTAB-135, E-MTAB-751). Single-cell data is available as Seurat Object on doi.org/10.1371/journal.ppat.1008408.
REFERENCES


36. Fallarino, F., Puccetti, P., Romani, L., Zelante, T. & De Luca, A. Pathogenic


56. PRIDE - Proteomics Identification Database. https://www.ebi.ac.uk/pride/.

57. ArrayExpress < EMBL-EBI. https://www.ebi.ac.uk/arrayexpress/.


Graphic abstract

TLR 1/2 or TLR2/6

C. albicans

Plasma membrane

MyD88
IRAK1/4
TRAF6

Endosome
TLRs (ex. TLR7/9)

MyD88

IRF9
STAT1
STAT2

Type I IFN
IFN-α
IFN-β

Type II IFN
IFN-γ

JAK1
STAT1

TK1
STAT1

Dendritic cells
E-MTAB-135 (Microarray)
E-MTAB-751 (Microarray)
GSE77969 (Microarray)

scRNAseq

PEMCs
GSE42606 (Microarray)
GSE154911 (RNAseq)

Whole Blood
GSE65088 (Microarray)
GSE114180 (RNAseq)

IKK complex

MAPKs

TRIF

NF-κB

AP-1

IRF7

Pro-inflammatory cytokines
Co-stimulatory molecules
Increased IRFs’ production

Type I IFNs
ISGs
**Graphic abstract.** Schematic view summarizing datasets and the interplay between TLR and IFN signaling pathways (based on references 48,95–101) in the immune response to *C. albicans* (created using BioRender.com). IFN, Interferon; TLR, Toll-like Receptor.
Figure 1

a) Cell clusters

b) Resting

c) C. albicans

d) dotplot for ORA

- Neutrophil degranulation
- Signaling by Interleukins
- Infectious disease
- MAPK family signaling cascades
- Influenza signaling
- Toll-like Receptor Cascade
- TCR signaling
- Antigen processing-Cross presentation
- C-type lectin receptors (CLRs)
- Downstream TCR signaling
- Interleukin-1 signaling
- CLEC7A (Dec205) signaling
- Interleukin-4 and Interleukin-13 signaling
- Interferon-gamma signaling
- Interferon-alpha/beta signaling
- Costimulation by the CD28 family
- Interleukin-17 signaling
- ISG15 antiviral mechanism
- Interleukin-10 signaling

Gene and protein expression by JAK-STAT signaling after IL-12 stimulation
Interleukin-2 family signaling
ROS and RNS production in phagocytosis
TNF signaling
DAP12 signaling
NOD1/2 Signaling Pathway
Phosphorylation of CD3 and TCR beta chains
Translocation of ZAP-70 to immunological synapse
Diseases associated with the TLR signaling cascade
Inflammation
CD209 (DC-SIGN) signaling
Interleukin-5 signaling

e) Box plots for gene expression

f) UMAP scatter plots for C. albicans and Resting

C. albicans
Resting

Min Max Min Max
Fig. 1. scRNAseq revealed common activation of TLR and IFN signaling pathways. a, UMAP visualization of scRNAseq profiles colored according to cell clusters. b, c, UMAP of resting and *C. albicans*-activated cells groups. d, Dot plot showing pathways associated with immune response to *C. albicans*, obtained by ORA of DEGs. e, Boxplot and f, UMAP of DEGs associated with both TLR and IFN signaling pathways, additional DEGs associated with these two pathways are described in Suppl. Table 3. DEGs, Differentially Expressed Genes; IFN, Interferon; ORA, Over representation analysis; scRNAseq, single-cell RNA sequencing; TLR, Toll-like Receptor; UMAP, Uniform Manifold Approximation and Projection.
Fig. 2. Modular gene co-expression analysis and the association of TLR and IFN signaling pathways. a, Co-expression modules significantly enriched (M1-M11, and M13) in PBMCs (resting n= 30; *C. albicans* infected n= 24; dataset GSE42606). b and d Network representation of M1 and M2 with hubs (most connected genes) colored based on co-expression (blue color), co-expressed and interactions (green color), or only interactions (dark-red color). c and e, Enrichment representation obtained by modular genes co-expression in M1 and M2 showing significantly (-Log10 transformed adjusted p-value) enriched signaling pathways. IFN, Interferon; TLR, Toll-like Receptor.
Fig. 3. *C. albicans* activates common TLR- and IFN-associated genes in peripheral blood leukocytes. **a,** The upset plot displays the number (set size) of DEGs present in each dataset (y-axis: WBCs, GSE65088, and GSE114180; PBMCs: GSE42606 and GSE154911) and their intersections. Black bubbles, present in the rows, mark the dataset which refers to the amount present in the blue columns, with intersections between two or more groups being shown. **b,** Hierarchical clustering of the 44 common DEGs demonstrating gene expression patterns across the different studies. The size and color of circles correspond to -Log10 transformed adjusted p-value and Log2 fold change (Log2FC), respectively. Blue represents downregulated and red indicates up-regulated DEGs. The cut-off applied to identify the down-/upregulated genes was Log2FC < -1/ > 1 and adjusted p-value < 0.05. Rows and columns were clustered based on cosine similarity between Log2FC values. **c,** GOplot of selected immunological pathways and associated gene. **d,** Heatmap of common and specific latent factors across the studies. Heatmaps contain genes presenting positive and negative loadings ranging from -1 to 1. DEGs, Differentially Expressed Genes; PBMCs, Peripheral Blood Mononuclear Cells; WBCs, White Blood Cells.
**Fig. 4.** *C. albicans* activate common TLR and IFN signaling pathways across different layers of immunity. 
a-c, Venn diagrams displaying the number of DEGs present in each dataset grouped by cell type and their intersections: datasets of WBCs (a), PBMCs (b), and moDCs (c). 
d, The intersection plot highlights the number of common DEGs across different cell groups. 
e, Hierarchical clustering exhibiting the pathways enriching common biological processes across the studies ([Suppl. Table 9]). 
f, Further analysis of TLR- and IFN-associated pathways. In both heatmaps the size of circles corresponds to adjusted p-value transformed into -Log10 and color intensity indicates the number of genes in each biological process and pathway across the studies, respectively. 
g, Network demonstrating the interactions between TLR- and IFN-associated DEGs/signaling pathways with other molecules and signaling cascades classically associated with the antifungal immune responses. Enrichment analysis was performed using Reactome. Circular nodes represent pathways and their size denote the number of genes enriching the pathways. Colored squares represent the cellular location of genes. The interaction network was built using the NAViGaTOR software. 
DEGs, Differentially Expressed Genes; moDCs, Monocyte-Derived Dendritic Cells; IFN, Interferon; PBMCs, Peripheral Blood Mononuclear Cells; TLR, Toll-like Receptor; WBCs, White Blood Cells.
Figure 5

a) Resting

b) C. albicans

c) TLR-associated genes

Correlation with canonical variate 1

Correlation with canonical variate 2

C. albicans

TLR genes

IFN genes

d) Variables - PCA

Dim1 (73.7%)

Dim2 (13.9%)

Groups: C. albicans, Resting

Dim3 (6.1%)

CONTRIBUTION

1.0

0.5

0.0

-0.5

-1.0
Fig. 5. Relationship between molecules associated with TLR and IFN signaling cascades. a, b, Correloplot of DEGs associated with TLR and IFN signaling cascades in PBMCs (GSE42606) in the absence or b, presence of C. albicans. Histograms of Pearson's correlation coefficient, containing negative and positive correlation from 1 to -1, respectively. c, Estimated correlations of TLR - and IFN-associated DEGs versus their corresponding first 2 canonical variates (x-CV1 and x-CV2, for IFN-associated genes; y-CV1 and y-CV2 for TLR-associated genes). Grey colored variables (with names omitted) are those with correlation coefficient ≤0.7 in its two corresponding canonical variates. Inner dotted lines limit the canonical correlation coefficient between -0.7 and 0.7, while outer dotted lines between -1 and 1. d, e, PCA was used for stratification analysis of resting and C. albicans infected PBMCs based on TLR- and IFN-associated DEGs. d, Of note, individuals with similar expression values for these DEGs are grouped together; e, Variables with positive correlation are pointing to same side of the plot, contrasting with negative correlated variables, which point to opposite sides. DEGs, Differentially Expressed Genes; IFN, Interferon; PBMCs, Peripheral Blood Mononuclear Cells; TLR, Toll-like Receptor.
Common genes modulated by *C. albicans* and *C. auris*

TLR-associated genes

IFN-associated genes

**Expression Z-scores**

**Figure 6**

**Image**

- **Panel a**: Dotplot for ORA
- **Panel b**: Gene expression profile
- **Panel c**: TLR-associated genes, *C. albicans* and *C. auris*
- **Panel d**: *C. albicans* (all datasets)
- **Panel e**: *C. auris* (GSE154911)
Fig. 6. Induction of the interplay between TLR and IFN signaling pathways by the multidrug-resistant *C. auris*. 

**a.** Dot plot showing the 30 most enriched signaling pathways obtained by ORA of DEGs using ClusterProfiler. Y-axis contains enriched pathways while the size of circles represents the number of genes (count) enriching each category and color (from blue to black) indicating how significantly enriched (when p-value <0.05) is the pathway. 

**b.** Network demonstrating interactions between pathways and their associated genes revealed by ORA. Circular nodes represent pathways, circle size is associated with the number of genes enriching each pathway, while colored squares represent the cell location of genes. The interaction network was build using the NAViGaTOR software. 

**c.** Circular heatmaps of RNAseq expression z-scores computed for log2 transformed DEGs (p-value adj < 0.05, fold change > 1 and < -1) compares the expression of TLR- (left panels) and IFN- (right panels) signaling pathways induced by *C. albicans* (green/grey heatmaps) or *C. auris* (yellow/grey heatmaps) all from GSE154911. Small circular heatmaps (blue/grey) demonstrate common DEGs modulated by *C. albicans* and *C. auris*. 

**d.** Venn diagram showing the transcriptional overlap (an intersection containing 237 shared DEGs) induced by *C. auris* and *C. albicans* (those 1096 genes found across all studies: Suppl. Table 10). 

**e.** Dotplot of enriched pathways by the 237 shared DEGs. DEGs, Differentially Expressed Genes; IFN, Interferon; ORA, Over representation analysis; TLR, Toll-like Receptor.
Fig. 7. The interactome of DEGs enriching signaling pathways involved in the anti-candida immune response and its association with inborn errors of immunity. a, Relationships (edges) among the 1096 DEGs (nodes) found across all studies (Suppl. Table 10). Subnetworks (semicircles) represent genes associated with IEI causing increased susceptibility to candidiasis, being 34 purple nodes genes shared with the group of 1096 DEGs, while 66 green nodes represent those not found in the Candida datasets. Colored squares and circles represent the cell location of genes. The interaction network was build using the NAViGaTOR software. b, Network of hubs present in a. c, Venn diagram of interferon types associated with the group of 1096 DEGs. Interferome analysis revealed 868 IFN-regulated genes modulated either by IFN type I, II, and III, as in the Venn Diagram. DEGs, Differentially Expressed Genes; IFN, Interferon; IEI, Inborn Errors of Immunity; TLR, Toll-like Receptor.
Figure 8
Fig. 8. Common TLR- and IFN-associated DEGs and signaling pathways across microarray, bulk, and single-cell RNA-seq datasets. **a**, Heatmap using expression value from scRNAseq of DEGs also present in microarray and bulk studies; cell condition and group are indicated by different colors. **b**, Hierarchical clustering of average expression comparing resting and *C. albicans*-activated cells. **c**, Hierarchical clustering showing common pathways selected from Fig. 1d, across the cell groups; the size of circles corresponds to adjusted p-value transformed into -Log10 and color intensity indicates the number of genes in each pathway across the cell groups, respectively. **DEGs**, Differentially Expressed Genes; **IFN**, Interferon; **TLR**, Toll-like Receptor; **scRNAseq**, single-cell RNA sequencing.
Suppl. Figure 1

(a) Dotplot for ORA

Interferon Signaling:
- Signaling by Interleukins
- Interferon alf/beta signaling
- Interferon gamma signaling
- Class A1 (Rhodopsin-like receptors)
- GPCR ligand binding
- Interleukin-10 signaling
- Peptide ligand-binding receptors
- Chemokine receptors binding chemokines
- Interleukin-4 and Interleukin-3 signaling
- Antiviral mechanism by IFN-stimulated genes:
  - Toll-like Receptor Cascades
  - Toll Like Receptor 4 (TLR4) Cascade
  - ISG15 antiviral mechanism
- DDX58/IFI19-mediated induction of interferon-alpha/beta
- Toll Like Receptor 7/8 (TLR7/8) Cascade
- MyD88 dependent cascade initiated on endosome
- Toll Like Receptor 3 (TLR3) Cascade
- Toll Like Receptor 9 (TLR9) Cascade
- MyD88-independent TLR4 cascade
- TRIF (TICAM1)-mediated TLR4 signaling
- TRAF6 mediated induction of NF-kB and MAPK upon TLR7/8/9 activation
- MyD88:MAL (TIRAP) cascade initiated on plasma membrane
- Regulation of IFNa signaling
- Negative regulators of DDX58/IFI19 signaling
- TRAF3-dependent IRF activation pathway
- RIPK1-mediated regulated necrosis
- Regulated Necrosis
- TRAF6 mediated NF-kB activation
- Regulation of IFNG signaling

(b) Toll-like Receptor Cascades

(c) Antiviral mechanism by IFN-activated genes

Node color: Go biological process
Suppl. Fig. 1. Functional clustering of DEGs associates TLR and IFN signaling pathways. 

**a,** Dot plot showing the 30 most enriched signaling pathways obtained by ORA of DEGs (dataset GSE42606) using ClusterProfiler. Y-axis contains enriched pathways while the size of circles represents the number of genes (count) enriching each category and color (from blue to black) indicates how significant (when p-value < 0.05) enriched is the pathway. **b,** Network of TLR- and IFN-associated genes enriching signaling pathways showed in the dot plot (a). The network includes genes upregulated and downregulated (Suppl. Table 4) when comparing resting with *C. albicans*-infected PBMCs. The size of circles represents the number (counts) of genes that enrich the pathway, and colored squares represent the cell location of genes. **c,** Network demonstrating interactions between pathways and their associated genes revealed by ORA. Circular nodes represent pathways and the circle size is associated with the number of genes that enrich each pathway, while colored squares represent the cell location of genes. The interaction network was build using the NAViGaTOR software. DEGs, Differentially Expressed Genes; IFN, Interferon; ORA, Over-representation Analysis; TLR, Toll-like Receptor.
Suppl. Figure 2

(a) Wnt signaling pathway

(b) Dotplot for ORA

(c) Regulatory networks of immune system signaling pathways
Suppl. Fig. 2. Inborn errors of immunity confirm the interplay of TLR- and IFN-associated genes. **a**, Overview (created using BioRender.com) of genes associated with inborn errors of immunity that are also found among the DEGs across the 7 datasets of WBCs, PBMCs, and moDCs (**Suppl. Table 14**). **b**, Dot plot showing the 30 most enriched signaling pathways obtained by ORA of DEGs seen in (a). Y-axis contains enriched pathways while the size of circles represents the number of genes (count) enriching each category and color (from blue to black) indicates how significantly enriched (when p-value <0.05) is each pathway. **c**, GOPlot displaying genes causing inborn errors of immunity and enriched pathways. DEGs, Differentially Expressed Genes; moDCs, Monocyte-Derived Dendritic Cells; IFN, Interferon; ORA, Over-representation Analysis; PBMCs, Peripheral Blood Mononuclear Cells; TLR, Toll-like Receptor; WBCs, White Blood Cells.