ABSTRACT

Title of Dissertation: IDENTIFICATION AND ANALYSIS OF GENETIC VARIANTS PERMITTING DISSEMINATED COCCIDIOIDOMYCOSIS

Amy Pepper Hsu, Doctor of Philosophy, 2022

Dissertation directed by: Dr. David M. Mosser, Professor
Department of Cell Biology & Molecular Genetics

Disseminated coccidioidomycosis (DCM) is caused by *Coccidioides*, pathogenic fungi endemic to the Southwestern United States and Mexico. While the majority of those infected have minor symptoms or remain asymptomatic, illness requiring medical attention occurs in approximately 30%, with <1% developing extrapulmonary dissemination. To address why some individuals allow dissemination, we performed whole-exome sequencing on an exploratory cohort of 67 DCM patients. Using standard genetic analysis for identification of novel or rare Mendelian mutations only two patients were identified, both with *STAT3* premature termination codons causing haploinsufficiency. Since *Coccidioides* are geographically isolated, I explored the possibility that dissemination could be a combination of more common genetic variants plus exposure. Defects in sensing and response to β-glucan, the major component of Coccidioides cell wall, were seen in 34/67 (50.7%) cases. Damaging variants in *CLEC7A*, encoding DECTIN-1,

(n=14) and *PLCG2* (n=11) were associated with impaired production of β-glucan-stimulated TNF from peripheral blood mononuclear cells compared to healthy controls (P<0.005). Using ancestry-matched controls, damaging *CLEC7A* and *PLCG2* variants were over-represented in DCM (P=0.0206, P=0.015, respectively) including *CLEC7A* Y238* (P=0.0105) and *PLCG2* R268W (P=0.0025). A validation cohort of 111 DCM patients confirmed over-representation of the specific variants, *PLCG2* R268W (P=0.0276), *CLEC7A* I223S (P=0.044), and *CLEC7A* Y238* (P=0.0656). Lastly, I identified a novel pathway of pulmonary-epithelial fungal recognition by DECTIN-1 leading to activation of the NADPH oxidase complex, DUOX1/DUOXA1. Stimulation with a DECTIN-1 agonist induced DUOX1/DUOXA1-derived H_2O_2 in transfected cells. Heterozygous *DUOX1* or *DUOXA1* variants which impaired H_2O_2 production were overrepresented in discovery and validation cohorts. Together these studies highlight the importance of fungal recognition and response for control of infections. Patients with DCM have impaired β-glucan sensing or response affecting TNF and H_2O_2 production. Impaired *Coccidioides* recognition and decreased cellular response are associated with disseminated coccidioidomycosis.

IDENTIFICATION AND ANALYSIS OF GENETIC VARIANTS PERMITTING DISSEMINATED COCCIDIOIDOMYCOSIS

by

Amy Pepper Hsu

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2022

Advisory Committee:

Professor David M. Mosser, Chair
Associate Professor Jason D. Kahn, Dean's Representative
Dr. Steven M. Holland
Professor Volker Briken
Dr. Michail S. Lionakis

Dedication

To my family – Elvin who always believed in me, Elizabeth who taught me to never stop moving forward, Joseph who encouraged me to stretch outside of my comfort zone, and Vasso who reminded me that family is what you make it. This wouldn't have happened without you!

Acknowledgements

It takes a village. In my case, returning to graduate school 30 years after college and raising a family, it took many villages.

To Dr. David Mosser who accepted me and mentored me with only a phone call – this would never have happened without you.

To the amazing physicians I have been honored to work for and with at NIH, Jennifer Puck, Steve Holland, Mihalis Lionakis, Alexandra Freeman, Jenna Bergerson, Gulbu Uzel and Andrea Lisco – your commitment to your patients and to putting their well-being first and good science second is unparalleled.

To my mentors, Jennifer, Steve, Mihalis, and Tom - for believing in me, letting me explore and discover, and ultimately learn how to think critically and openly about science and strive to do the next, right experiment. We have had some amazing adventures together, here's to many more.

To Tom Leto – thank you for welcoming me into your world of biochemistry and reactive oxygen species, for generously sharing your knowledge and being patient as I learned.

To the women in my science circles – Jennifer Puck, Polly Matzinger, Alexandra Freeman and Jenna Bergerson – thank you for reminding me that there is a place for us at this table, even if some try to push us out.

To Volker Briken and Jason Kahn who agreed to serve on my committee – you challenged me and pushed me outside of my comfort zones with some of your questions. You opened my eyes to different aspects of my project I had never considered.

To my collaborators at the Valley Fever Center for Excellence in Tucson, AZ for teaching me and challenging me to think about a unique organism.

To my cohort of graduate students – thank you for accepting me and being my friends even though I was old enough to be your mother. You have the future in front of you – I can't wait to see what you all do.

To Glennys, Vasilis, and Andrea – for the laughter, tears, long discussions, and coffee. Lots of coffee.

And last, but not least, to Pat. You were there from the start after a chance meeting on the trail. We laughed and cried together through the ups and downs, grief, surgeries, rejections and acceptances, and you always told me I'd get here. Thank you for helping me look forward to the goal.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

Ψ	pseudogene
1000G	1000 Genomes
AA	Amino acids
AECII	Type II Alveolar epithelial cells
AFR	African / African American ancestry
AMR	Latino/Admixed American ancestry
AQP3	Aquaporin-3
BAL	Bronchoalveolar lavage
CADD	Combined Annotation-Dependent Depletion
CDD	Conserved domain database
CF	Cystic fibrosis
CFU	Colony forming units
CMC	Chronic mucocutaneous candidiasis
CNS	Central nervous system
Cp1038	*Coccidioides posadasii* strain 1038
CTLD	C-type lectin domain
DCM	Disseminated coccidioidomycosis

DHR	dihydrorhodamine oxidation assay
EAS	East Asian ancestry
EBV	Epstein Barr virus
EMT	epithelial-mesenchymal transition
ENCODE	Encyclopedia of DNA Elements
EUR	European ancestry
FPKM	fragments per kilobase per million mapped reads
GATK	Genome analysis toolkit
gnomAD	Genome Aggregation Database
gnomAD freq	allele frequency across all gnomAD samples
H_2O_2	Hydrogen peroxide
HC	Healthy control
HCV	Hepatitis C virus
het	heterozygous
hg19	Human genome build 37
HGC	Human Gene Connectome
HIES	Hyper IgE syndrome
hom WT	homozygous wild-type
HPV	Human papilloma virus
HSCT	Hematopoietic stem cell transplantation

HSV	Herpes simplex virus
IBD	Inflammatory bowel disease
IFNγ	Interferon gamma
JAK	Janus kinase
LPO	Lactoperoxidase
LPS	lipopolysaccharide
M. tb	*Mycobacterium tuberculosis*
NCBI	National Center for Biotechnology Information
NMD	Nonsense-mediated decay
o/e	Observed / expected
OR	Odds ratio
$OSCN^-$	hypothiocyanite
pLoF	putative Loss of function
PMA	phorbol 12-myristate 13-acetate
Pop Freq	allele frequency across specific population within gnomAD
PRR	Pattern recognition receptor
ROS	Reactive oxygen species
RR	Relative risk
SAS	South Asian ancestry
SCN^-	thiocyanate

snRNP	small nuclear ribonucleoprotein
STAT	Signal transducer and activator of transcription
STAT3-HIES	STAT3 mutated Hyper IgE syndrome
TMAP	Torrent Mapping Alignment Program
TSS	Transcription start site
UCSC	University of California Santa Cruz
VCF	Variant callset file
WES	Whole exome sequencing
WGS	Whole genome sequencing
WT	Wild-type
Y_n	polypyrimidine tract
yr	year

Chapter 1: *Coccidioides*: Natural history and human infections

1.1 Introduction

Coccidioides spp. are dimorphic fungi endemic to the desert Southwestern United States and Mexico. In the soil the fungus exists as arthroconidia, viable spores resulting from hyphal segmentation. When the soil is disturbed by construction, farming, winds, etc. the arthroconidia may become airborne (Figure 1.1A) allowing inhalation by animals and humans. Due to the small size, the 2-5 μM arthroconidium can enter deep into the lower airway, settling in terminal bronchioles (Figure 1.1B). Once inhaled, the arthroconidium swells becoming a developing spherule, undergoing free nuclear division and developing endospores (Figure 1.1C).

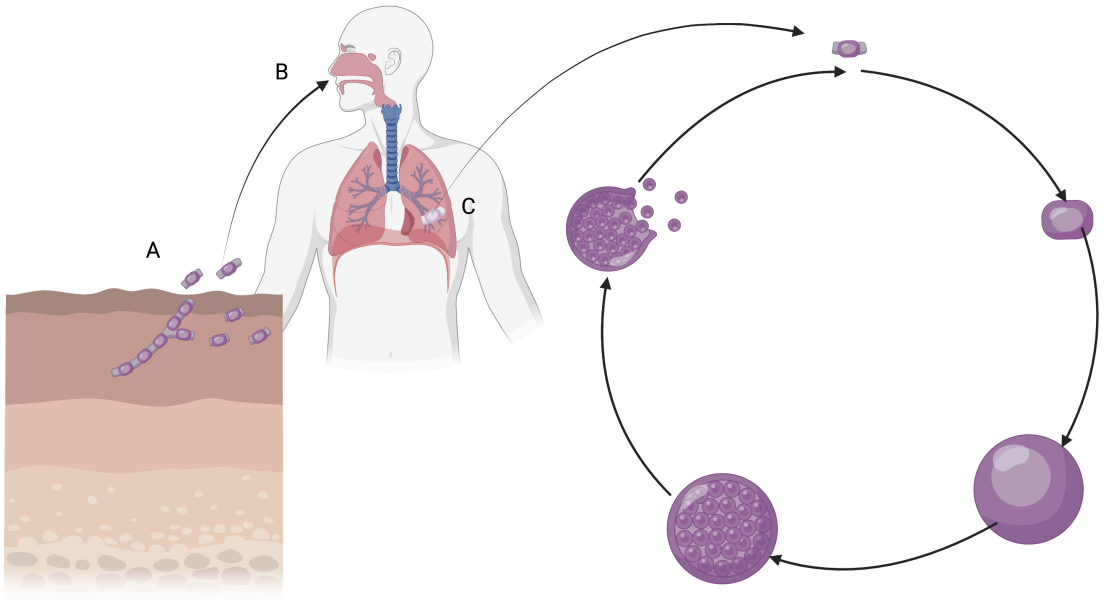


Figure 1.1. *Coccidioides* life cycle in human infection. A. Hyphal forms of the fungus are present in soil. The hyphae septate allowing separation of individual arthroconidia which may become airborne. B. Mammalian host inhales arthroconidia. C. A single arthroconidia swells to become an immature spherule then undergoes internal cell division to produces hundreds of endospores which are released upon spherule rupture. Figure made in BioRender.

Critically, at this point, the developing spherule is now too large for neutrophil phagocytosis and is approaching the limits of macrophage phagocytosis. Over the course of 3-5 days, the endospores mature (Figure 1.2A,B), causing spherule rupture (Figure 1.2C) and release of hundreds of endospores (Figure 1.2D,E), each of which can develop into a spherule establishing a logarithmic growth of the fungus. Spherule rupture leads to rapid influx of immune cells.

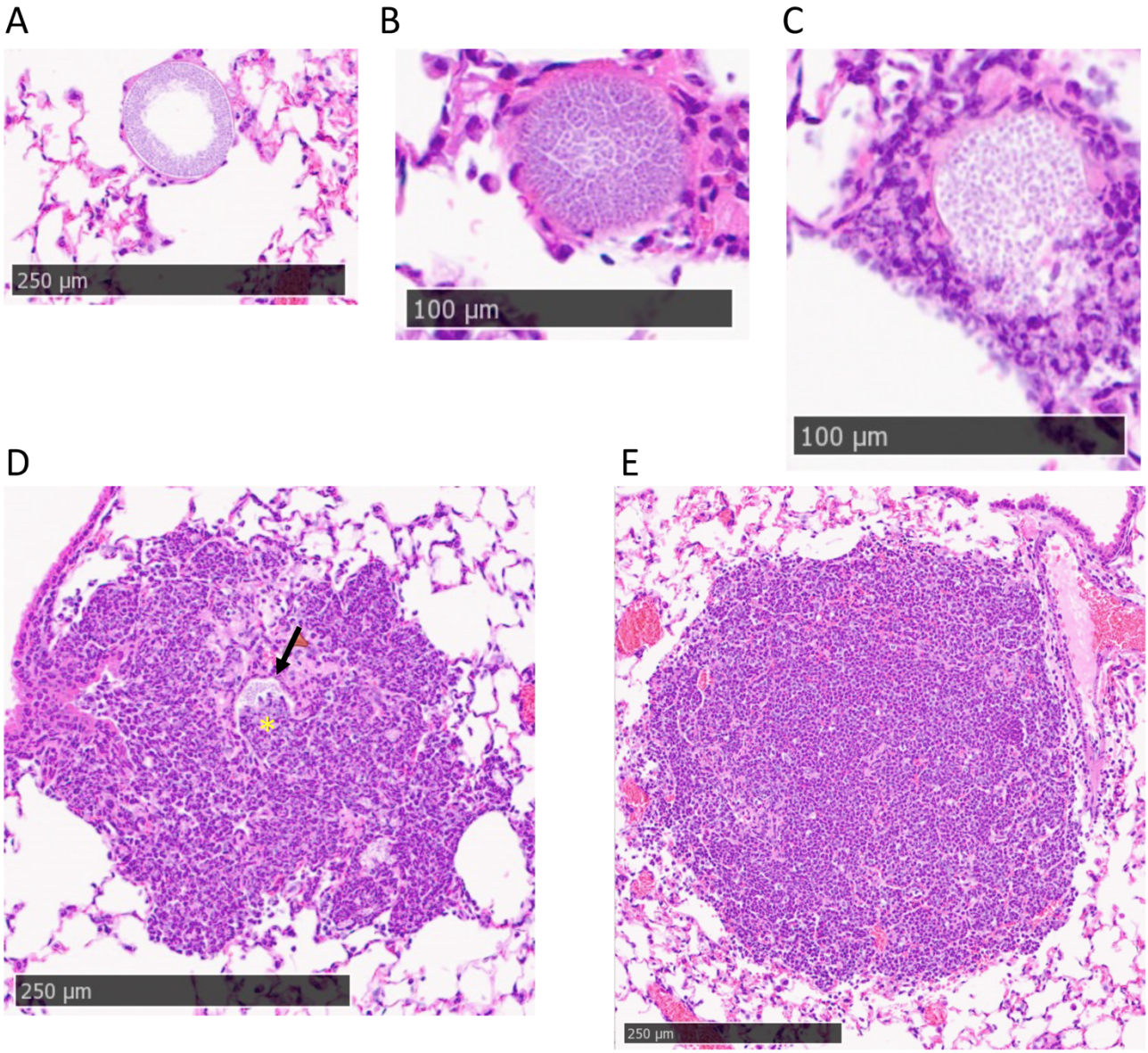


Figure 1.2. Spherule development. A. Immature spherule showing endospores developing internally and lack of immune recognition. B. Mature spherule with immune cells beginning to surround it. C. Rupturing spherule showing influx of neutrophils concurrent with expulsion of endospores. D. Ruptured spherule with immune cell recruitment; both neutrophils (asterisk) and endospores (arrow) seen within the spherule. E. Site of spherule post-rupture, inflammatory region is ~500 µm in diameter. H&E staining of paraffin embedded lungs. Images scanned using a Hamamatsu NanoZoomer S60 slide scanner and NDP.view 2 software.

1.2 Endozoan theory

Early studies from 1950-1970 attempted to identify the ecological niche for *Coccidioides.* Low annual rainfall, sandy soils and periodic dust storms characterize the lower Sonoran Desert life zone, the originally identified Coccidioides endemic area[94]. Soil conditions reported included salinity[34,36], temperature[34], alkalinity[83] and texture – fine sand and silt but not clay[40]. A more recent study attempted to merge these factors and found that only temperature and soil texture were associated with the presence of *Coccidioides* while pH, vegetation types, and electrical conductivity were not associated[40]. Most recently, a soil modeling study[31] predicted suitable habitats with primary features of temperature, soil salinity and moisture. This model identified not only endemic areas but also two locations, south-central Washington State and Dinosaur National Monument in northeastern Utah where individuals have developed coccidioidomycosis from infected soil.

The difficulty establishing suitable habitat for Coccidioides may be indicative that soil is not the primary criteria for survival. In 1965, Maddy and Crecelius[95] demonstrated burial of infected mammals led to infected soil, offering the first hint that Coccidioides are able to grow on animal matter. Infected small mammals will often retreat to burrows to die at which point the fungus switches to its hyphal phase, utilizing nutrients from the dead host and infecting the soil. This retreat to burrows or burial of infected animals may account for the patchy distribution of *Coccidioides* in the soil[155] with *Coccidioides*-positive and negative sites interspersed with each other and remaining positive or negative for extended periods of

time[155]. Correlations between the presence of *Coccidioides immitis* and rodent burrows were first demonstrated in 1956[35]. The importance of *Coccidioides*/animal connection was highlighted when genome sequencing of members of the Onygenales family – including the human pathogens, *Coccidioides spp*, *Histoplasma capsulatum* and *Blastomyces dermatiditis,* and related, non-pathogenic fungi - identified evolutionary changes in gene families[145]. Strikingly, there was a contraction of gene families involved in metabolism of plant material and expansion of proteases and keratinases associated with nutritional utilization of animal products[145]. Molecular analysis of soil samples from animal burrows in the endemic region found 95/105 Coccidioides-positive soils (90.5%) were taken from animal burrows while only 10/105 positive samples were from top soil[80]. Together these support the small-mammal reservoir hypothesis in which the fungus exists within granulomas of living mammals. Upon death of the animal, the endospore switches to hyphal growth and converts to septated mycelium with alternating segments becoming arthroconidia available to infect a new host[159].

1.3 Human infections

In 1892 Alejandro Posadas first described coccidioidomycosis in an Argentinian soldier with recurrent, recalcitrant dermatological problems. What began as a single lesion on his face in 1889, became a series of red, itchy spots which ulcerated and discharged pus. When Posadas examined the patient in 1891, the lesions had spread become a large fungal-like mass on his face, numerous lesions on his nose, growths on his forearm as well as lesions on his trunk and abdomen. In a classic demonstration of Koch's postulates, Posadas used material from the

soldier to infect several small mammals, recapitulating the disease. Notably in 1896, Rixford and Gilchrist reported the causative organism was a protozoan resembling Coccidia and named it *Coccidioides immitis*[131]. A decade later, Ophüls conclusively demonstrated the organism was a fungus and documented pus taken from a lesion on a guinea pig contained spherules on a slide and become mycelia the following day[116].

To date there have been no reported cases of transmission of Coccidioidal infections between humans. Cocci pneumonias are found in single lobes with focal consolidation suggesting single point of infection (i.e. single arthroconidia) [JN Galgiani, personal communication]. As mentioned above, the primary route of infection is inhalation of airborne arthroconidia. With the rise of population in the desert southwest has come extensive construction, road work and off-road recreational activities, each of which can result in soil disruption, allowing arthroconidia to become airborne. Disease presentation can range from asymptomatic in the majority of infections (~2/3[150]), primary pulmonary (self-limiting flu-like symptoms or progressive pneumonia which clears with treatment and does not recur), chronic pulmonary (pneumonia which, despite antifungal therapy remains present after >12 months of treatment), or extra-pulmonary dissemination (presence of fungus outside of the lungs, frequently skin, bone, brain, soft tissue). Dissemination is estimated to occur in 600-1000 individuals of the ~150,000 believed to become infected each year (0.4-0.6%)[46].

1.4 Cost of disease

Although 27 states consider Coccidioidomycosis a reportable disease (CDC https://www.cdc.gov/fungal/fungal-disease-reporting-table.html, accessed 12/2021), 97% of reported cases occur in Arizona or California with roughly 2/3 in Arizona (CDC https://www.cdc.gov/fungal/diseases/coccidioidomycosis/statistics.html, accessed 12/2021) (Figure 1.3A). As a result, the majority of the economic burden exists in those two states. There has been a documented increased incidence of Coccidioides infections in endemic areas both by absolute number of cases (Figure 1.3B, data from https://www.cdc.gov/fungal/diseases/coccidioidomycosis/statistics.html, accessed 12/2021) and incidence normalized to population from 5.3/100,000 in 1998 to 42.6/100,000 in 2011[161]. In Arizona specifically, the rate increased from 84.4/100,000 in 2014 to 144.1/100,000 in 2019[51].

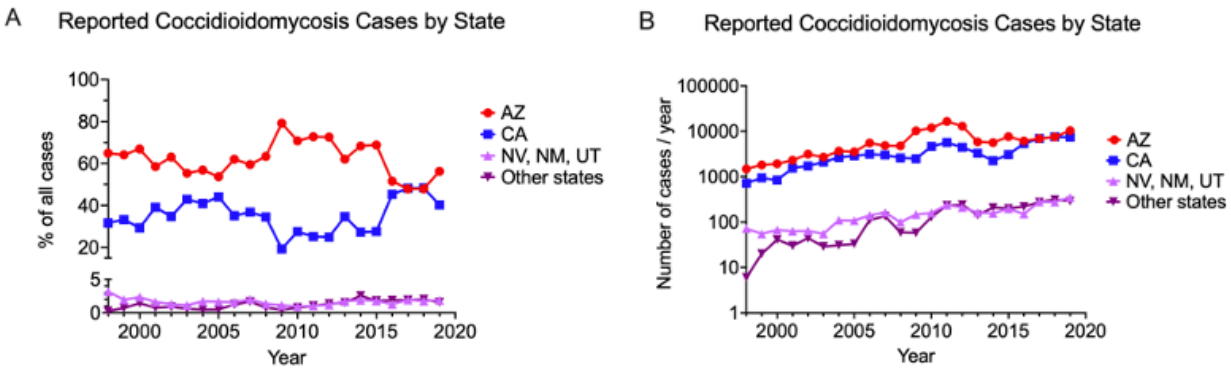


Figure 1.3. Coccidioidomycosis cases by state. A. Cases of coccidioidomycosis reported to CDC 1998 – 2019. B. Number of reported coccidioidomycosis cases per year by state, 1998 - 2019.

The economic burden of disease has also increased. Using lifetime costs for 7466 cases diagnosed in California in 2017, Wilson et al.[178] found total lifetime costs in CA to be $700 million ($429 million direct and $271 million indirect). Highest per person costs were for DCM ($1.025 million) with even uncomplicated Cocci pneumonia costing >$22,000/person. Similar calculations were performed for Arizona[51] in which they calculated the incidence-based lifetime cost for the 10,359 cases in Arizona in 2019 to be $736 million. Not surprisingly, the highest economic burden came from patients with disseminated disease, calculated to be $1.26 million direct (healthcare costs for diagnosis, treatment, procedures and care) and $137,400 indirect (cost of lost work and mortality) per person[51].

1.5 Dissemination risks

External factors such as use of biologics (TNFα or IL-6 inhibitors), immunosuppression for medical reasons (transplant recipients), immunocompromised hosts (HIV-AIDS) or pregnancy [reviewed in Odio[114]] are seen in many individuals with severe or disseminated disease. Apart from those known risk factors, epidemiologic studies have demonstrated occupational exposure, gender and ancestry. In a given setting, there is no reason to suspect differences in exposure to arthroconidia across individuals. Despite that, numerous studies have documented differences in disease severity across people of different ancestries. A prospective study followed military personnel in Southern California for three months using coccidioidin skin-test positivity as a marker for infection. Despite similar rates of conversion to positivity, 4/49 (8%) of African Americans had disseminated disease compared to 0/34

Caucasians[175]. After a 1977 dust storm near Bakersfield, CA, there was a disproportionate amount of dissemination among African Americans (25-50%[41,117,176]) and Asians (38%[176]) compared to 0% dissemination among Caucasian patients[176]. Similar burden of severe disease is seen in epidemiologic studies focusing on hospitalization. In one study, African Americans had a 12-fold higher rate of hospitalization for disseminated coccidioidomycosis in Arizona than did Caucasians despite equivalent all-cause hospitalizations[143]. Those findings were matched in California with 8.8 fold higher hospitalization for DCM for African Americans compared to Caucasians[143]. A later study evaluated data for hospitalizations in Arizona and California from 2005-2011 and compared the annual coccidioidomycosis hospital incidence rate per 100,000 persons[82]. While the rates fluctuated year to year, consistently Asian/Pacific Islander (11.8/100,000 persons, 95% CI 9.1-14.4) and African American (12.8/100,000 persons 95% CI 10.9-14.6) individuals had higher incidence of hospitalization for coccidioidomycosis than individuals of other ancestries (Caucasians 6.2/100,000 persons 95% CI 5.2-7.2; Native American, 5.6/100,000 persons (95% CI 4.2–7.1); Hispanics 5.2/100,000 persons (95% CI 4.3–6.1); and 0.5/100,000 persons (95% CI 4.1–7.1) among other race/ethnicities[82]. While these studies do not adjust for socioeconomic factors, there was no significant difference across quartiles of household incomes in hospitalization for coccidioidomycosis compared to all-reason hospitalization[82] suggesting some inherent basis for susceptibility to dissemination or severe disease.

In multiple epidemiologic studies increased risk of infection is seen in distinct populations including men[150,151], individuals of African descent[133,143,150,151], pregnant women and immune-compromised individuals[114]. In a recent study the relative risk (RR) of coccidioidomycosis in California (2000-2018) was similar for Blacks and Hispanics (1.76, 1.81 respectively)[151]. When that risk was stratified and adjusted for region, Black individuals had a RR of 2.13 while Hispanic individuals dropped to 1.21. The authors suggest this may be due to higher environmental exposure among Hispanics. This hypothesis is supported by a recent case/control CDC report on occupational risk of coccidioidomycosis among Hispanic workers in Kern County, CA. The risk of developing symptomatic coccidioidomycosis was two-fold higher after dust exposure and three times higher for those individuals working with root and bulb vegetable crops. In contrast, individuals working with leaf crops such as grapes had a 60% decrease in risk[99].

1.6 Genetic mutations associated with Coccidioides dissemination

Identification of causative mutations in patients with invasive fungal disease in isolation or as part of a broader primary immune deficiency have identified key genes, proteins, and pathways critical to fungal resistance. Despite the 600-1000 cases of disseminated infection per year[46], there are only 13 patients in the literature with identified monogenic mutations. The first DCM patient with an identified mutation was a young girl with *Coccidioides* meningitis and hyper-IgE syndrome (HIES) caused by a dominant-negative mutation in *STAT3*[124]. Patients with HIES have susceptibility to invasive fungal disease, however in one series of 64 STAT3-mutated

HIES patients, invasive fungal infections only occurred among those with existing lung damage[168]. Since the original patient, two additional HIES patients were reported with *Coccidioides* meningitis[113,153] while we have diagnosed one additional STAT3-mutated patient with chronic, refractory pulmonary coccidioidomycosis [Hsu, unpublished). Additional patients reported include cytokine receptor mutations: one with dominant-negative mutation in *IFNGR1*[167], two siblings with homozygous *IL12RB1* mutations[169], one patient with heterozygous *IL12RB2*[114], as well as transcription factors: two with gain-of-function *STAT1* mutations[136], three members of a single family with dominant-negative *STAT4* mutation[120,122]. One additional patient has been reported with a dominant mutation in *GATA2*[72] which causes the loss of B-cells, NK-cells, monocytes[63] and dendritic cells[29]. It is noteworthy that with the exception of *GATA2*, the mutations occur within components of the IL-12/IFNγ pathway (Figure 1.4) suggesting initial recognition and response to the fungus is critical to the control of the infection.

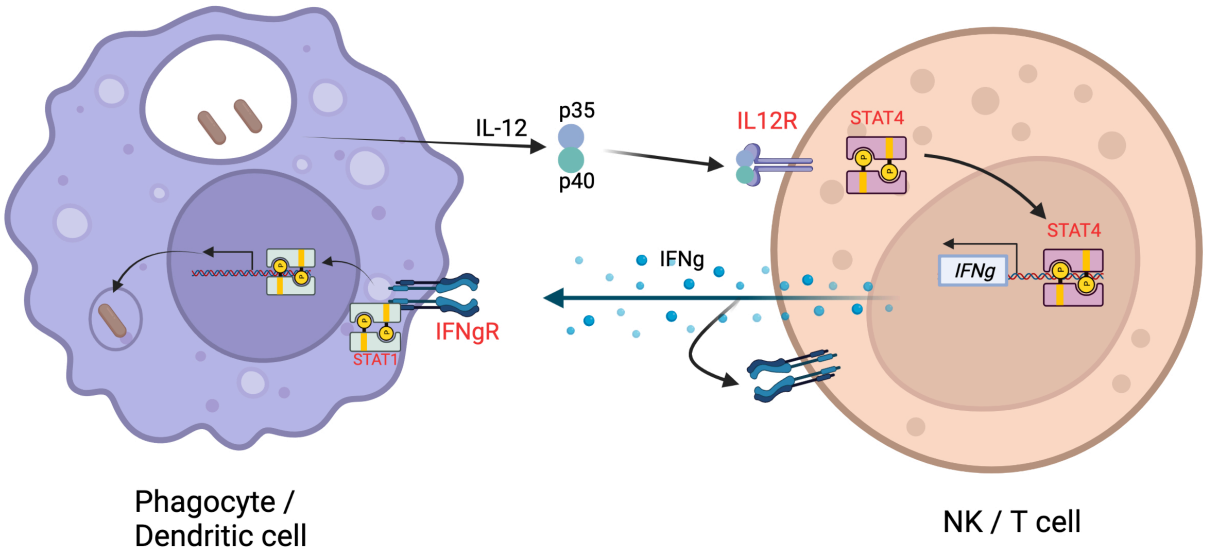


Figure 1.4. IL-12/IFNγ pathway. Upon recognition of certain bacterial or fungal pathogens, antigen presenting cells including macrophages and dendritic cells will release the dimeric cytokine, IL-12. Engagement of the IL-12 receptor, found on T- and natural killer (NK-) cells, results in phosphorylation of STAT4 leading to activation of transcription of interferon gamma (IFNγ). Figure made in BioRender.

The focus of this work is to examine the genetic differences between the fraction of 1% of individuals who disseminate *Coccidioides* and the two-thirds who never come to medical attention. First I will discuss how genetic variants are analyzed to predict mutations, highlighting online tools and resources as well as considerations of transcriptional and translational impact using a patient with disseminated *Mycobacteria tuberculosis* as a case. Next I will discuss the role of the transcription factor, STAT3, in fungal infection and demonstration of STAT3 haploinsufficiency in two patients of our cohort. Since there are few DCM patients in our cohort or the literature with identified mutations I present an alternative genetic analysis suggesting more common variants in the setting of a highly pathogenic

organism with limited geographic distribution and identification of a fungal-recognition/response pathway in the pulmonary epithelia. Finally I summarize the data, suggest additional avenues of study proceeding from this work and place this in context of the challenges facing us today with other infectious diseases such as SARS-COV-2.

Chapter 2: Identifying genetic mutations

2.1 Introduction

The advent of high-throughput, next-generation sequencing has revolutionized the ability to identify genetic causes for patient phenotypes. In a trade-off to maximize the likelihood a damaging variant would be identified, whole exome sequencing (WES) was developed in which the 1-2% of DNA corresponding to coding exons is captured and sequenced. This allowed deeper coverage of a smaller amount of the genome, improving the ability to identify heterozygous variants. Now whole genome sequencing (WGS) which covers the majority of the genome as well as mitochondrial DNA, is becoming the norm in research and even some clinical laboratories. The difficulty facing researchers and clinicians is no longer how to obtain sequence data, the challenge is how to interpret variants identified by these techniques. In this chapter I will summarize how analysis of variants can be performed, some critical elements which should be considered, and additional testing which can be used. Historically, a mutation was defined as a change in DNA that resulted in a change in phenotype in an individual. Changes in DNA may occur through errors during DNA replication or demethylation, or exposure to environmental factors such as chemicals or radiation. Changes which are found in germ cells and can be passed down to offspring are termed germline or inherited while DNA changes occurring in only a subset of cells and not affecting gametes are termed somatic. Mutations may be dominant, in which case a single mutant allele is sufficient

to cause disease, or recessive, in which case both copies of the gene need to be mutated in order to cause phenotype. In either case, determination of pathogenicity is first predicted *in silico* then demonstrated *in vitro* or *ex vivo*.

2.2 In silico annotation and variant nomenclature

After performing WES, the sequencing reads are aligned to the reference genome. Those locations which differ from the reference are collated into a list of variants which can then be annotated to include chromosomal location, associated gene, location within the gene (i.e. intron or exon, coding or untranslated region) and predicted effect - alteration of canonical splice sites or effect on translation (nonsense – substitution of amino acid codon with termination codon, missense – replacement with different amino acid codon, or synonymous – alteration of codon but retaining the original amino acid). The variants are then named according to Human Genome Variation Society standards (https://varnomen.hgvs.org/). First, the specific transcript being used must be noted since several genes have multiple splice forms or alternative 5' untranslated exons. Nomenclature for a cDNA variant is to begin with "c." to denote cDNA, followed by the cDNA base number counting the A of the ATG start codon as 1. This is followed by the reference base and the variant base. For example a C to T transition at base 358 would be c.358C>T. Similarly, protein nomenclature begins with "p." to specify protein, followed by the reference amino acid, the amino acid number - counting from the initial Methionine as 1 - and the variant amino acid. If the base change above replaced an Arginine (CGC) with Cysteine (TGC) the reported protein variant would be p.R120C. Either single

letter or three letter amino acid codes are acceptable for protein variants. Intronic variants after the exon are denoted with the last cDNA base of the exon followed by "+" and how many bases into the intron, for example c.189+2T>A would indicate the second base in the intron was changed from a "T" to an "A". Similarly, variants occurring before the exon are denoted by the number of the first base of the exon followed by "-" and the number of bases preceding the exon, c.190-3C>G. Using standard nomenclature provides a framework for discussion of genetic variation.

2.3 In silico predictions of deleteriousness

Additional annotations are usually included which predict how damaging the variant is. Some commonly used predictors include SIFT which compares the physicochemical properties between the reference and alternate amino acids (AA)[111] suggesting more dramatic AA changes are more damaging, PolyPhen2 which compares the conservation of the AA across species[1] based on the premise that conservation of AA is indicative of importance in the protein so less conserved AA are more amenable to change than highly conserved residues. Similarly, conservation at the nucleotide level across species, predicted by GERP++[25] may suggest constraints affecting transcription or splicing. More recently, ensemble predictors of pathogenicity have been developed which integrate and weight predictions based on multiple factors including nucleotide and AA conservation, location within the protein to identify conserved domains, effect on splicing, transcription factor binding, and variant frequency in population databases. The most commonly referenced pathogenicity predictor is Combined

Annotation-Dependent Depletion (CADD)[79,129] which, after validating the method on known and artificial variants, computed scores for the majority of single nucleotide changes that could exist in the human genome, estimated to be 8.6 million. These scores were then log-scaled such that a CADD score of >10 corresponds to the top 10% of variants predicted damaging in the human genome, CADD > 20 is the top 1% while CADD > 30 is top 0.1%. Using this method, CADD scores have immediate context for how damaging that variant is predicted to be in silico.

2.4 Variant frequency and gene constraints

In order to begin evaluating variants, some common baseline values need to be established, especially which version of the reference genome sequence was used and what natural variation exists across different populations. The first such reference for gene variation from large scale sequencing was the 1000 Genomes project[160], begun in 2008 and designed to provide an overview of human genetic variation. This was followed in 2016 by ExAC[75], a compendium of 60,000 WES assembled by the Broad Institute and analyzed using the same bioinformatic pipeline. ExAC was the first easily accessible, large-scale, web-based browser that provided frequencies of individual variants in the dataset as a whole as well as across multiple populations. This has since been supplanted by the larger, gnomAD, Genome Aggregation Database that includes exomes or genomes on more than 141,000 individuals[76]. The use of variant frequencies has become commonplace in genetic diagnostics since it is unlikely a severe phenotype is caused by a variant widely present in a population. An additional metric established by the large-scale assembly of variants is the constraint any given gene has for

variation. It is possible to predict how many variants a gene would be expected to have given its size and a common rate of change. Genes that have fewer variants than expected (observed / expected or o/e) are considered to be constrained against variation implying a requirement for full function of the gene product. Constraint scores are then calculated for each gene and for synonymous variants (no change in AA, should not be constrained), missense variants (change in AA possibly affecting protein folding or function) or null variants (in this case classified as premature stop codons or splicing mutations that would lead to loss of the protein product). Figure 2.1 shows gnomAD constraint metrics for three genes, *TLR5*, *CFTR* and *STAT3*. The number of "observed" *TLR5* variants equals "expected" for all three categories: synonymous, missense and loss of function/null alleles (collectively "predicted loss of function" abbreviated pLoF) (Figure 2.1A), indicating there is average variation within the gene and suggesting variants identified would not be causative unless bi-allelic mutations were identified. In keeping with this observation, there have been 3 studies[57,58,173] associating a specific SNP with disease state but no mutations have been causative in human disease and mouse studies have relied on homozygous knockout mice to identify phenotypes. Autosomal recessive disease genes, those requiring both alleles to carry damaging mutations, may exhibit more variation than expected as seen for *CFTR* (Figure 2.1B). While synonymous and pLoF variants occur at the expected level, missense variants are more common in the population than would be expected. Finally, genes with autosomal dominant inheritance are constrained against variation, both missense and pLoF as shown for the transcription factor *STAT3* (Figure 2.1C). Constraint metrics for a

gene may therefore support a causal role for damaging variant such as in *STAT3*, or may suggest a search for a second mutation if the gene does not appear constrained. It is worth noting however that constraints are not the sole measure predicting inheritance as phenotypes may be associated with carrying mutations in known autosomal recessive genes such as *CFTR*[100].

A *TLR5* Observed = Expected

Category	Expected SNVs	Observed SNVs	Constraint metrics
Synonymous	172.5	178	Z = -0.33 o/e = 1.03 (0.91 - 1.17)
Missense	429.9	407	Z = 0.39 o/e = 0.95 (0.87 - 1.03)
pLoF	21	20	pLI = 0 o/e = 0.95 (0.67 - 1.38)

B *CFTR* Observed > Expected

Category	Expected SNVs	Observed SNVs	Constraint metrics
Synonymous	272.2	256	Z = 0.77 o/e = 0.94 (0.85 - 1.04)
Missense	753.5	996	Z = -3.14 o/e = 1.32 (1.25 - 1.39)
pLoF	77	84	pLI = 0 o/e = 1.09 (0.91 - 1.31)

Note More missense variants than expected

C *STAT3* Observed <<< Expected

Category	Expected SNVs	Observed SNVs	Constraint metrics
Synonymous	161.9	158	Z = 0.24 o/e = 0.98 (0.86 - 1.11)
Missense	418.5	131	Z = 4.99 o/e = 0.31 (0.27 - 0.36)
pLoF	50.1	1	pLI = 1 o/e = 0.02 (0.01 - 0.1)

Figure 2.1. Comparison of constraint scores from gnomAD. A. *TLR5* demonstrates similar numbers of observed variants as predicted for all three categories, synonymous, missense and null alleles (pLoF). B. *CFTR* demonstrates more missense variants than expected indicating a large amount of variation within the gene. Synonymous and pLoF variants occur at the expected frequency. C. *STAT3* is highly constrained against both missense and pLoF variants characteristic of a gene with dominant inheritance and susceptible to haploinsufficient mutations.

2.5 Filtering variant lists

Using annotated variant files, the hunt for causative variants usually begins with novel or rare variants, with thresholds set at < 1/100 (1%) or lower and those predicted damaging (CADD score > 15-20). These two constraints alone will often reduce the number of variants to be considered by orders of magnitude, from hundreds of thousands to several hundred. Figure 2.2 is an example of WES analysis for a patient with disseminated *Mycobacterium tuberculosis* (*M. tb*). Initial sequencing identified 400,630 variants. Filtering those variants for read-depth (>10, enabling accurate variant calling), genotype quality (>75, a statistical measure of zygosity determination), presumed germline inheritance (variant allele frequency > 0.25, eliminating somatic variants) there were 47,009 variants. Under the assumption that *M. tb* infection is not uncommon but severe disease and dissemination is rare, we would pursue a rare, damaging variant model and filter for population minor allele frequency < 1% and CADD > 20, producing 368 variants (Figure 2.2). This is a bit more manageable but still a large number of variants.

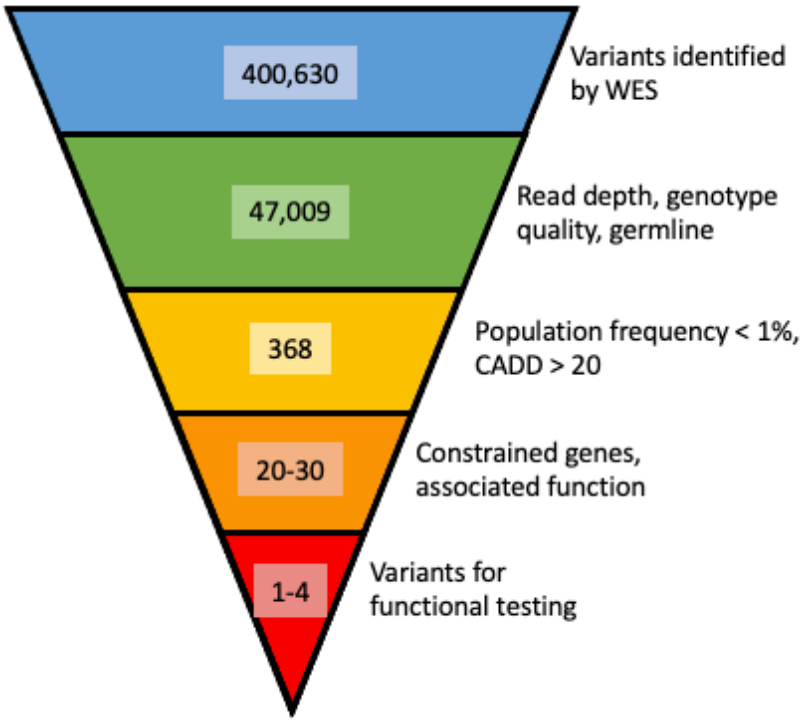


Figure 2.2. Filtering of variants from whole exome. Starting with all variants identified by whole exome sequencing, each successive filter decreases the number of variants considered causative until reaching a small enough number to perform functional studies.

Utilizing the prediction of pathogenicity provided by CADD, the target list of genes can be evaluated. One caveat of using CADD scores is a high weight is given to pLoF variants regardless of how tolerant the gene is to pLoF. In the example provided, 8 of the 20 highest CADD scores are mutations creating a premature stop codon or altering a canonical splice site (Table 2.1).

Table 2.1. 20 Variants with highest CADD scores identified in patient with disseminated *M. tb*.

GeneName	cDNA variant	Protein variant	CADD	pLoF o/e	Missense o/e
ZRANB1	c.1240G>T	p.E414*	41	**0.11**	
FAM8A1	c.1015C>T	p.R339*	41	**0.35**	
URB2	c.4480C>T	p.Q1494*	40	0.64	
GEMIN4	c.2599C>T	p.Q867*	39	0.72	
FAM8A1	c.1183C>T	p.R395*	37	**0.35**	
CNTNAP3B	c.3811C>T	p.Q1271*	36	1.27	
VPS13B	c.9506-1G>T		34	0.55	
ELN	c.1150G>A	p.G384R	33		0.99
ARID4B	c.1912C>T	p.R638W	32		**0.75**
CKMT1B	c.1018C>T	p.R340C	32		0.86
NBR1	c.2230C>T	p.R744C	32		0.83
SPTBN4	c.4093C>T	p.R1365W	32		**0.70**
TRAPPC10	c.2381A>G	p.D794G	32		0.84
ACAA1	c.904G>T	p.G302W	32		0.78
SHPRH	c.2326C>T	p.R776C	32		0.78
PRSS3	c.337C>T	p.Q113*	32	1.41	
FN3KRP	c.446C>T	p.T149M	31		1.07
ARAP2	c.4159C>T	p.R1387C	31		1.08
MINDY4	c.1553C>G	p.S518W	31		1.00
HAX1	c.830G>A	p.R277Q	29.6		0.97

In this case, examination of the pLoF score from gnomAD – those genes with low observed/expected (o/e) ratios are more likely to be damaging than those with high o/e. Of those in Table 2.1, only two genes in the list, *ZRANB1* and *FAM8A1,* have pLoF o/e ratios < 0.35 (bold) indicating very low tolerance for loss of function variants. A similar score exists for missense changes in which o/e < 0.75 indicates higher level of constraint (bold). Again, only two genes from Table 2.1 reach that threshold, *ARID4B* and *SPTBN4*. These can then be examined for expression patterns (https://gtexportal.org/home/) to see if they are expressed in a cell or

tissue type that may be relevant. *FAM8A1* and *ZRANB1* are expressed at low levels across all tissues, *ARID4B* is expressed in most tissues but minimally expressed in whole blood while SPTBN4 is exclusively expressed in brain, specifically in the cerebellum. Searching PubMed for each gene revealed *ZRANB1* encodes the deubiquitinase ,TRABID, and deletion of *ZRANB1* in dendritic cells caused epigenetic modifications of *IL12b* inhibiting expression after toll-like receptor stimulation[74]. Since IL12B is a component of both IL-12 and IL-23 cytokines, each involved in the response to *Mycobacteria*, *ZRANB1* should be considered a candidate worthy of further study.

2.6 Genotype – Phenotype correlations

Additional filtering of the target variant list may be accomplished by incorporating the patient phenotype and knowledge about genes involved in *M. tb* infection. VarElect (ve.genecards.org) is part of the GeneCards suite of NGS analysis tools developed by the Weissman Institute[154]. This web-based tool (figure 2.3) integrates genes and known disease associations, as well as indirect connections with intermediate genes to rank the input genes by relevance with input phenotypes. Ideally, this may help narrow the list of genes down to a manageable number for further consideration with some genes excluded – an immune deficiency is less like to be caused by a variant in myosin than an immune gene for example. A second web-based tool is the Human Gene Connectome (HGC) from Rockefeller University (hgc.rockefeller.edu)[73]. This allows input of a list of core genes – those known to be associated with some part of the patient's phenotype – and a list of "Genes of Interest" – i.e. those

identified with variants (Figure 2.4). The output of HGC is a ranking of genes with degrees of separation from the input core genes. This can be used to identify genes related to known genes which may not have been obvious at first glance. Using these two tools, the gene list of a few hundred may now be narrowed down to a manageable number such as 5-10. The question then is, how to evaluate the remaining variants to determine which to pursue functional studies?

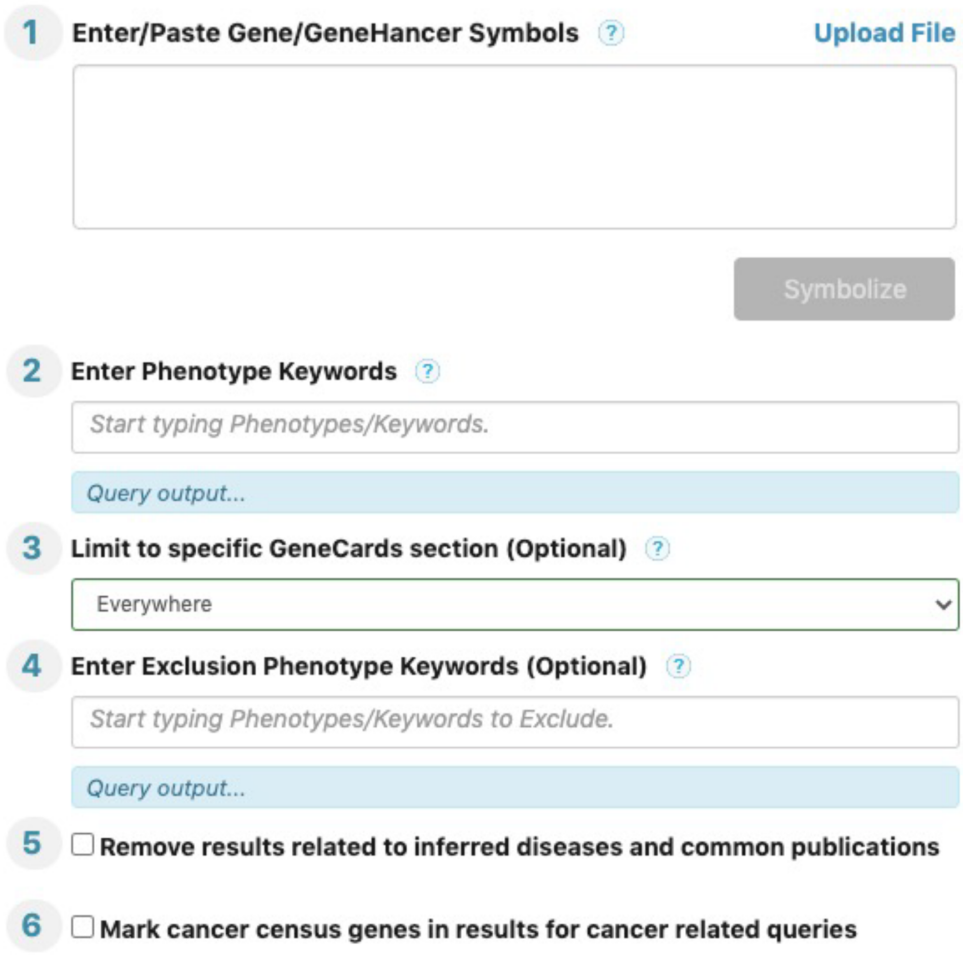


Figure 2.3. VarElect online tool for connecting patient phenotypes with identified genes with variants. VarElect inputs a list of genes containing variants in one field and phenotype keywords in another. Using datamining of literature as well as disease databases, the tool returns a ranked list of genes and the associated phenotypes.

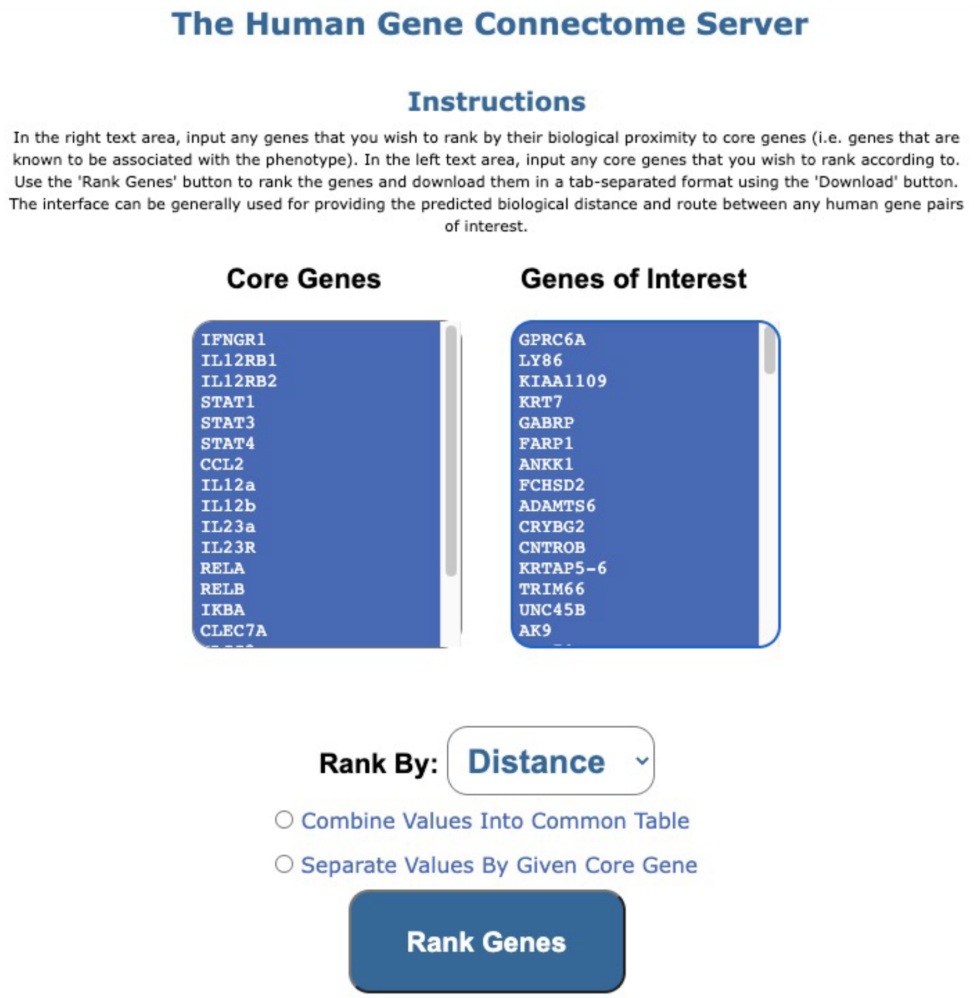


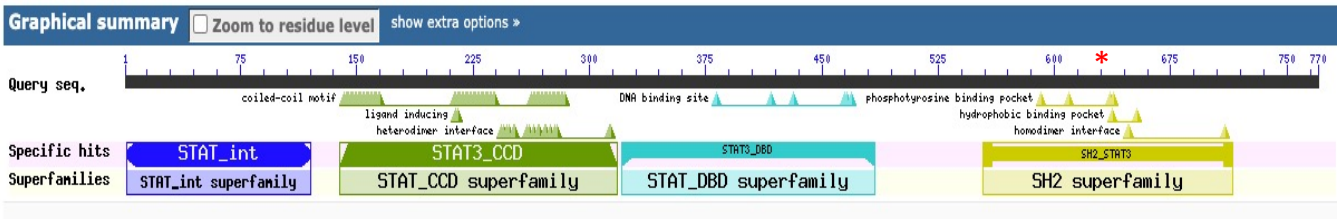
Figure 2.4. Human Gene Connectome connects genes within suspected pathways with patient genes with variants. Human Gene Connectome inputs a list of core genes known to be related to the phenotype of the patient in one field and the list of genes with variants in another. The output is a list of genes ranked by degree of separation between each gene and the core list entered. This may highlight additional genes in a pathway not previously considered.

2.7 Evaluating amino acid substitutions

While LoF mutations cause loss of protein product from the mutant allele, missense mutations may or may not alter the function of the protein. Three things to consider for

missense changes are the physicochemical properties - size, side chains, hydrophobicity, charge – of the reference and variant AA, location within domains or functional sites of the protein and conservation across species or across the protein family. Numerous references or tables exist for individual amino acid properties, although one that is especially useful is http://russelllab.org/aas. It is a compilation of properties but also substitution preferences, and roles in structure and function. Identification of domain or functional sites of proteins can be obtained using National Center for Biotechnology Information (NCBI) Protein search (https://www.ncbi.nlm.nih.gov/protein/). One example of using the conserved domain database (CDD) search is highlighted in my identification of a STAT4 mutation, p.E626G, in a three generation family with disseminated coccidioidomycosis[120,122]. Using the CDD (Figure 2.5A), glutamic acid at amino acid 626 (E626) was identified as a highly conserved amino acid within the conserved SH2 domain of the protein.

A



B

Figure 2.5. Use of protein domains and conservation to predict damaging variants. A. Conserved domains for STAT4 showing a graphic with each conserved domain highlighted as well as specific features within the domains. E626 is noted by a red asterisk. B. BLAST alignment of STAT4 across species showing conservation of the amino acid.

Further, it is located within the phosphotyrosine binding pocket (Figure 2.5B) and by homology is predicted to form Hydrogen bonds with the phosphotyrosine[138] on the partner STAT molecule, forming an active dimer which translates to the nucleus activating transcription. Finally, conservation of the amino acid across protein families such as signal transducers and activators of transcription (STATs) can be identified by inputting the reference protein sequence into the NCBI BLAST algorithm (https://blast.ncbi.nlm.nih.gov/Blast.cgi) and aligning selected proteins from the output using the multiple sequence alignment tool found on the results page.

This provides not only conservation of the individual amino acid where the mutation occurs, but context of the surrounding amino acids (Figure 2.5B). If they are highly conserved across the protein family or across species then that region may be important even if it does not fall with a conserved domain annotated in CCD.

2.8 Additional considerations

In some cases the mutation is not easily identified by WES. This can be due to the mutation itself – a deletion spanning a full exon will not be captured by exonic capture probes resulting in sequence from only the wild-type allele[63], a variant in an intron creating a new splice site leading to the incorporation of a new exon in the transcript such as seen in X-linked chronic granulomatous disease[132] or a mutation in an untranslated exon not included in the exome capture affecting splicing such as seen in *IKBKG* encoding NEMO[66]. In the case of patients in whom the phenotype matches a known disease gene but no mutation is identified, analysis of full-length cDNA from primary patient cells is a useful tool to examine transcript sequence and integrity. A heterozygous polymorphism identified by genomic sequence should also appear heterozygous by cDNA sequencing. Mutations affecting transcript stability will result in the loss of the heterozygosity at the cDNA level indicating the presence of transcript from only one allele as seen in GATA2 deficiency[64] while mutations creating an in-frame insertion or deletion due to exon skipping or inclusion will show heterozygous sequence for the mutant transcript. Further, Consideration of the effect of the variant on splicing is critical. Mutations which alter the canonical splicing motifs at the start and end of an intron (GT and AG

respectively) are classified in variant annotation software. Frequently missed however are those variants that alter extended motifs required for splicing (Figure 2.6): variants affecting the first or last base of an exon, the +5 site at the 5' end of the intron which is a critical nucleotide for U1 small nuclear ribonucleoprotein (snRNP) binding[96], the branch point A site (bound by U2 snRNP[110]) 18-40 nucleotides 5' of the downstream exon[174], shortening of the polypyrimidine tract (Y_n) prior to the exon[19] or the -3 site at the 3' end of the intron which binds to U2AF (reviewed in Wahl[172]).

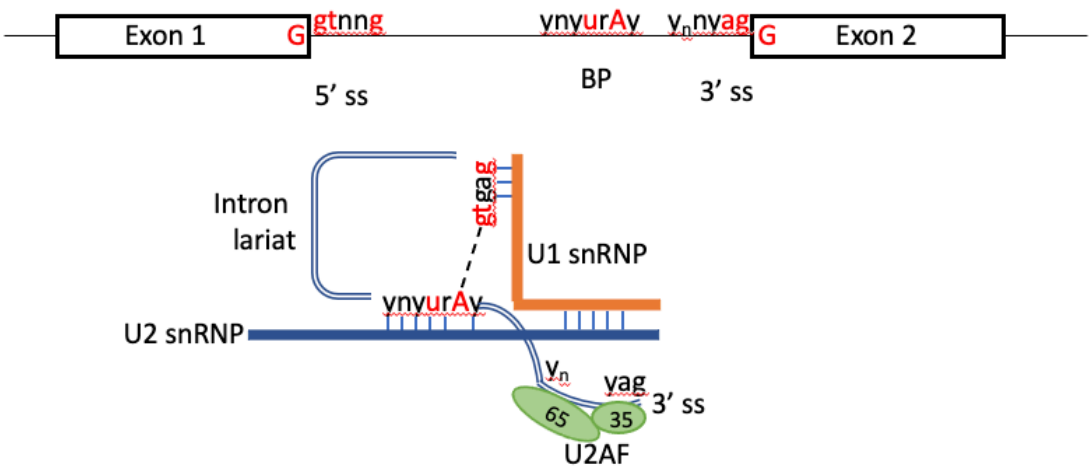


Figure 2.6. Splicing motifs which may be mutated leading to disease. Top linear cartoon of two exons. Thicker boxes indicate exonic sequence while thin line denotes intronic sequence. Bases shown in red have been identified with mutations leading to alteration of splicing. ss – splice site, BP – branch point. Lower – cartoon showing assembly of the intron and associated snRNPs recognition. U1 snRNP binds to both 5'ss and U2 snRNP while U2 snRNP recognizes the BP . U2AF complex binds to polypyrimidine tract (y_n) and 3' splice site.

Changes at any of these may affect proper splicing and should be evaluated by cDNA. An additional use of cDNA sequencing is in the case of pseudogenes (Ψ) with high degree of

identity to the canonical transcript. In these cases, sequences obtained by WES do not map to the reference sequence correctly resulting in no analyzable read data for the duplicated exons. When a patient has a phenotype consistent with *IKBKG* (NEMO) mutation, cDNA sequence is the preferred technique to identify the mutation. The use of cDNA removes the issue of the pseudogene since *IKBKG*Ψ lacks the first 3 exons including transcription and translation starts and is not expressed.

2.9 Somatic mutations and patient mosaicism

Rarely, a patient may have a milder phenotype than expected if they carry the mutation in only a portion of their cells, called mosaicism. This may happen when a mutation occurs early in development affecting only a subset of tissues. One example of this is seen in somatic mosaicism for mutations in the pleiotropic transcription factor, STAT3. Two men were identified with phenotypes resembling STAT3-deficiency, however both then had children with extensive STAT3-deficient phenotypes. Careful examination of isolated cells and tissues revealed the presence of the mutation in only a subset of cells with enough wild-type protein to ameliorate many symptoms[65]. The converse of this is patients who inherit a mutation and have a somatic reversion event in a progenitor cell resulting in a milder phenotype. A recent case of this was identified in a patient with an inherited *IL2RG* mutation encoding the common cytokine receptor gamma chain responsible for signaling via IL-2, IL-4, IL-7, IL-15 and IL-21. The patient had a reversion in CD8+ T cells but not in the NK-lineage resulting in a recurrent HPV driven

nasal carcinoma which resolved after successful hematopoietic stem cell transplant restored full NK-cell immunity[87].

In summary, next-generation, high-throughput sequencing has made identification of variants within patients substantially faster. Analysis of the sequencing results however is complicated and, while the list of variants may be filtered and ranked by numerous *in silico* tools, ultimately, well-planned functional studies need to be performed to establish causation. My recent review of newly identified genes affecting innate immunity[67] highlights the amount of work required to validate variants in novel genes and the reason thoughtful approaches to narrowing variants for further testing is useful.

Chapter 3: STAT3 and fungal disease

3.1 Introduction

Dominant-negative mutations in the signal transducer and activator of transcription-3 (STAT3) are the genetic cause of a multi-systemic syndrome commonly referred to as STAT3-Hyper-IgE (STAT3-HIES) or Job's syndrome[61,102]. Patients present with elevated levels of serum IgE, recurrent staphylococcal infections, skeletal disorders including scoliosis, craniosynostosis, and minimal trauma fractures, and recurrent pneumonias leading to formation of pneumatoceles, air filled lesions in the lung parenchyma[61]. Patients often develop fungal infections including mucocutaneous candidiasis, a fungal infection of the oral and vaginal mucosa. Pulmonary aspergillosis is not uncommon and is usually seen after structural lung damage caused by repeated pneumonias[168]. The identification of causative mutations and patient phenotypes has driven the study of STAT3 signaling pathways in health and disease. STAT3 signaling occurs in response to extracellular cytokine signals transduced by transmembrane receptors. In its inactive state, homo-dimeric STAT3 is found in the cytoplasm. Upon cytokine signal transduction by receptor engagement, Janus kinase (JAK) molecules phosphorylate the receptor allowing STAT binding and phosphorylation by JAKs on a conserved tyrosine residue. Depending on the cytokine, STAT3 homo-dimerization (i.e. after IL-6) or hetero-dimerization with other STATs such as STAT4 (i.e. after IL-23) occurs followed by nuclear translocation, DNA binding and activation of transcriptional targets. Dominant mutations in

STAT3 are predicted to result in 75% of homodimers and 50% of heterodimers containing a mutant protein and correspondingly decreased STAT3 driven transcription (Figure 3.1).

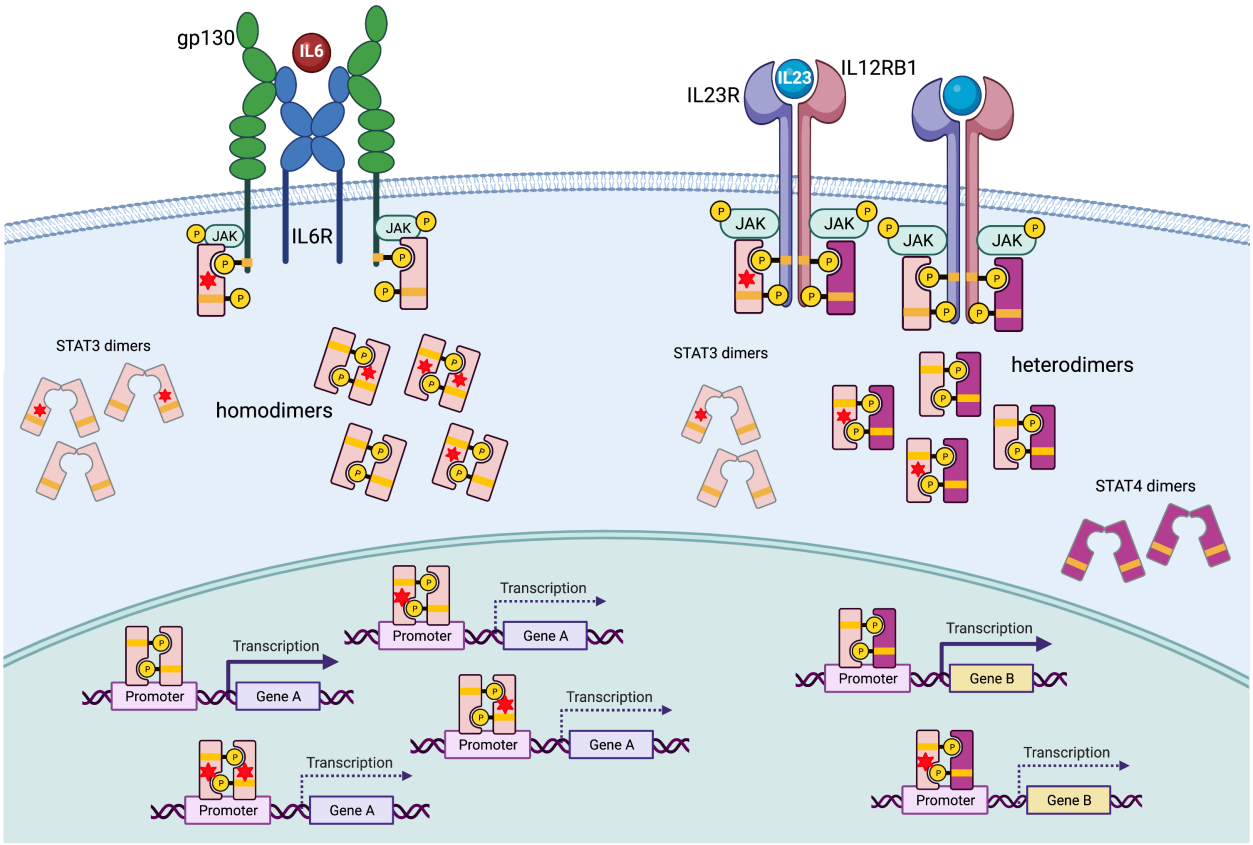


Figure 3.1. STAT3 signaling after cytokine receptor engagement. A. IL-6 signals through a receptor consisting of two IL6R and two gp130 chains. Both IL6R and gp130 activate STAT3 leading to phospho-STAT3 homodimers which enter the nucleus and activate transcription. In patients with dominant-negative STAT3 mutations, half of all unphosphorylated STAT3 molecules are mutant (red star) and 75% of phospho-STAT3 dimers are mutant. B. IL-23 signals through an IL23R/IL12RB1 heterodimeric receptor. This leads to recruitment and phosphorylation of both STAT3 (IL23R) and STAT4 (IL12RB1) and subsequent heterodimeric signaling complex of phospho-STAT3/STAT4. Half of the heterodimers contain a mutant STAT3 molecule. Figure made in BioRender.

3.2 Methods

3.2.1 PCR and Sanger sequencing

PBMCs were isolated from whole blood by means of Ficoll-Hypaque gradient centrifugation and lysed either in STAT-60 (Tel-Test, Friendswood, Tex) followed by RNA isolation or with the PureGene DNA Isolation kit cell lysis solution (Qiagen, Valencia, Calif) for DNA extraction. Amplification of full-length cDNA was performed with Superscript III One-Step RT-PCR kit (Invitrogen) using cDNA primers 54F- 5'-CGCCCGTCCCCGGCACACG-3' and 2764R- 5'-AGGAGGGCAGGGGAACAAAACAACACAAGA-3'.

For genomic DNA amplification Platinum Taq HI Fidelity supermix (Invitrogen) and the following primer sets were used: Exons 12 - 14 amplification: 12F amp- 5'-TAGTTTAAAGAAATGCCCAGGAGCACAGAGGTTTT-3', 14R amp- 5'-TTTGGCCTGAAGTGACTTTTTGGAATAACTACAGC-3'. Exons 15-17 amplification used 15F amp- 5'-CCAAAAGAATTCATTGGAGTCCATAAGTGTAAGGT-3' and 17R amp- 5'-AAGAGTCAAGTAGTACATTTTCAGCTTGGGGTCAA-3'.Sequencing was performed with Big Dye Terminators version 3.1 (Applied Biosystem) and the following primers: cDNA: 222F- 5'-GCGGCAGTTTCTGGCCCCTTGGATTG-3', 511F- 5'-GGCCAGGCCAACCACCCCACAGC-3', 1170R- 5'-AGTGAACTGGACGCCGGTCTTGATGACGAG-3', 958F- 5'-CAGACCCGTCAACAAATTAAGAAACTGGAG-3', 1391F- 5'-ATGGGGGCCGAGCCAATTGTGATGCT-3', 2199R- 5'-TGGCCGACAATACTTTCCGAATGCCTCCTC-3', 1900F- 5'-AGTAAGGAGCGGGAGCGGGCCATCTTG-3', 2232F- 5'-TAGCGCTGCCCCATACCTGAAGACCAAGTT-

3'. Genomic: 13F seq- 5'-CTGGGGACGTTGCAGCTCTCAGAGGGTAAGT-3' and 16F seq- 5'-CTCGCCTAGAGTTGGCAGCAGGTGTGGTT-3'. Reactions were purified using DTR spin plates (Edge Biosystems) and run on an ABI 3730 capillary sequencer (Applied Biosystems). Data was analyzed using Sequencher (Gene Codes).

3.2.2 Western blot analysis

Transfected cells were lysed in RIPA buffer with protease inhibitors. 20 μg total protein per sample was denatured and loaded on a 10% polyacrylamide gel. After electrophoresis, proteins were transferred to nitrocellulose membranes, blocked for 1 hour in TBS-t with 5% powdered milk then probed with specific antibodies. Membrane was probed using primary antibodies for STAT3a, total STAT3 or β-actin (all from Cell Signaling). Antibody dilution was 1:1000 in 5% w/v milk, 1X TBS, 0.1% Tween20, at 4^{o}C overnight. Secondary antibodies: Anti-rabbit IgG-HRP-linked & Anti-mouse IgG-HRP-linked (antibody dilution 1:1000 in 5% w/v milk, 1X TBS, 0.1% Tween20, 1 hour room temperature). Chemiluminescence detection was performed with Clarity™ Western ECL Substrate (Bio-Rad), using the ChemiDoc™ MP Imaging System (Bio-Rad) and quantification was obtained by densitometry image analysis using Image Lab 6.0 software (Bio-Rad).

3.2.3 Mouse Coccidiodes infection studies

Groups of 10 mice were infected intranasally under ketamine (80 mg/kg)-xylazine (8 mg/kg) anesthesia with 56 spores of *C. posadasii strain 1038* administered in 30 μl of sterile

saline dropwise into the nares with a micropipettor. Infectious doses were verified by plate culture of inoculum following infection of the mice. Mice were monitored daily for activity level and hydration and weighed weekly. Animals exhibiting dehydration, hunched posture, isolation from cage-mates, weight loss >25%, and lethargy or weakness were euthanized as needed prior to planned end of study. Studies were terminated on day 70 post-infection.

3.3 Cytokine signaling through STAT3

STAT3 is a pleiotropic transcription factor, active throughout the life of the organism and across multiple tissue types including immune, vascular, skeletal and connective tissue (reviewed in Tsilifis[162]). STAT3 is activated downstream of cytokine receptors including IL-12 family members IL-23 and IL-27, gp130, the common cytokine receptor involved in IL-6, IL-11, IL-27, Oncostatin-M and LIF signaling, IL-21, and IL-10. Since STAT3 is the nexus of multiple pathways, identification of individuals with specific cytokine or receptor mutations or mouse models of cytokine / receptor knockouts has elucidated the role of individual cytokine signaling pathways driving specific portions of the STAT3-HIES phenotype. Patients with bi-allelic IL-6 receptor (*IL6R*)[152,108] or gp130 (*IL6ST*)[141,144,18] mutations have bacterial skin and lung infections, eczema, and elevated IgE, all components of the STAT3-HIES phenotype. Unlike STAT3-HIES, neither *IL6R* nor *IL6ST* patients have reported fungal infections suggesting signaling through gp130 is not directly related to fungal immunity. Additional cytokines signaling through gp130 include LIF, OSM and IL-11. In each case, specific receptor mutations have been identified in humans. *LIFR* mutations cause a severe skeletal dysplasia, Stüve-Wiedemann syndrome which is

often embryonically fatal[23]. *OSMR* mutations, which alter the signaling of OSM and IL-31 cause familial primary localized cutaneous amyloidosis[5]. *IL11RA* mutations phenocopy some of the bone phenotypes seen in STAT3-HIES with patients presenting with craniosynostosis and supernumerary teeth[112]. An *IL6ST* mutation affecting only IL-11 signaling also presented with craniosynostosis and retained primary teeth[142] two common features of STAT3-HIES. Roughly half of reported *IL21R* mutated individuals have fungal infections, mostly *Pneumocystis jerovecii* pneumonia or esophogeal- or systemic- candidiasis[15] suggesting signaling via IL-21 is important for antifungal immunity. IL-21R is expressed on hematopoietic cells as well as mucosal epithelia[17] and keratinocytes[30]. Replicating the fungal susceptibility seen in IL21R deficient patients, a murine model of *Pneumocystis* infection demonstrated IL21R signaling in CD4+ T-cells is required for STAT3 driven IL-22 production[37]. IL-21 induced activation of STAT3 controls the epigenetic status of the *IL22* promoter and triggers IL-22 production by CD4+ T-cells[181]. In intestinal epithelial models, the immune cell-mediated production of IL-22 acts on epithelial cells leading to production of proteins required for epithelial barrier maintenance[139] (reviewed in Alcorn[3]). In the absence of IL21R, exogenous IL-22 could drive the production of the fungicidal peptide, cathelicidin[37] suggesting the loss of IL-22 production may account for fungal infections in IL21R deficiency.

3.4 Pathways highlighting non-immune phenotype of STAT3-HIES patients

STAT3-HIES patients are also notable for connective tissue disorders including hyperextensibility, scoliosis, tortuous arteries and aneurysms. These phenotypes are similar to

those seen in Loeys-Dietz syndrome[90] caused by activating mutations of the TGFβ signaling pathway including receptor (*TGFBR1, TGFBR2*)[90,91], cytokine (*TGFB2*[86], *TGFB3*[130]) or transcription factor (*SMAD3*[164]). Work by Lyons et al.[92] elucidated the connection between TGFβ signaling and STAT3 with the discovery of bi-allelic mutation in *ERBB2IP* encoding ERBIN. They demonstrated STAT3-driven *ERBB2IP* transcription and formation of a STAT3 - ERBIN - SMAD2/3 complex in the cytoplasm, sequestering SMAD2/3 and thereby inhibiting TGFβ signaling[92] (Figure 3.2). Importantly, Loeys-Dietz patients do not display infection susceptibilities seen in STAT3-HIES. TGFβ signaling has a crucial role in epithelial-mesenchymal transition (EMT) of alveolar epithelia cells (reviewed in Willis & Borok[177]), possibly contributing to the development of bronchiectasis and pneumatocele formation, common sequelae of the recurrent pneumonias seen in STAT3-HIES.

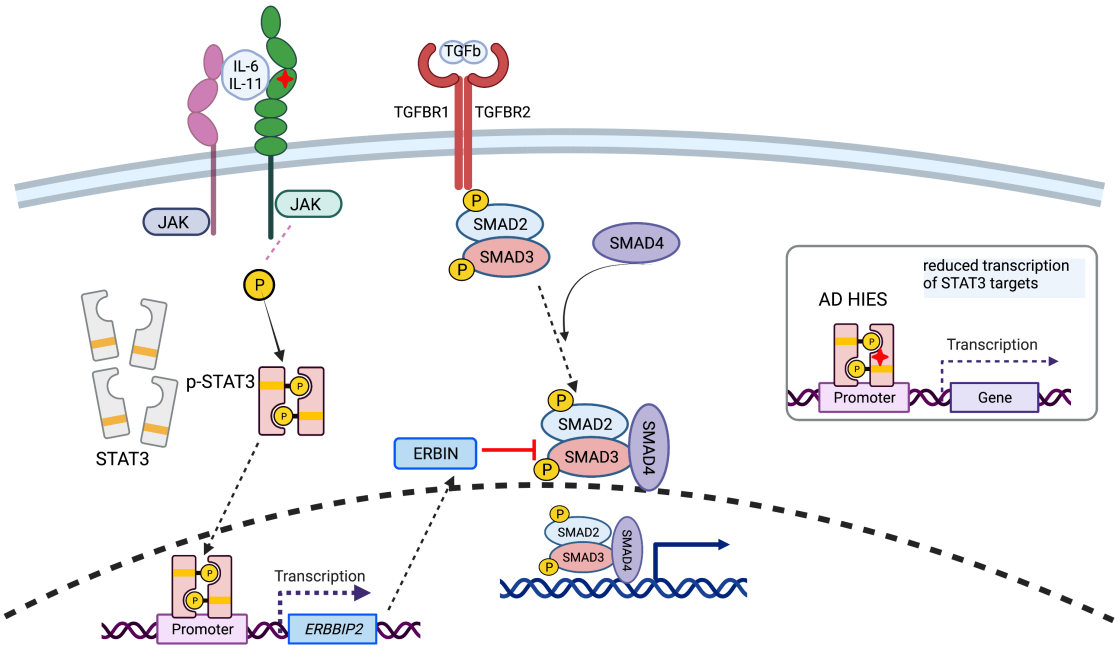


Figure 3.2. STAT3 activates transcription of *ERBB2IP* encoding ERBIN. ERBIN acts as an inhibitor of TGFβ signaling by sequestering SMAD2/3 in the cytoplasm. STAT3 mutations lead to decreased ERBIN expression and increased TGFβ signaling. Figure made in BioRender.

3.5 STAT3 signaling and Th17 cells

Shortly after the identification of STAT3-HIES, it was reported that STAT3-HIES patients have decreased numbers of Th17 cells[93,101], hypothesizing reduced Th17 numbers lead to fungal susceptibility. Th17 cells were identified in 2005[56,84,118] and have been associated with fungal immunity, mostly in relation to *Candida albicans*-induced chronic mucocutaneous candidiasis (CMC)[21] and their role in mucosal infections (reviewed in Khader[78]). Manel[97] and Volpe[171] demonstrated the differentiation of naïve CD4+ T cells to Th17 cells in the presence of TGFβ, IL-

6, IL-1β, and IL-23, most of which require STAT3 for signaling. Th17 cells produce IL-22 in response to intestinal microbiota. Epithelial cells respond to IL-22 with increased production of bactericidal compounds, strengthening of tight junctions and alterations to the mucus layer all of which work together to prevent pathogen translocation from the lumen of the intestine[139]. More recently, in a mouse model of oral candidiasis, STAT3-dependent IL-22 signaling was shown to specifically enhance proliferation of the basal epithelial layer in the oral mucosa increasing the ability of IL17R expressing epithelial cells to respond to fungal challenge[2]. Taken together, these studies suggest STAT3 signaling itself is required for fungal immunity, particularly at the mucosa or epithelia and reduced Th17 cells may be a symptom of decreased STAT3 signaling and coincident to, rather than causative of, fungal infection.

3.6 STAT3 levels are finely tuned

Patients with autosomal dominant STAT3 mutations have severely reduced STAT3 signaling with 75% of STAT3-STAT3 homodimeric and 50% of STAT3-STAT1/4 heterodimeric STAT molecules containing a mutant STAT3 protein (as shown in Figure 3.1). This level of signaling disruption leads to a syndromic phenotype spanning multiple organ systems beyond immunity. Complementing the STAT3-dependent pathways identified by cytokine receptor mutations, there are additional mutations in both mice and humans which affect the global level of STAT3 protein. Using the amount of wild-type STAT dimers as a reference, we can estimate level of STAT3 signaling required in different systems based on phenotype. Heterozygous knockout mice were generated with targeted disruption of the murine *Stat3*

locus, replacing exons 20-22 (of 24 exons) with a NEO cassette[156]. Crossing heterozygous mice resulted in no *Stat3*$^{-/-}$ offspring despite appropriate 1:2 ratios of wild-type and heterozygous pups[156]. Examination of embryonic development showed presence of *Stat3*$^{-/-}$ egg cylinders at embryonic day 6.0 (E6.0) but complete loss of all *Stat3*$^{-/-}$ embryos by E7.5 indicating a critical role of Stat3 in embryogenesis prior to hematopoiesis[156]. They did not further explore the phenotype of *Stat3*$^{+/-}$ animals.

Bi-allelic mutations of the transcription factor, ZNF341 which binds to the *STAT3* promoter and activates transcription, result in significantly decreased *STAT3* transcript and protein levels in patients cells. In these patients, all STAT3 protein made is wild-type, however the level of protein made is significantly reduced[6,44]. These patients exhibit a phenocopy of the STAT3-HIES phenotype indicating sufficient STAT3 levels for development but insufficient levels for immune system maintenance, or skeletal and connective tissue development[6.44]. In these two series of patients, 5/11[44] and 5/7[6] had recurrent *Candida* infections including oral and esophageal thrush, CMC or cutaneous infections. Patients with ZNF341 deficiency had roughly 50% the level of *STAT3* transcript in PBMCs, EBV lines and virally-transformed T-cells as healthy volunteers while primary skin fibroblasts had slightly higher transcript levels[44]. STAT3 protein levels varied across cell types, with bulk PBMCs from an affected patient having roughly one quarter the total STAT3 protein of healthy controls[44].

Two males were reported with somatic mutations in STAT3 resulting in a mosaic mutation pattern in the patients but germline transmission to their children. In each case, the

phenotype was somewhat milder than traditional STAT3-HIES[65] with different cell lineages containing differing percentages of STAT3 mutant cells (Figure 3.3). Both men had chronic onychomycosis and one had CMC despite both having normal levels of Th17 cells[65]. It is worth noting that the patient with CMC appeared heterozygous in a biopsy of the gastric mucosa (stomach biopsy) (Figure 3.3) suggesting the digestive tract and oral mucosa contained more mutant STAT3 thereby affecting IL-22 and IL-23 signaling required for control of candida.

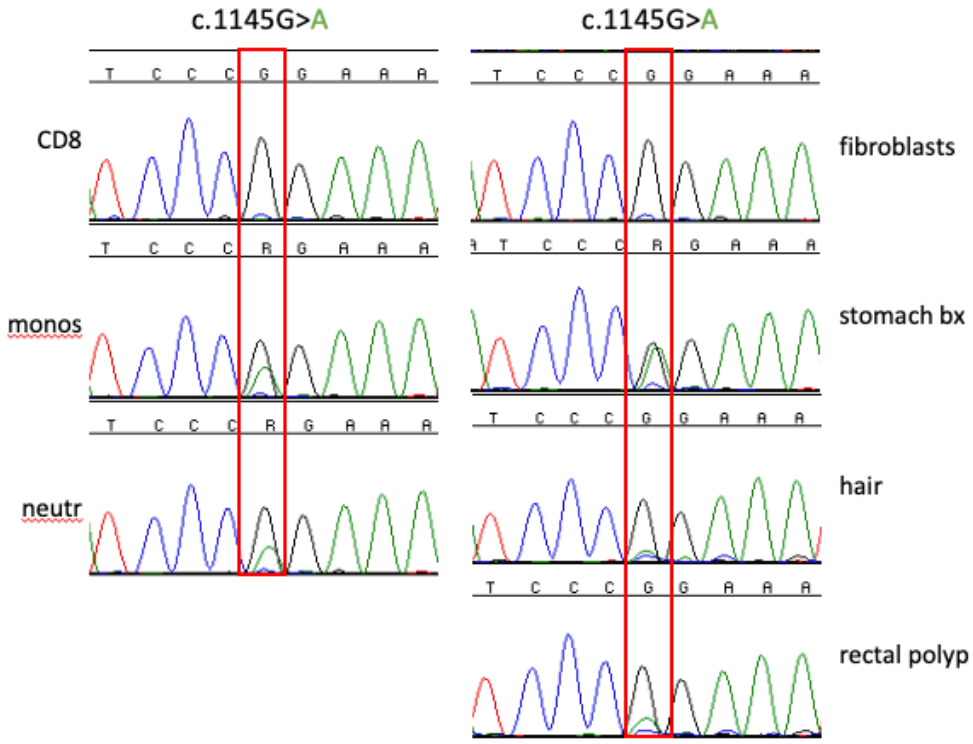


Figure 3.3. Somatic mosaicism leading to milder phenotype.
Mosaicism for STAT3 c.1145G>A in a patient with mild STAT3-HIES phenotype. A. Isolated hematopoietic lineages showing different levels of wild-type G (black peak) and mutant A (green peak). CD8 – CD8+ T-cells, monos – monocytes, neutr – neutrophils. B. Comparison of mutation levels across several tissue types. Mutant peak ranges from undetectable in fibroblasts to apparently heterozygous in stomach gastric mucosa biopsy.

As discussed in chapter 2, when considering pathogenicity of a genetic variant, one metric utilized is the constraint identified for that gene in large sequencing databases such as gnomAD (gnomad.broadinstitute.org). Generating an observed / expected ratio (o/e) of null alleles due to premature stop codons, frameshift or splicing mutations, the presence of purifying selection against these types of mutations can be demonstrated and calculated for each gene[76]. Genes with a low o/e are strongly intolerant of null alleles and include known haploinsufficient genes – those genes for which dosage is critical and protein derived from one allele is insufficient. For *STAT3*, the o/e ratio is 0.02 (90% CI of 0.01-0.1) (gnomAD database, accessed November, 2021) indicative of severe constraint against loss of function alleles. Taken together with the presence of fungal infections in patients with mutations in *ZNF341* and *STAT3* suggest there is a threshold level of functioning STAT3 required for adequate immunity against fungi.

3.7 STAT3 in disseminated coccidioidomycosis

The first patient identified with a monogenic immune deficiency and DCM was a young girl with STAT3-HIES[124] with two additional patients since reported (reviewed in Odio[113]). With the aim of identifying genetic susceptibility factors associated with DCM, whole exome sequencing data on 67 individuals with DCM were analyzed, searching for monogenic mutations accounting for the severe disease presentation. Using the techniques outlined in Chapter 2, the data were filtered for rare or novel variants predicted damaging, focusing on known immune genes and genes related to immunity. Only two patients were identified with monogenic

mutations, both carrying premature stop codons in *STAT3*. Importantly, neither of the patients in the DCM cohort exhibited syndromic features of STAT3-HIES, in fact, both were healthy until becoming infected with *Coccidioides*.

The first patient identified in our cohort with a monogenic mutation in a known immune gene was heterozygous for *STAT3* c.1267C>T leading to p.R423*. She became infected during pregnancy, disseminated to her central nervous system (CNS) and despite antifungal therapy, succumbed to her infection 5 years later. On autopsy she had severe, destructive lung disease with cavitary lesions in both upper lobes and dense consolidation of the right lower lobe. The destructive lung disease is a common feature in STAT3-HIES where defective STAT3-directed remodeling frequently leads to cavitary lesions and pneumatoceles[42,43]. Strikingly, other organs appeared normal with no signs of dissemination except for the brain which displayed hyphae (Figure 3.4), a rare, but not unseen, form of the fungus, previously reported in CNS[182], lung[140], and patients with Type 2 diabetes[107].

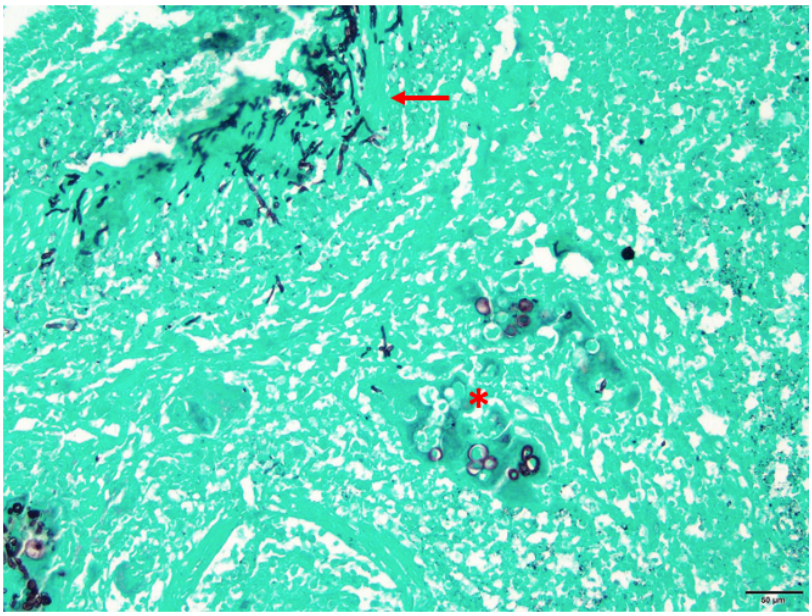


Figure 3.4. Fungal forms in brain of STAT3 haploinsufficient DCM patient. Autopsy brain section from DCM patient with haploinsufficient mutation, STAT3 p.R423*. GMS fungal stain shows both hyphal forms (upper left, arrow) in the artery wall and yeast forms (asterisk) in clusters outside the artery. GMS fungal stain, 200X, size bar = 50 μm. Figure courtesy of Dr. David Kleiner, NIH.

The second *STAT3* patient was a Native American man who presented with DCM meningitis. The disease was treated with ongoing antifungal therapy and the patient was stable for several years until succumbing to infection. Analysis of WES data on the second patient identified a heterozygous c.247C>T causing p.R84*. Samples were not available on either patient to perform functional studies however, given the early stop codon and continued splicing of the transcript, these mutations are predicted to lead to nonsense mediated decay (NMD) of the mutant transcript.

3.8 Demonstration of STAT3 haploinsufficiency

A third patient who carried the same mutation as the first patient above, c.1267C>T, p.R423*, was identified with a 50 year history of disseminated fungal infection after traumatic *Trichophyton* inoculation. EBV transformed B-cells from this patient were used for functional studies. While the patient was heterozygous in genomic DNA, examination of full-length cDNA transcripts revealed only wild-type sequence consistent with NMD of the mutant transcript (Figure 3.5A, top). To confirm the presence of the mutant transcript, patient EBV cells were treated with 100 ng/ml Emetine, a proteasomal inhibitor which decouples transcription and translation and inhibits NMD. Under these conditions, full-length cDNA derived from both alleles was present (Figure 3.5A, bottom). In contrast, a STAT3-HIES patient carrying a missense mutation in STAT3, p.V713M which is not associated with splicing, demonstrates presence of both alleles before (Figure 3.5B, top) and after (Figure 3.5B, bottom) emetine treatment.

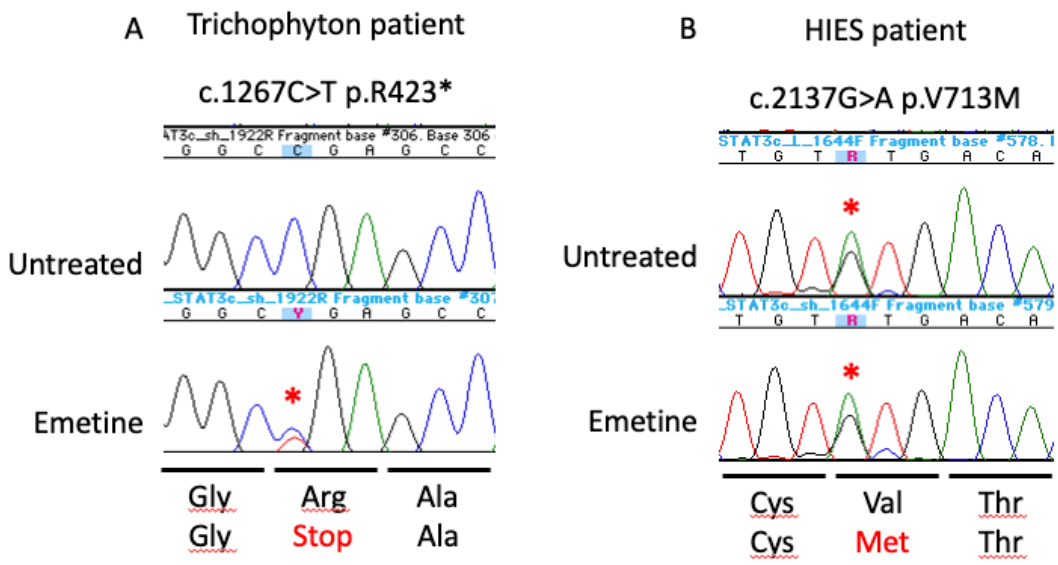


Figure 3.5. *STAT3* cDNA sequence from EBV cell lines. A. Sequence from a patient with disseminated Trichophyton carrying the same STAT3 p.R423* mutation as the first STAT3 DCM patient in our cohort. Top panel shows untreated cells with only wild-type transcript presence. Lower panel is sequence from cells treated for 4 h ours with 100 ng/ml emetine, blocking NMD. The mutant transcript (red asterisk) is present after emetine treatment. B. *STAT3* cDNA sequence from a STAT3-HIES patient with DCM carrying a missense mutation, p.V713M, not predicted to alter transcript stability. In untreated and emetine treated cells both mutant and wild-type transcripts (red asterisks) are present.

Western blot analysis of lysates from EBV-B cell lines revealed significantly decreased levels of STAT3α protein in the p.R423* patient while the p.V713M patient had STAT3 protein levels comparable to healthy controls (Figure 3.6A, P=0.0007). Normalizing the STAT3α/β-actin ratio to healthy controls quantified the p.R423* patient STAT3 levels as approximately 1/3 of controls (mean 32%, range 24-45%, P=0.0001) (Figure 3.6B).

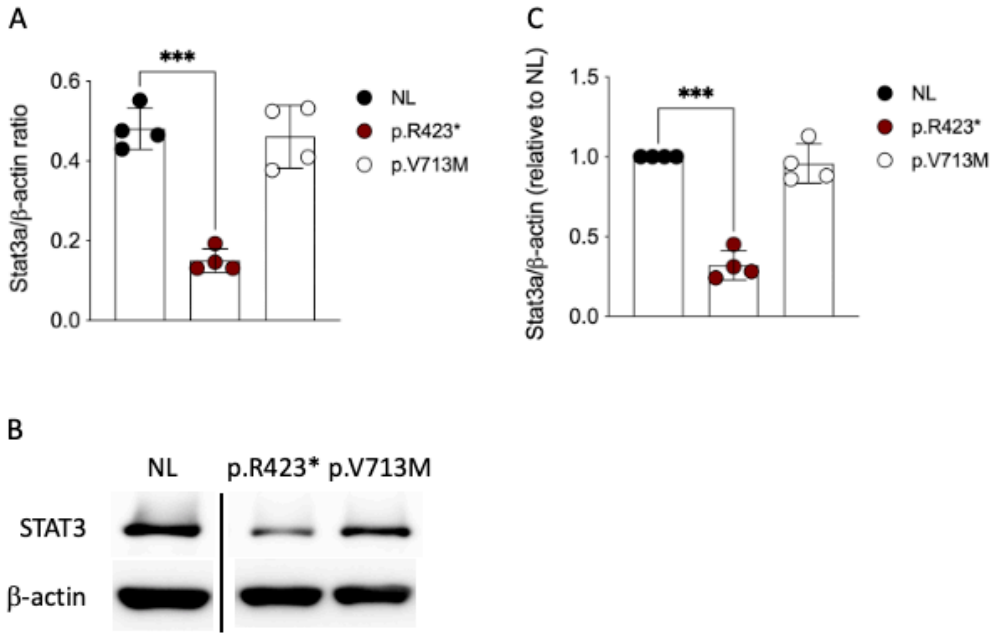


Figure 3.6. STAT3 protein levels in EBV-transformed B cells. Western blot analysis of STAT3a protein levels in healthy controls (NL), a patient carrying STAT3 p.R423* causing NMD and a patient heterozygous for p.V713M which does not alter transcript stability. A. Ratio of densitometry measurement of STAT3α and β-actin bands detected by Western blot. Data are from four independent experiments. *** P = 0.0001. B. Representative Western blot from one experiment. C. Data from A with the STAT3α/β-actin ratio expressed as percentage of healthy controls. Data courtesy of Dr. Vasileios Oikonomou, NIH.

Having demonstrated loss of expression of the mutant allele by cDNA analysis and decreased protein levels by Western blot, the question remained if the decreased STAT3 levels were related to the disseminated fungal disease. To generate heterozygous *Stat3* knockout mice we crossed *Stat3*^{flox} mice with the ubiquitously expressed CMV-Cre mouse. The result is mice with one deleted allele and one wild-type allele (Figure 3.7).

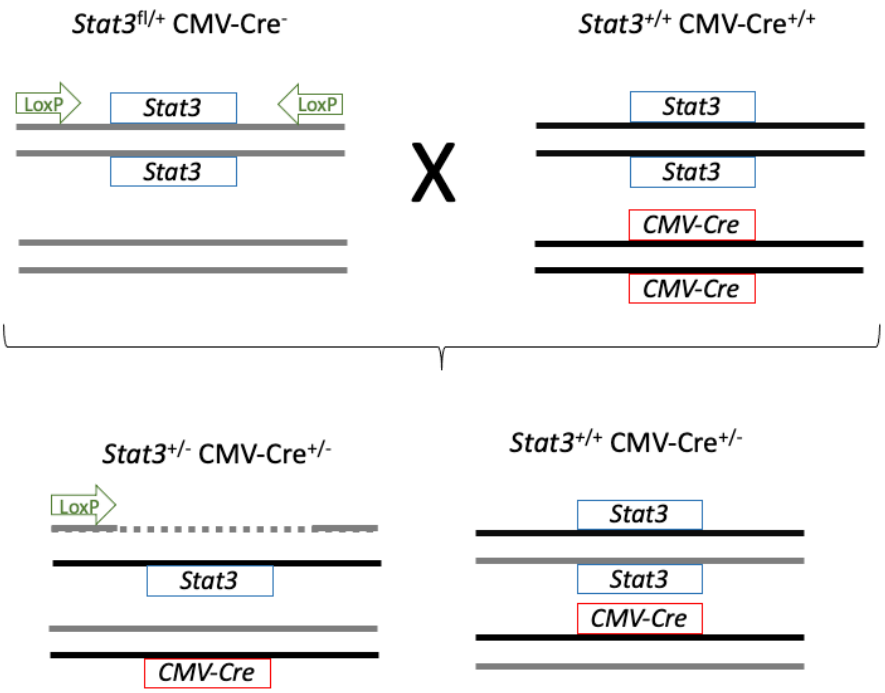


Figure 3.7. Strategy for generating Stat3 haploinsufficient mice. Mice heterozygous for a floxed-Stat3 allele (Stat3^{fl})and not carrying CMV-Cre (top left) were crossed with mice wild-type for Stat3 and homozygous for CMV-Cre (top right). The offspring (bottom) are either heterozygous for *Stat3*^{flox} and CMV-Cre causing deletion of the *Stat3*^{fl} allele while maintaining the wild-type allele, or homozygous for wild-type *Stat3* and carrying the CMV-Cre allele allowing the use of littermate controls.

Using Stat3flox^{-} CMV-Cre^{+} littermates as controls, the mice were infected with 56 arthroconidia of the attenuated *Coccidioides posadasii* strain 1038 (Cp1038)[147]. A 10 week survival study was performed with animals observed daily. Kaplan-Meier survival curve (Figure 3.8A) shows decreased survival of the *Stat3* haploinsufficient mice compared to littermate controls (P = 0.0302) despite similar terminal fungal burdens in lung and spleen (Figure 3.8B).

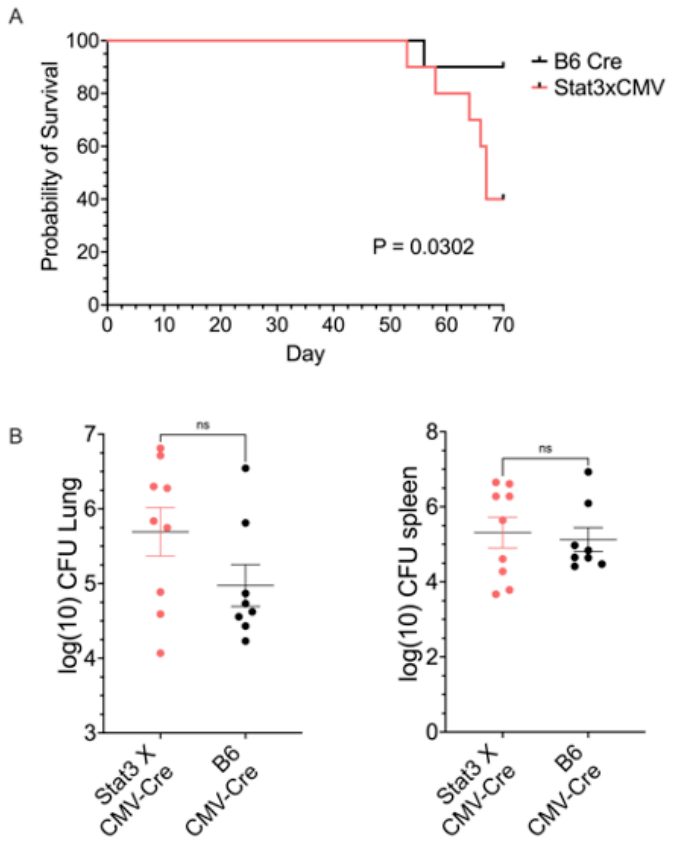


Figure 3.8. Intranasal infection of Stat3-haploinsufficient mice with Cp1038. A. Kaplan-Meier survival plot of Stat3-haploinsufficient (Stat3xCMV) versus C57BL/6 CMV-Cre (B6 Cre) mice. P = 0.0302 Kruskall Wallace. B. Log_{10} colony forming units (CFU) in lung and spleen as a measure of fungal burden at time of death or at conclusion of study on day 70. Data courtesy of Dr. Lisa Shubitz, University of AZ.

Additional experiments are required to explore timing of fungal expansion in the two groups as well as timing of dissemination from the lungs to spleen and central nervous system. These are important since both STAT3 haploinsufficient patients as well as the first two identified STAT3-HIES patients with DCM had dissemination to the CNS[113,124].

3.9 Conclusions

As highlighted in the mouse studies above, additional work is required to determine hematopoietic versus non-hematopoietic contribution of STAT3 and the role of STAT3 in development and immunity. Comparing the genetic causes of reduced functional STAT3 protein provides a spectrum model of STAT3 disease, with the severity of phenotype inversely proportional to amount of functional STAT3 (Figure 3.9). This supports STAT3 haploinsufficiency in these two DCM patients as the monogenic cause leading to their DCM.

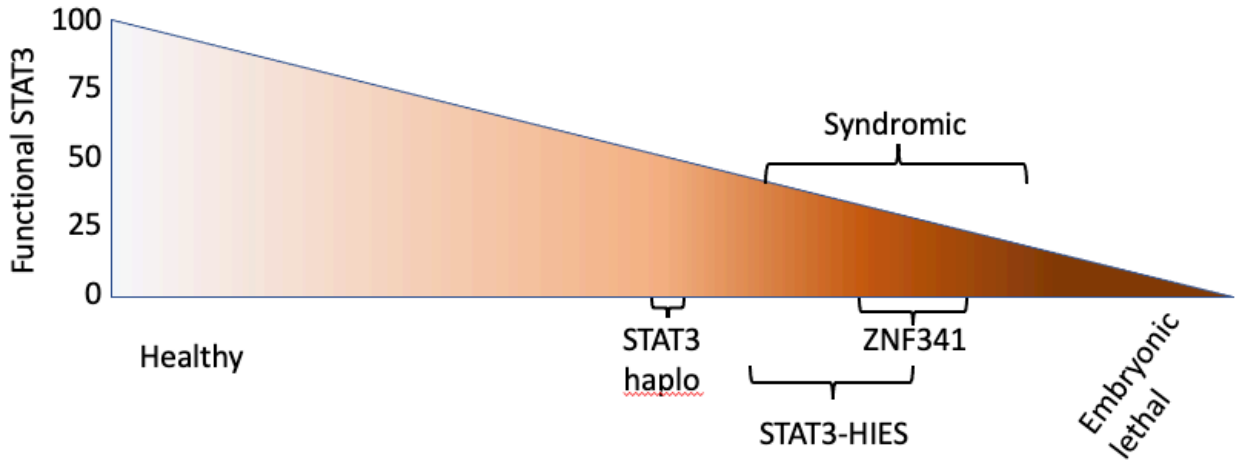


Figure 3.9. Model of disease spectrum across mutations affecting STAT3 quantity and quality. Severity of disease appears to be inversely proportional to the amount of functional STAT3 protein.

Chapter 4: Immunogenetics of severe or disseminated coccidioidomycosis

4.1 Introduction

Coccidioidomycosis, "Valley fever", is caused by infection with the soil-dwelling, dimorphic fungi, *Coccidioides immitis* and *C. posadasii*, prevalent in the Western United States, Mexico, Central and South America. Infection occurs after inhalation of arthroconidia, which morph into spherules that release hundreds of endospores 3-5 days later, becoming new spherules, establishing infection. Roughly 1/3 of the 150,000 US individuals infected annually will develop a symptomatic, self-limited pneumonia, while <1% of those will develop disseminated disease[11,46,69] mostly affecting meninges, bones, skin or joints[163]. The median hospitalization costs are more than $70,000[125], 2017 costs in California surpassed $700M[178] and total lifetime costs for 2019 Arizona cases were estimated to be $736M[51].

Similar to *Mycobacterium tuberculosis*, *Coccidioides* can cause illness in otherwise healthy hosts. Infection is more likely to disseminate in immunocompromised patients, including people with AIDS, those receiving chemotherapy, organ transplants, immunomodulatory biologics (e.g., anti-TNF and anti-IL-6), or women in the third trimester of pregnancy[114]. Interestingly, only a small number of DCM patients have identified mutations, all within the IL-12/IFN-γ and signal transducer and activator of transcription-3 (STAT3)

pathways[114]. Therefore, we used whole-exome sequencing (WES) to explore genetic risk factors in disseminated coccidioidomycosis.

4.2 Methods

4.2.1 Patients and Controls

Patients with biopsy-proven DCM but without comorbidities predisposing to DCM were enrolled in studies approved by the institutional review board of the National Institute of Allergy and Infectious Diseases. Samples were obtained from healthy donors from the National Institutes of Health blood bank. DNA was isolated from whole blood. PBMC were isolated using density centrifugation from heparinized whole blood. Autosomal-dominant Hyper-IgE patients (AD-HIES) and controls were enrolled in approved studies.

Validation cohort consisted of known coccidioidomycosis patients treated at UC Davis. Exclusion criteria included abnormal CBC, positive HIV, evidence of rheumatologic or other potentially immunocompromising disease or immunosuppressant medication use. Patient charts were reviewed. Disseminated infection was defined as disease outside the thorax. All patients had normal CBC, no HIV or evidence of invasive fungal disease other than coccidioidomycosis.

4.2.2 Whole exome sequencing and analysis

Whole exome sequencing in the discovery cohort was performed using the Ion Torrent AmpliSeq RDY Exome Kit (Life Technologies) and the Ion Chef and Proton instruments (Life

Technologies). Briefly, 100 ng of gDNA was used as the starting material for the AmpliSeq RDY Exome amplification step following the manufacturer's protocol. Library templates were clonally amplified and enriched using the Ion Chef and the Ion PI Hi-Q Chef Kit (Chef package version IC.4.4.2, Life Technologies), following the manufacturer's protocol. Enriched, templated Ion Sphere Particles were sequenced on the Ion Proton sequencer using the Ion PI chip v3 (Life Technologies). Read mapping and variant calling were performed using Ion Torrent Suite software v4.4.2. In short, sequencing reads were mapped against the University of California, Santa Cruz (UCSC) hg19 reference genome using the Torrent Mapping Alignment Program (TMAP) map4 algorithm. Single-nucleotide polymorphisms (SNPs) and insertions/deletions (INDELS) were called by the Torrent Variant Caller plugin (v.4.414-1) using the "Generic-Proton-Germ Line: Low Stringency" configuration. Only reads that were unambiguously mapped were used for variant calling. Variants were annotated using ANNOVAR (http://annovar.openbioinformatics.org/).

Validation cohort was sequenced using the Nextera Rapid Capture Exome Kit. Samples were sequenced at the Broad Institute on Illumina HiSeq sequencers using the Illumina Nextera exome capture kit. Each sample's sequencing reads were aggregated into a BAM file and processed through a pipeline based on the Picard set of software tools. The BWA aligner mapped reads onto the human genome build 37 (hg19). Single nucleotide polymorphism and insertions/deletions were jointly called across all samples using Genome Analysis Toolkit (GATK) HaplotypeCaller package version 3.4[165] to produce a version 4.1 variant callset file (VCF).

Variant call accuracy was estimated using the GATK Variant Quality Score Recalibration (VQSR) approach.

4.2.3 Genomic boundaries for variants

Variants falling within the below regions were extracted from the VCF files of the validation cohort and 1000 Genomes. Since *DUOX1* and *DUOXA1* are adjacent on the chromosome, a single interval was used to capture both genes. Variants were annotated using CADD v1.6[129] (https://cadd.gs.washington.edu/score). The following hg19 boundaries were used for variants within genes: *CLEC7A* chr12:10269376-10282868, *PLCG2* chr16:81812899-81991899, *DUOX1/DUOXA1* chr15:45409564-45457774, *STAT3* chr17:40,465,343-40,540,513.

4.2.4 Variant nomenclature

The following transcripts were used for variant nomenclature: *STAT3* NM_003150, *CLEC7A* NM_197947, *PLCG2* NM_002661, *DUOXA1* NM_144565, *DUOX1* NM_017434.

4.2.5 Cytokine production

PBMC were plated in 96-well plates at $5x10^5$ cells/well in 100 ml complete RPMI and stimulated for 24 hours with 100 mg/ml purified, particulate b-glucan[115] or 100 ng/ml lipopolysaccharide (LPS) (Invitrogen). Supernatants were harvested and cytokine levels measured by ELISA (R&D Systems). Normalized cytokine levels were compared using an unpaired t-test (Graphpad Prism v8.0).

4.2.6 Transfection studies and H_2O_2 measurements

Wild-type, HA-tagged DUOX1 and V5-tagged DUOXA1 expression constructs[104] were used for site-directed mutagenesis to create patient-identified variants. Flp-In-293 cells stably expressing wild-type DUOXA1 or DUOX1 were transfected (Fugene HD, Promega) with HA-tagged DUOX1 or V5-tagged DUOXA1 expression constructs, respectively. Alternatively, HEK-293 cells were co-transfected with DUOX1, DUOXA1, DECTIN-1 and PLCG2 or empty vectors. 48 hours post-transfection, cells were stimulated with 1 mM ionomycin (Invitrogen) or 100 mg depleted Zymosan (Invivogen). H_2O_2 release was continuously measured (Luminoskan Ascent plate reader, Thermo Fisher) using luminol and Horseradish Peroxidase (Sigma) for 60 minutes. Unstimulated, transfected cells were harvested for immunoblot analysis.

4.2.7 Immunoblotting

Transfected cells were lysed in RIPA buffer with protease inhibitors. 20 μg total protein per sample was denatured and loaded on a 10% polyacrylamide gel. After electrophoresis, proteins were transferred to nitrocellulose membranes, blocked for 1 hour in TBS-t with 5% powdered milk then probed with specific antibodies. Antibodies used were anti-HA (Covance, 1:1000), Anti-V5 (ThermoFisher, # R960-25, 1:1000) and anti HSP-90 (Santa Cruz, #sc-13119, 1:3000). Blots were imaged using IBRIGHT FL100 imaging system (ThermoFisher).

4.2.8 Confocal microscopy

Five µm paraffin sections were incubated at 60°C for 60 minutes, dewaxed in xylene and rehydrated in ethanol. Antigen retrieval was performed in pH 7.4 .00356 M citraconic acid solution in a steamer for 40 minutes. Slides were blocked in a solution of 1M Tris, 0.1% Tween 20, 0.5% 40%-50% gelatin from cold water fish skin (Sigma), and 0.1% of .00356 M citraconic acid solution. Slides were stained with the following Abs: Dectin-1 (1:20, clone RH1, Biolegend, cat: 144302, lot: B218624), LAMP-1 (1:50, Abcam, ab208943, lot: GR3213103-10) followed by fluorescent- conjugated secondary antibodies (anti-rat AF647, Invitrogen #A11077, anti-rabbit AF488 Jackson Immunoresearch # 711-545-152, both 1:100) for 2 hours at room temperature. chitin and cellulose (calcofluor white reagent B, Remel, R40015, Lot: 179886), TO-PRO™-3 iodide (Invitrogen, lot: 2069619) was used for nuclear stain while Calcofluor white reagent B (Remel) was used to stain *Coccidiodes* and mount slides. Images were acquired on a Leica SP5 inverted confocal microscope equipped with HyD hybrid detectors.

4.2.9 Image analysis

Images were deconvolved using Huygens Professional (Scientific Volume Imaging) and colocalization analyzed using the coloc feature of Imaris 9.7.2 (Oxford Instruments).

4.2.10 Mouse Coccidiodes infection studies

Groups of 10 mice were infected intranasally under ketamine (80 mg/kg)-xylazine (8 mg/kg) anesthesia with 56 spores of *C. posadasii strain 1038* administered in 30 µl of sterile

saline dropwise into the nares with a micropipettor. Infectious doses were verified by plate culture of inoculum following infection of the mice. Mice were monitored daily for activity level and hydration and weighed weekly. Animals exhibiting dehydration, hunched posture, isolation from cage-mates, weight loss >25%, and lethargy or weakness were euthanized as needed prior to planned end of study. Studies were terminated on day 70 post-infection.

4.2.11 Bulk RNA-seq of nasal and bronchial epithelial tissues

Nasal and bronchial epithelial tissues were collected from nasal scrapes and bronchial brushing biopsy specimens of healthy donors and AD-HIES patients following the IRB protocol (NHLBI protocol number: 07-H-0142). Total RNAs were isolated by RNA purification kit (Direct-zol RNA miniprep, Zymo Cat#R2051), followed by RNA quality verification, library generation and sequencing performed by Novogene Inc. RNA-sequencing data in FASTQ format were checked for quality using the FastQC tool (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/). The sequence reads in FASTQ files were mapped to human genome reference assembly hg38, and Gencode.v21 gene and transcript annotation. The expression values in fragments per kilobase per million mapped reads (FPKM) were filtered by mean >1 among all samples, and further normalized by quantile normalization, before log2 transformation.

4.2.12 Case-Control matching

Principal component analysis of the genotype data was performed using ~10,000 ancestry-informative variants after which 4 population controls from the 1000 Genomes project (n=2504) were selected for each DCM case using the R package Optmatch (https://cran.r-project.org/web/packages/optmatch/index.html), based on PC1-PC5.

4.2.13 Statistical analysis of variant burden

A logistic regression model for the binary response of disseminated disease status was used to estimate the odds ratio (OR) of variant burden in our genes of interest, adjusting for principal components of ancestry background. Two gene association tests (*CLEC7A* and *PLCG2*) were performed. Predicted damaging variants were selected based on gnomAD_controls_AF_popmax < 0.1, CADD (v1.6) phred score ≥ 20, and with PASS filter status in the VCF. The association was first examined in the discovery matched cohort and the significant gene was tested in the validation cohort. A p-value <0.05 (two tailed) was considered significant and a gene OR greater (less) than 1 is interpreted as indicating positive (negative) association. We used Bonferroni adjustment to correct for multiple testing. Association tests were then performed for recurring *CLEC7A* or *PLCG2* variants.

4.2.14 Statistics

Statistical analyses were performed using GraphPad Prism 8.0 software. Details of individual tests are included in the figure legends. Data was analysed by unpaired *t*-tests. In

cases where multiple data sets were analysed, one- or two-way ANOVA was used with Dunnett's correction as stated in figure legends. In all cases, *P* values <0.05 were considered significant.

4.3 Results

4.3.1 Description of cohort

We enrolled 67 DCM patients, predominantly from Arizona, in an exploratory cohort. Patients with histories of other invasive fungal infections, HIV infection or receiving immunosuppression were excluded. Median lymphocyte count was 1.8 x 10^3/μl (range 1.28-4.3), median monocyte count was 0.53 x 10^3/μl (range 0.3-0.74). Consistent with previous reports[114,151], most were males (n=44, 65.7%). Median age at enrollment was 41 years (range 10 months – 85 years). The cohort included individuals with the following genetically-determined ancestries: 20 European, 20 admixed American/Latino, 18 African/African-American, 4 East Asian, 3 South Asian and 2 indeterminate. Since *Coccidioides* exhibits tissue tropisms, we classified dissemination sites as bone (n=13), central nervous system (CNS) (n=28) or soft tissue (n=17). There were 9 patients with dissemination to two different tissues (6 bone and soft tissue, 1 bone and CNS, 2 CNS and soft tissue). Forty patients reported other previous infections including 15 with non-invasive fungal infections (11 dermatophytoses); 27 with viral infections (7 shingles); and 3 with bacterial infections, (nontuberculous mycobacteria,

Staphylococcus and *Salmonella*) (Table 4.1). Twelve patients reported recurrent upper respiratory infections while 7 reported asthma. There were 27 patients without previous invasive infections (Table 4.2).

Table 4.1. Demographics, genetic variants, and infections of disseminated coccidioidomycosis patients with other infections.

Pt	Gender	Ancestry1	Site	Variants2	Fungal Infection	Bacterial Infection	Viral Infection	Upper Resp5	Asthma
22	F	AFR	Bone	DUOXA1 c.1369G>T p.D457Y	-	-	varicella	-	-
25	M	AMR	soft tissue	DUOX1 c.3656G>A p.R1219Q	-	-	varicella	-	-
27	F	AFR	bone	-	-	-	-	otitis	-
28	F	AMR	CNS	-	-	mycobacteria	-	otitis, sinusitis	-
29	M	AFR	CNS	CLEC7A c.668T>G p.I223S; DUOXA1 c.56C>T p.P19L	Dermatophyte3	-	HPV, varicella, shingles	-	asthma
30	M	AFR	bone / soft tissue	-	-	-	HCV*	-	-
31	M	AMR	soft tissue	-	MCC4	-	EBV	-	-
32	M	SAS	soft tissue	-	-	-	varicella, shingles	-	-
33	M	ND	CNS	PLCG2 c.802C>T p.R268W	-	-	varicella	sinusitis	-

36	M	AFR	bone	PLCG2 c.1712A>G p.N571S; DUOX1 c.1853G>A p.R618Q; DUOXA1 c.56C>T p.P19L	-	-	varicella	-	-
37	M	AFR	bone / soft tissue	DUOXA1 c.937A>G p.S313G	-	-	varicella	-	-
38	M	AMR	CNS	-	-	-	EBV	-	-
39	F	EUR	CNS	CLEC7A c.714T>G p.Y238*	-	-	varicella	sinusitis	asthma
41	F	AFR	Bone / Soft Tissue	DUOX1 c.133G>T p.G45C; DUOX1 c.2911C>T p.R971C	-	-	-	sinusitis	-
45	F	AMR	soft tissue	CLEC7A c.714T>G p.Y238*	MCC	-	HSV	otitis	-
47	M	EAS	CNS	-	-	-	HSV	-	-
49	M	EUR	CNS	-	dermatophyte	-	varicella, shingles	-	-
50	M	EUR	CNS	CLEC7A c.714T>G p.Y238*; PLCG2 c.1712A>G p.N571S	-	recurrent staph skin infections	-	-	-
51	F	AMR	Bone	-	MCC dermatophyte	-	varicella, shingles	sinusitis, otitis	-

52	M	ND	bone / soft tissue	CLEC7A c.714T>G p.Y238*; PLCG2 c.1712A>G p.N571S			varicella	otitis	asthma
56	M	EUR	bone				varicella, shingles		
57	M	AMR	CNS	CLEC7A c.714T>G p.Y238*; DUOX1 c.1778T>C p.F593S			varicella		asthma
59	F	AMR	CNS		dermatophyte				
61	M	AMR	soft tissue	PLCG2 c.950C>T p.P317L; PLCG2 c.2161G>A p.E721K				otitis	
63	F	EUR	CNS / soft tissue	PLCG2 c.802C>T p.R268W	MCC		varicella		asthma
66	M	AFR	CNS	DUOXA1 c.638C>T p.T213M	dermatophyte		HSV		
69	F	EUR	soft tissue	CLEC7A c.714T>G p.Y238*; PLCG2 c.802C>T p.R268W	dermatophyte		varicella	sinusitis	
70	M	EUR	CNS	PLCG2 c.802C>T p.R268W			varicella		
71	F	AMR	soft tissue	CLEC7A c.714T>G p.Y238*			varicella	otitis	

72	M	AMR	bone	PLCG2 c.1712A>G p.N571S	-	-	varicella	-	-
73	M	AFR	CNS	PLCG2 c.802C>T p.R268W	Tinea corporis	-	-	-	-
76	M	AFR	CNS	DUOXA1 c.397C>T p.R133C	dermatophyte	-	-	-	-
80	F	EUR	CNS / soft tissue	CLEC7A c.714T>G p.Y238*; PLCG2 c.802C>T p.R268W; DUOXA1 c.647T>G p.F216C;	-	-	HPV, varicella, shingles	-	-
81	M	AFR	bone / soft tissue	-	dermatophyte	-	-	-	-
82	M	AMR	soft tissue	-	dermatophyte	-	HPV	-	-
84	M	AFR	CNS	CLEC7A c.714T>G p.Y238* DUOX1 c.3415G>A p.V1139I	dermatophyte	-	-	-	asthma
86	M	EUR	Bone	PLCG2 c.802C>T p.R268W	-	Salmonella	varicella	-	-
90	M	AMR	soft tissue	-	-	-	-	otitis	-
91	M	EUR	CNS	CLEC7A c.714T>G p.Y238*; PLCG2 c.802C>T p.R268W;	-	-	shingles	-	-

				PLCG2 c.1712A>G p.N571S					
94	M	EUR	Bone	PLCG2 c.82A>T p.M28L	MCC	-	-	-	-

[1]Genetically determined ancestry. AMR – Admixed American, AFR – African American, EAS – East Asian, EUR – European, SAS – South Asian, ND – not determined., [2]Homozygous variants are listed in **bold** type, [3]Fungal infection of skin, nails or hair, [4]Mucocutaneous candidiasis, [5]Upper respiratory infections

Table 4.2. Demographics and genetic variants of disseminated coccidioidomycosis patients without additional infections.

Pt	Gender	Ancestry[1]	Site	Variants	Asthma
23	M	EUR	soft tissue		
24	F	EUR	bone / soft tissue		
42	M	EUR	bone	*DUOXA1* c.167G>A p.R56Q	
43	F	EAS	soft tissue		
44	F	EAS	soft tissue		
53	F	AMR	soft tissue		
55	M	AFR	bone		
60	M	AFR	soft tissue	*CLEC7A* c.714T>G p.Y238*	
67	M	SAS	soft tissue	*CLEC7A* c.714T>G p.Y238*; *DUOX1* c.3656G>A p.R1219Q	
83	M	SAS	soft tissue		
87	M	AFR	Bone		
88	M	AFR	bone		
89	F	EUR	bone		
92	M	AMR	soft tissue		
35	M	AFR	bone / CNS		
40	F	EUR	CNS		
46	M	AMR	CNS		
54	F	EUR	CNS	*STAT3* c.1267C>T p.R423*	
58	M	AMR	CNS		asthma
65	M	AFR	CNS	*DUOXA1* c.56C>T p.P19L	asthma
68	M	AMR	CNS		
74	F	EAS	CNS		
75	F	EUR	CNS		
77	F	EUR	CNS	*PLCG2* c.802C>T p.R268W; *PLCG2* c.2324A>G p.K775R	
78	M	AMR	CNS		
79	M	AMR	CNS	*STAT3* c.250C>T p.R84*	
93	F	EUR	CNS	*CLEC7A* c.714T>G p.Y238*	

[1]Genetically determined ancestry. AMR – Admixed American, AFR – African American, EAS – East Asian, EUR – European, SAS – South Asian, ND – not determined.

4.3.2 Monogenic mutations

To date, only 12 DCM patients have been described carrying mutations associated with primary immunodeficiencies, all within the IL-12/IFN-γ and STAT3 axes[72,114,120,122]. Initial analysis of our patients considered rare, damaging variants with a focus on the International Union of Immunological Societies list of genes associated with human inborn errors of immunity[158] identifying two patients with novel nonsense mutations in *STAT3* (c.250C>T p.R84* and c.1267C>T p.R423*). No biallelic or dominant mutations were identified in *IL12RB1*, *IFNGR1*, *STAT1, STAT4,* or *GATA2*. While the first identified monogenic mutation leading to DCM was identified in a patient with dominant-negative STAT3 leading to autosomal-dominant Hyper-IgE (AD-HIES)[124], our two STAT3 haploinsufficient patients did not exhibit AD-HIES characteristics. Similar to the single previously reported STAT3 haploinsufficient patient[109], these two patients had unremarkable clinical histories until DCM; both cases developed fatal CNS dissemination.

4.3.3 Identification of DECTIN-1 pathway mutations

Given the paucity of identified monogenic mutations and the limited geographic presence of *Coccidioides*, we examined our DCM cohort for more common variants. We filtered our WES data using 10% minor-allele-frequency cut-off and CADD score >20 as a prediction of

deleteriousness. Hypothesizing that even common defects in fungal recognition could be important, we specifically queried *CLEC7A* c.714T>G; p.Y238*. *CLEC7A* encodes DECTIN-1, the C-type lectin PRR for the fungal cell wall component β-glucan. DECTIN-1 mediates innate fungal recognition and p.Y238* is associated with familial mucocutaneous candidiasis[39], increased susceptibility to invasive aspergillosis after hematopoietic stem cell transplantation (HSCT)[22], and chronic lung allograft dysfunction after lung transplantation[16]. We found homozygous *CLEC7A* c.714T>G; p.Y238* in 3/67 (4.5%) DCM patients compared to 680/141265 (0.48%) in the reference Genome Aggregation Database (gnomAD v2.1[76]) (P=0.0027, Fisher's Exact; Figure 4.1A). Additionally, 10 DCM patients were heterozygous for p.Y238*. Therefore, 13/67 (19.4%) DCM patients carried p.Y238*, significantly higher than that found in gnomAD 16450/141265 (11.6%) (P=0.0303, Fisher's Exact) (Figure 4.1B).

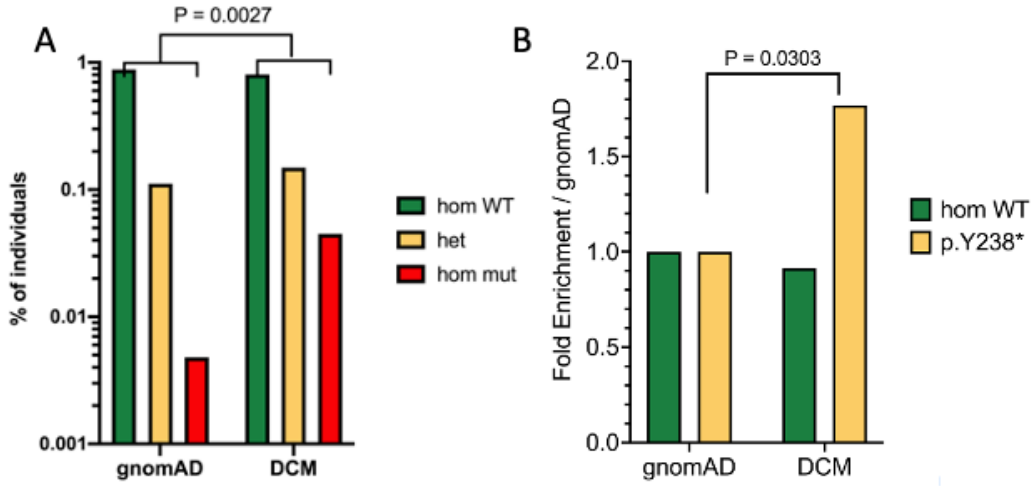


Figure 4.1. Fold enrichment of *CLEC7A*, c.714T>G; p.Y238* genotype in DCM compared to gnomAD. A. Comparison of frequency of homozygous wild-type (hom WT), heterozygous (het) or homozygous Y238* (hom mut) genotypes. B. Fold-enrichment of *CLEC7A*, c.714T>G; p.Y238* variant in DCM patients compared to gnomAD. Normalized to frequency of homozygous WT or variant carriers in gnomAD. P = 0.0303, Fisher's exact test.

p.Y238* results in loss of the terminal 10 amino acids of the C-type lectin domain (CTLD) of DECTIN-1, including the key structural residue, Cys241, which forms a disulfide bridge with Cys147[12] (Figure 4.2).

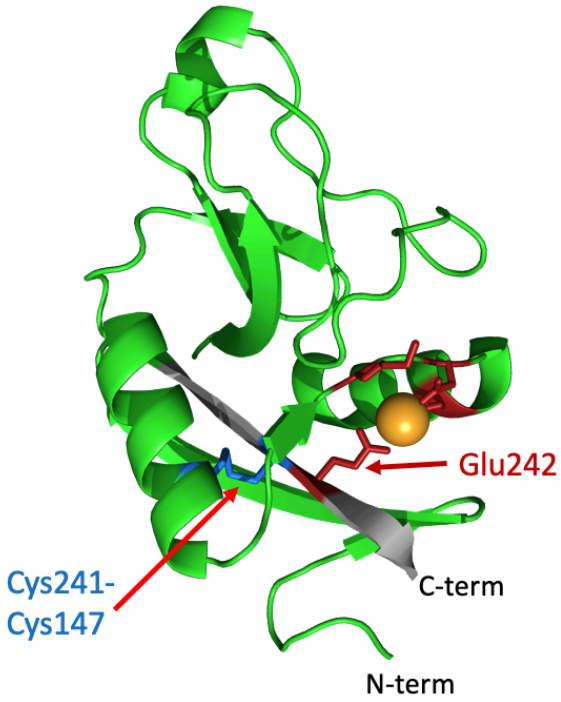


Figure 4.2. Structure of murine DECTIN-1 C-type lectin domain. Murine DECTIN-1 C-type lectin domain (PDB ID 2BPE)[19]. Amino acids deleted by Y238* are shown in grey, residues forming disulfide bonds (blue) and binding Ca2+ (red) are highlighted.

Fungal recognition by DECTIN-1 is the first step in a signaling cascade (Figure 4.3) that activates SRC kinase, leading to phosphorylation of spleen tyrosine kinase (SYK), assembly of the CARD9-BCL10-MALT1 complex and NFkB activation. Concurrently, SYK phosphorylates PLCγ2, increasing intracellular Ca^{++} and inducing non-CARD9 dependent NFAT pathways[27,49,179]. Upon engagement, DECTIN-1 is endocytosed[10].

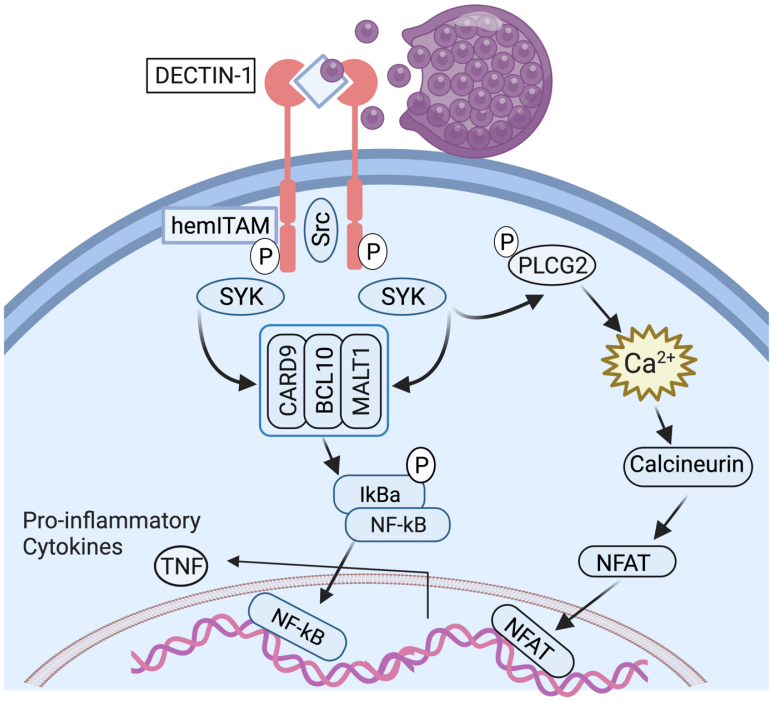


Figure 4.3. Parallel signaling pathways after β-glucan recognition by DECTIN-1 leading to activation of NFkB and NFAT transcription factors and production of TNF. β-glucan recognition by DECTIN-1 results in phosphorylation of the hemITAM domain, multimerization of DECTIN-1 receptors and activation of SYK. Activated SYK recruits CARD9-BCL10-MALT1 complex resulting in NFkB activation. SYK will also phosphorylate PLCγ2 increasing intracellular calcium leading to activation of NFAT. Figure created using BioRender.

While the 70-100 micron size of spherules prohibits engulfment, endospores recognized by DECTIN-1 are phagocytosed, localizing to DECTIN-1/LAMP-1-positive phagolysosomes (Figure 4.4).

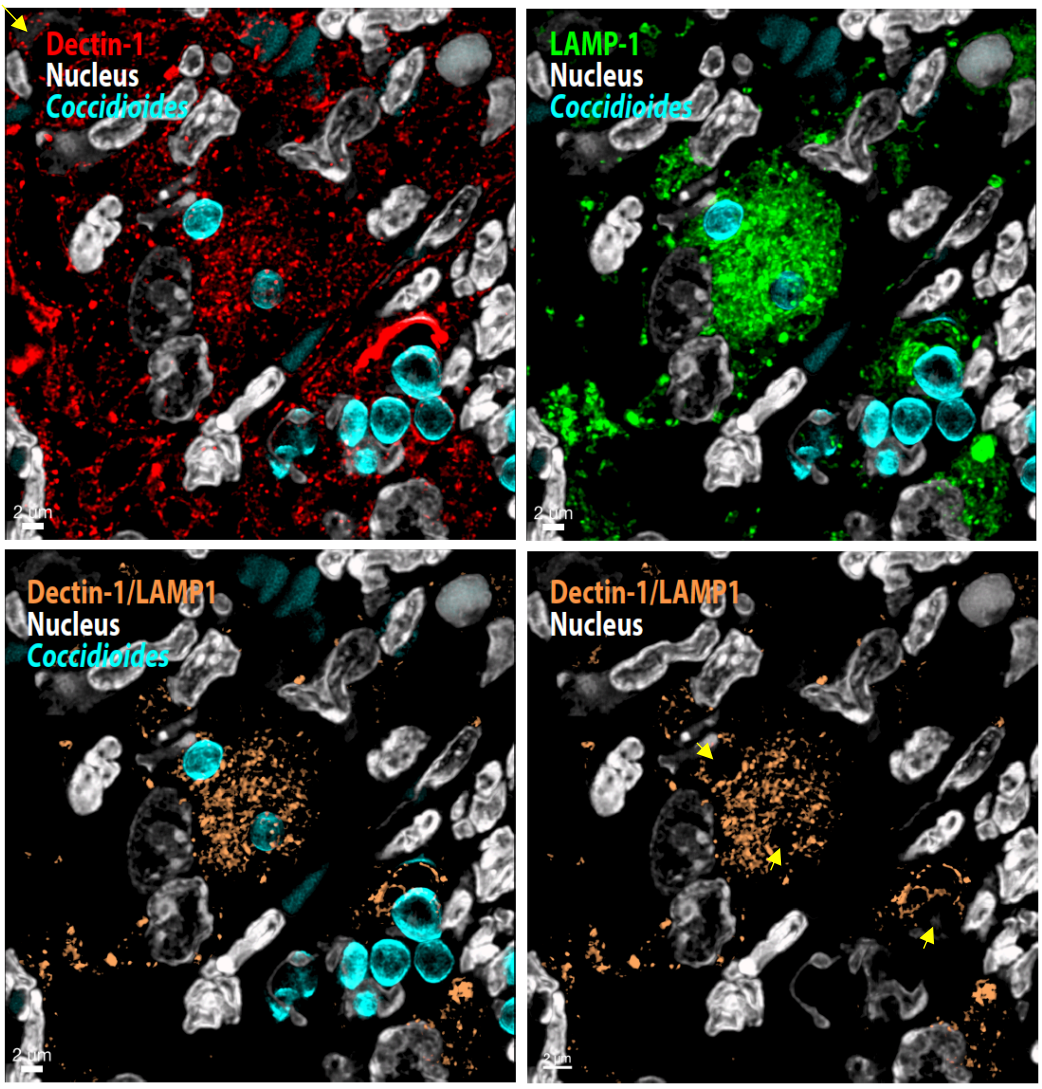


Figure 4.4. *Coccidioides* endospores found within DECTIN-1/LAMP1 positive phagolysosomes. Confocal microscopy of lung from *C. posadasii* infected C57BL/6 mouse showing DECTIN-1 (red) localized near endospores (blue) (top left) nuclei are shown in white. LAMP1 (green) localization near endospores (top right), colocalization of DECTIN-1 and LAMP1 (tan) around endospores (lower left) and DECTIN-1/LAMP1 colocalization with blue channel removed showing the colocalization surrounding the endospore (yellow arrowheads) (lower right). Image courtesy of Cynthia Aguilar, MS, NIH.

In view of the *CLEC7A* Y238* mutation abundance in our DCM cohort, we screened for additional damaging variants in this fungal recognition pathway. One patient had heterozygous DECTIN-1 p.I223S while 15 patients carried predicted damaging *PLCG2* (encoding PLCγ2) variants. *PLCG2* p.R268W was seen in 9/67 patients, including 3 with *CLEC7A* p.Y238*. Seven of the 9 individuals with *PLCG2* p.R268W were among the 20 Europeans in our cohort, significantly higher than seen in the non-Finnish European gnomAD population [7/20 (35%) vs 7917/62273 (12.66%), P=0.009, Fisher's Exact] (Figure 4.5).

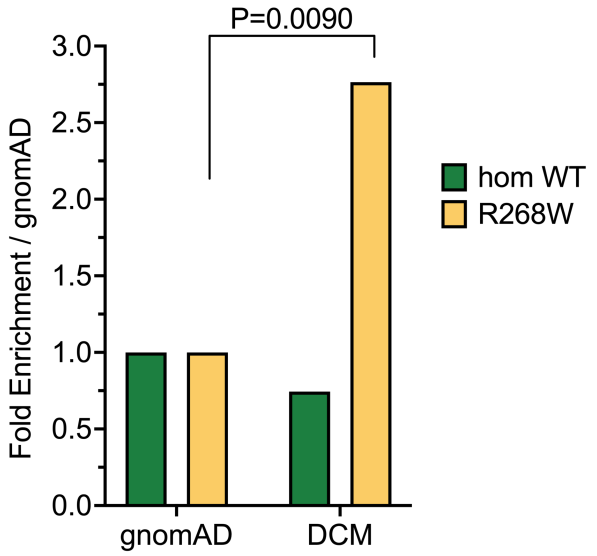


Figure 4.5. Increased PLCG2 p.R268W in DCM cohort. Frequency of EUR DCM patients with *PLCG2*, c.802C>T; p.R268W genotype normalized to the non-Finnish European population in gnomAD. P = 0.0077 Fisher's exact test.

4.3.4 Decreased TNF production in response to β-glucan

To identify relevant biologic effects of these variants, PBMCs from patients and healthy controls were stimulated with purified, particulate β-glucan, a DECTIN-1 agonist, or LPS, a

bacterial component that signals through TLR4. In response to LPS, all patients produced normal levels of TNF (Figure 4.6A). However, in response to β-glucan, cells from patients with either heterozygous or biallelic DECTIN-1 pathway variants produced significantly less TNF than controls (Figure 4.6B). DECTIN-1 p.Y238* homozygous (filled symbol) and heterozygous (open symbols) patient cells showed similar defects in TNF-α production, suggesting that DECTIN-1 p.Y238* is dominant-negative[12,33].

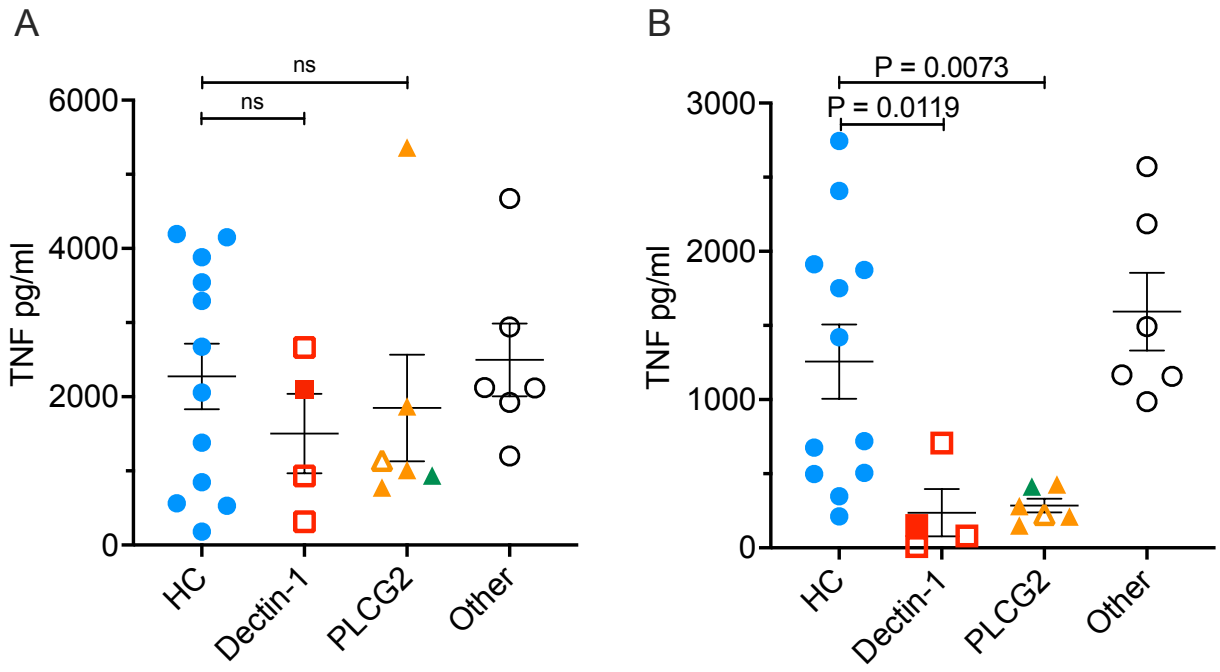


Figure 4.6. TNF production from peripheral blood mononuclear cells from patients or healthy controls. A. LPS-induced TNF production measured by ELISA. Individuals are grouped as healthy controls (n = 12), DECTIN-1 (n = 4) or PLCG2 (n = 6) variants or "Other" (n = 6) for those DCM patients without an identified causal variant. B. Particulate β-glucan induced TNF from same individuals in A. DECTIN-1 variants include homozygous p.Y238* (filled) and heterozygous p.Y238* (open). PLCG2 patients include p.R268W heterozygotes (yellow symbols), p.M28L heterozygote (green symbol), and p.R268W and p.K775R compound heterozygote (open yellow symbol). P values calculated using Brown-Forsythe and Welch ANOVA with Dunnett's T3 multiple comparisons test.

Patient cells tested, including *PLCG2* R268W and M28L, were impaired in β-glucan-induced TNF production while production of IFNγ, IL-12p70, and IL-17 did not differ between patients and healthy controls (Figure 4.7). Two control samples, recruited from the NIH blood bank in Maryland, also produced minimal TNF-α, suggesting these individuals might also be at risk if exposed.

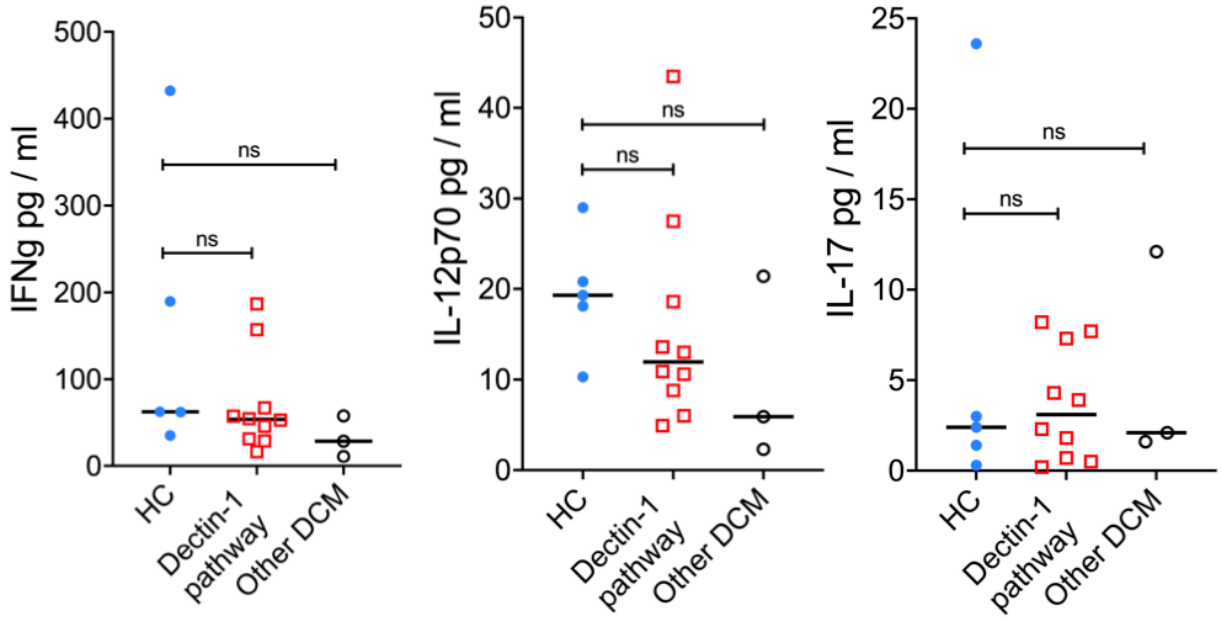


Figure 4.7. β-glucan-induced cytokine production. Individuals are grouped as healthy controls, DECTIN-1 / PLCG2 variants (Dectin-1 pathway) or "Other DCM" for those DCM patients without an identified variant. P values calculated using Brown-Forsythe and Welch ANOVA.

These data show that cells from patients carrying DECTIN-1 pathway variants produced significantly less TNF than controls ($P<0.005$, Figure 4.8) in response to β-glucan while retaining normal IFNγ, IL-12p70, and IL-17 production and normal responses to LPS.

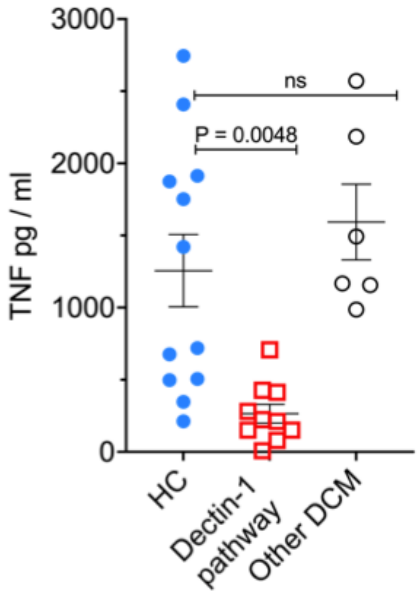


Figure 4.8. TNF production after β-glucan stimulation comparing healthy controls to patients with (Dectin-1 pathway) or without (Other DCM) Dectin-1 pathway variants.

4.3.5 Challenge of mixed population dataset

When screening for population based variants, consideration of the genetic ancestry of the individual is crucial. The identified PLCG2 p.R268W variant is primarily seen in Europeans whereas CLEC7A p.I223S is primarily seen in individuals of African descent. Since our discovery cohort consisted of patients recruited with a common disease, disseminated coccidioidomycosis, but not common ancestry, we used 10,000 genetic markers from the whole exome data to determine genetic ancestry of the patients. We compared the 10,000 markers to those from individuals in the reference genetic dataset, 1000 Genomes[160] (1000G) consisting of

2504 individuals of known ancestry. Principle component analysis can then be used to group individuals (Figure 4.9), a more accurate assessment of genetic ancestry compared to relying on self-reported race and ethnicity.

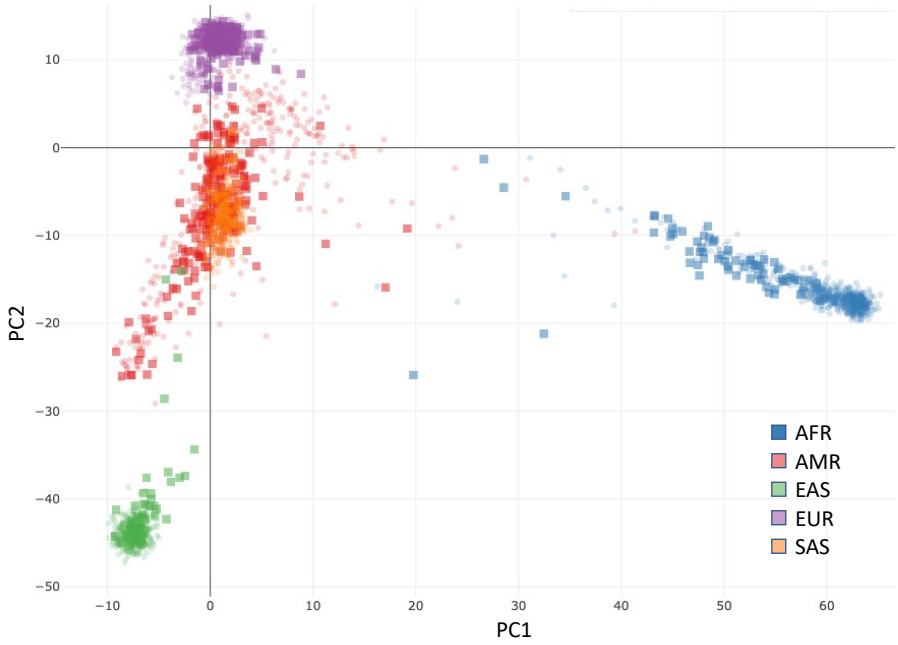


Figure 4.9 Principle component analysis of individuals from 1000G. Individuals with related genetic ancestry cluster together in multiple dimensions, shown here are PC1 vs PC2. Samples are colored by genetically determined ancestry. AFR – African / African American, AMR – Admixed American / Latino, EAS – East Asian, EUR – European, SAS – South Asian.

While Figure 4.9 shows only two dimensions, the principle component analysis encompasses 10 dimensions. Individuals who cluster closer in PCA space have more similar underlying genetics than those who are more separated.

4.3.6 Variant burdens in validation and reference cohorts

Given the mixed population of our cohort and the identification of multiple variants in the genes, we selected ancestry-matched controls from 1000 Genomes[160] (1000G) and a generalized linear model regression analysis including ancestry principal components. Using the PCA analysis shown above, each patient in the cohort was matched to the closest controls from 1000G. This decreases underlying genetic differences and provides the ability to ask whether the identified variants are over-represented in DCM patients or if the two cohorts have different ratios of individuals of various ancestries.

Variant burden across genes and for recurrent variants within the DECTIN-1 signaling pathway were evaluated using the 4 most closely matched controls from 1000G for each patient (case). DCM patients carried more damaging DECTIN-1 variants (P=0.0206, OR 3.45, 95% CI 1.21-9.84). Specifically, the Y238* variant was seen in 13 members of our cohort (P=0.0105, OR 4.05, 95% CI 1.39-11.84). Damaging variants across *PLCG2* were also over-represented (P=0.015, OR 3.05, 95% CI 1.24-7.51) including R268W (P=0.0025, OR 5.46 95% CI 1.82-16.37) and N571S (P=0.0166, OR 26.22, 95% CI 1.81-379.37). Within our cohort of 67 patients, 24 carried at least one heterozygous variant in the DECTIN-1/PLCγ2 signaling pathway (35.8%).

To validate our exploratory analysis we analyzed an independent DCM cohort consisting of patients with meningitis (n = 31; 54 yr, range 24-75 yr) or without meningitis (n = 80; 51 yr, range 21-86 yr), predominantly from California. Patient ancestry was determined using principal component analysis and 1000 Genomes[160] as reference. Previously identified

monogenic susceptibility genes, *IL12RB1, IFNGR1, IFNGR2, STAT1, STAT4,* and *STAT3* were screened and no causative mutations were found.

In the DCM validation cohort specific variants in both *CLEC7A* and *PLCG2* were again over-represented: *CLEC7A* I223S (P=0.0444, OR 3.44 95% CI 1.03-11.50); *PLCG2* R268W (P=0.0276, OR 2.49 95% CI 1.11-5.59). In the overall validation cohort CLEC7A Y238* did not reach significance; however 100% of East-Asian DCM patients carried DECTIN-1 pathway variants. These DECTIN-1 pathway variants included Y238* (12/14, 85.7%), I223S (1/14) and PLCG2 R268W (1/14). In contrast, only 4/504 (0.79%) East-Asians in 1000G (P<0.0001) and 2/9973 (0.02%) of East-Asians in gnomAD (P<0.0001) carried Y238* (Figure 4.10). Among these East-Asian individuals with DECTIN-1 Y238*, one also carried PLCG2 M28L.

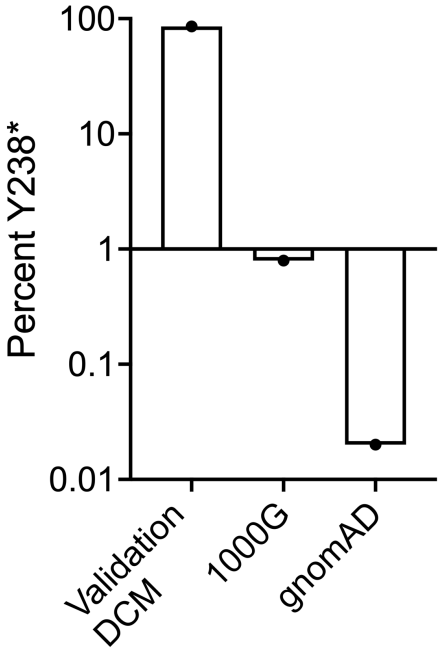


Figure 4.10. Frequency of Y238* among East-Asian patients from DCM validation cohort compared to 1000G and gnomAD.

Increased risk of dissemination among African-Americans and East-Asians has been documented extensively[134]. Multivariate analysis of coccidioidomycosis in Kern County, California, comparing patients with mild coccidioidomycosis to DCM showed an increased risk among African Americans to develop disseminated disease (OR 4.6, 95% CI 1.4 - 15)[133]. Hospitalization data from Arizona and California 2005 – 2011 documented increased incidence of hospitalization for coccidioidomycosis among African Americans (8.1 – 16.1/100,000) and Asian / Pacific Islanders (8.71 – 14.72 / 100,000) compared to other ethnicities (3.53 - 7.92 / 100,000)[82]. Within our validation cohort, presentation of disease by ancestry is consistent with the epidemiologic data, with 80% of the African-Americans in the cohort presenting with DCM. In contrast, 75-80% of those with European or Admixed-American ancestry presented with primary pulmonary disease (Figure 4.11). These disparities suggest underlying factors affecting fungal control, such as the ones we have demonstrated among the East-Asians within our cohort.

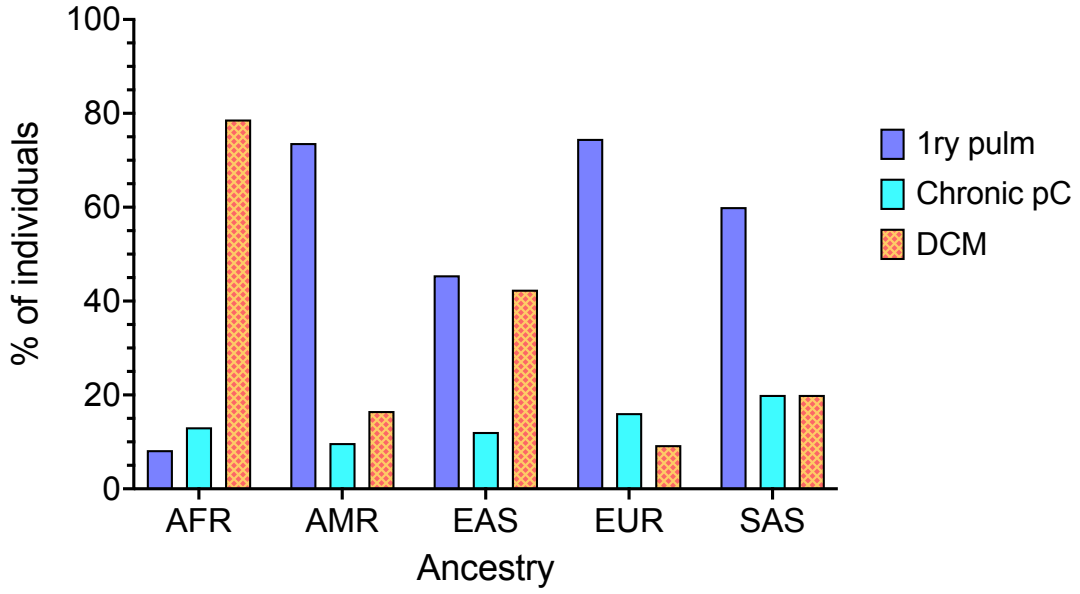


Figure 4.11. Coccidioidomycosis disease presentation by ancestry. Disease presentation of individuals from validation cohort grouped by genetically determined ancestry. Presentations include primary pulmonary (1ry pulm), chronic pulmonary (Chronic pC) and disseminated disease (DCM).

4.3.7 Non-hematopoietic fungal recognition

Heterozygous DECTIN-1 p.Y238* in either donors or recipients is associated with invasive aspergillosis after HSCT[22], demonstrating both hematopoietic and non-hematopoietic contributions of DECTIN-1 p.Y238* to antifungal defense. DECTIN-1 knockdown also impaired β-glucan-induced cytokine production in a pulmonary epithelial cell line[22]. We next considered some downstream effects of DECTIN-1 activation. Several NADPH oxidases in the NOX/DUOX family participate in innate immunity by generation of reactive oxygen species (superoxide or hydrogen peroxide, H_2O_2) after stimulation. Recognizing that DECTIN-1 engagement on

neutrophils produces superoxide via NADPH-oxidase complex[77], we examined the pulmonary epithelial NADPH-oxidase complex DUOX1/DUOXA1. Heterodimeric assembly of DUOX1 and its obligate accessory maturation factor, DUOXA1, in the endoplasmic reticulum enables its transit to the apical surface of specific epithelial cells[104]. DUOX1 releases H_2O_2 in response to calcium-mobilizing agonists due to its two intracellular Ca^{++}-sensing EF-hand domains that activate its NADPH oxidase activity. *Duox1*$^{-/-}$ mice have increased morbidity and mortality when influenza challenged[137]. While bi-allelic mutations in *DUOX2* and *DUOXA2* cause congenital hypothyroidism, heterozygous *DUOX2* variants have been implicated in inflammatory bowel disease (IBD)[50]. Similar to DUOX2/DUOXA2 in IBD, variants were found throughout DUOX1 and DUOXA1 in our exploratory cohort (Figure 4.12, yellow stars).

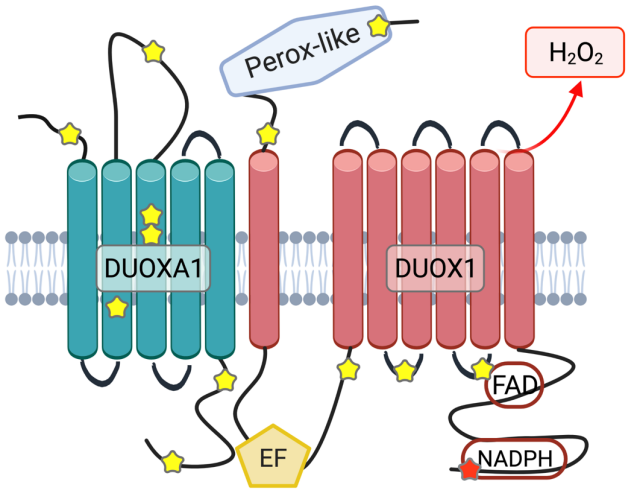


Figure 4.12. DECTIN-1 activation of DUOX1/DUOXA1. DECTIN-1 activated PLCγ2 releases intracellular Ca^{++} which activates the EF Hand domains of DUOX1 leading to H_2O_2 production. DUOX1 and DUOXA1 patient variants are highlighted by yellow stars. Figure made in BioRender.

In contrast to the recurrent DECTIN-1 and PLCG2 variants, DUOX1/DUOXA1 variants were mostly identified in single patients and were found in 10/18 (55.6%) African Americans, 3/20 (15.0%) non-Hispanic Caucasians, 1/20 (5.0%) Admixed Americans/Latinos and 1/3 (33.3%) South Asians. In the validation cohort, 10/48 African Americans (20.8%) carried damaging *DUOX1/DUOXA1* variants compared to 1/14 (7.1%) Europeans, 4/34 (11.8%) Admixed Americans/Latinos, and 2/14 (14.3%) East-Asians. For each individual variant we compared the cohort frequency to the gnomAD v2.1 population frequency corresponding to patient ancestry (Table 4.3).

Table 4.3. DUOX1/DUOXA1 variants overrepresentation by ancestry.

Gene	Variant	CADD	Ancestry	Disc DCM	Val DCM	gnomAD Pop count*	P value@	gnomAD Pop freq&	gnomAD freq$	LOF
DUOX1	p.R1219Q	23	SAS	1/3		9/15308	**0.002**	0.000294	0.0008273	y#
DUOX1	p.R1219Q	23	AMR	1/20		25/17720	**0.0289**	0.0007054	0.0008273	y
DUOX1	p.G1521R	28.6	AMR		1/34	17/17717	**0.0339**	0.0004798	0.000514	y
DUOXA1	p.R56Q	33	EUR	1/20		3/61134	**0.0013**	0.00002454	0.00002611	y
DUOX1	p.W172C	27.8	AFR		2/48	0/20737	**<0.0001**	0	0	
DUOXA1	p.F216C	23.7	EUR	1/20		1/56870	**0.0007**	0.000008792	0.000003977	
DUOX1	p.G45C	27.3	AFR	1/18		2/12474	**0.0043**	0.00008017	0.00000708	
DUOXA1	p.D457Y	21.9	AFR	1/18		6/12378	**0.0101**	0.0002424	0.00002525	
DUOX1	p.R618Q	19.8	AFR	1/18		10/12476	**0.0157**	0.0004008	0.00004247	
DUOX1	p.V1139I	17.44	AFR	1/18		29/12473	**0.0424**	0.001163	0.0001204	
DUOX1	p.H105Y	24.9	AMR		1/34	4/17268	**0.0098**	0.0001158	0.0000482	
DUOXA1	p.T213M	27.3	AFR	1/18		39/12472	*0.0561*	0.001564	0.003756	
DUOXA1	p.T213M	27.3	EUR		1/15	362/64553	*0.0811*	0.002835	0.003756	
DUOX1	p.F593S	22.4	AMR	1/20		75/17712	*0.0824*	0.002117	0.0006047	
DUOXA1	p.S313G	23.1	EAS		2/14	344/9976	*0.0827*	0.01734	0.003572	
DUOXA1	p.S313G	23.1	AFR	1/18	2/48	457/12473	0.4898	0.01859	0.003572	
DUOX1	p.R857H	24.2	AFR		1/48	48/12484	0.1717	0.001922	0.0001803	
DUOX1	p.W16C	22.2	AFR		1/48	50/12505	0.1778	0.002007	0.0001851	

DUOXA1	p.P249L	21.8	AMR		1/33	76/17713	0.1375	0.002145	0.0003113
DUOXA1	p.R133C	25.2	AFR	1/18		89/12471	0.1221	0.003568	0.0006048
DUOX1	p.R1481Q	27.6	EAS		1/14	106/9977	0.14	0.005312	0.0006753
DUOX1	p.P219R	24.6	AFR		1/48	189/8063	>0.9999	0.01178	0.001257
DUOXA1	p.P19L	22.1	AFR	3/18	4/48	954/12453	0.1557	0.03919	0.003831
DUOXA1	p.P19L	22.1	AMR		1/34	68/17641	0.1246	0.001927	0.003831

* Number of variant positive individuals / Number of individuals sequenced for given ancestry-specific population

@ Fishers exact test for population frequency between DCM cohort and ancestry-specific population in gnomAD v2.1. $P < 0.05$ denoted in **bold**, $P < 0.10$ in *italics*.

[&] Frequency of variant alleles in ancestry-specific population in gnomAD v2.1

[$] Frequency of variant alleles in gnomAD v2.1

Demonstrated loss of function in heterologous overexpression assay

In the discovery cohort 10/13 variants had a gnomAD population-specific frequency less than 2.5 per 1000 with 7/13 less than 1 per 1000. The frequency of these rare alleles was higher in both the discovery and validation cohorts compared to either gnomAD v2.1 ancestry-specific population frequency or the gnomAD v2.1 frequency for all populations (Figure 4.13).

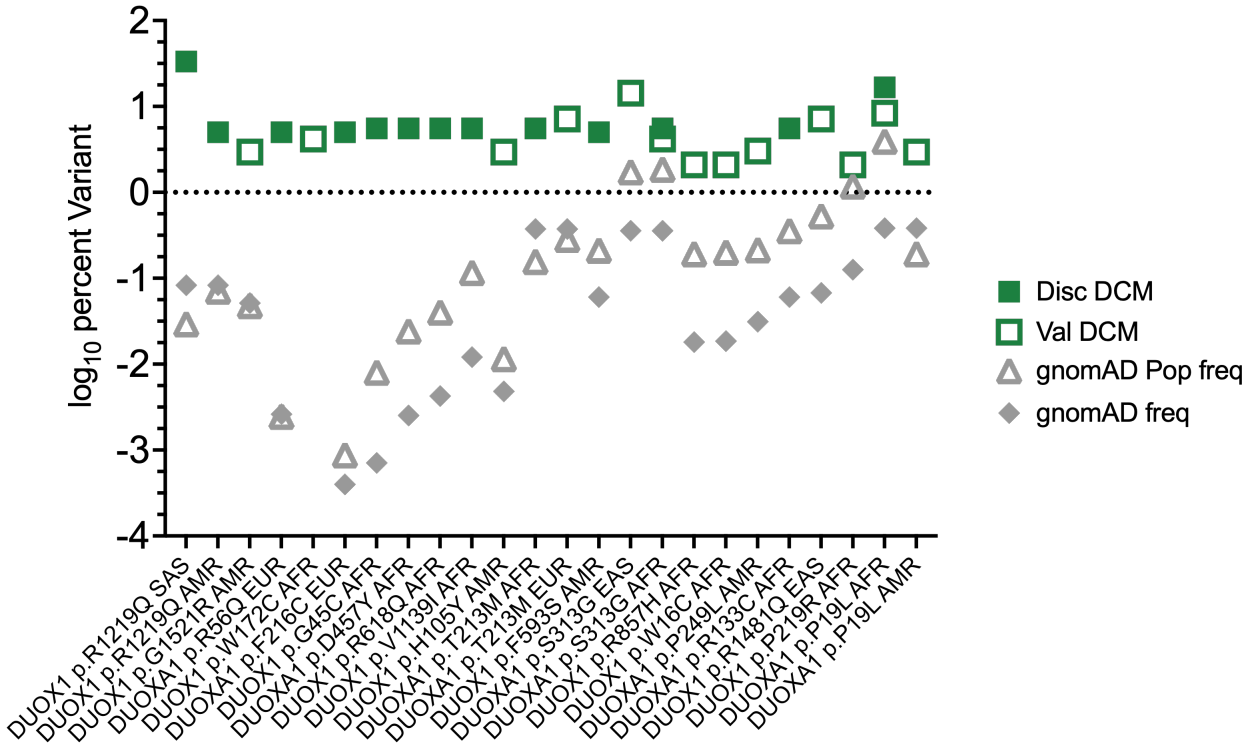


Figure 4.13. Log_{10} DUOX1/DUOXA1 variant frequency in DCM cohorts and gnomAD. Variant frequencies from discovery (Disc DCM) and validation (Val DCM) compared to gnomAD v2.1 ancestry-specific population frequency (gnomAD Pop freq) or total gnomAD v2.1 samples (gnomAD freq).

4.3.8 Evaluation of DUOX1/DUOXA1 variants

To evaluate the identified DUOX1/DUOXA1 variants, Flp-In 293 cells stably expressing wild-type DUOXA1 were transfected with HA-tagged DUOX1; three DUOX1 variants exhibited decreased protein stability and failed to produce H_2O_2 after stimulation with ionomycin despite overexpression (Figure 4.14A). Similarly, Flp-In 293 cells expressing DUOX1 were transfected with V5-tagged DUOXA1 variants. One identified variant, p.R56Q, previously reported to have compromised H_2O_2 production[88] showed significantly decreased H_2O_2 production (Figure 4.14B).

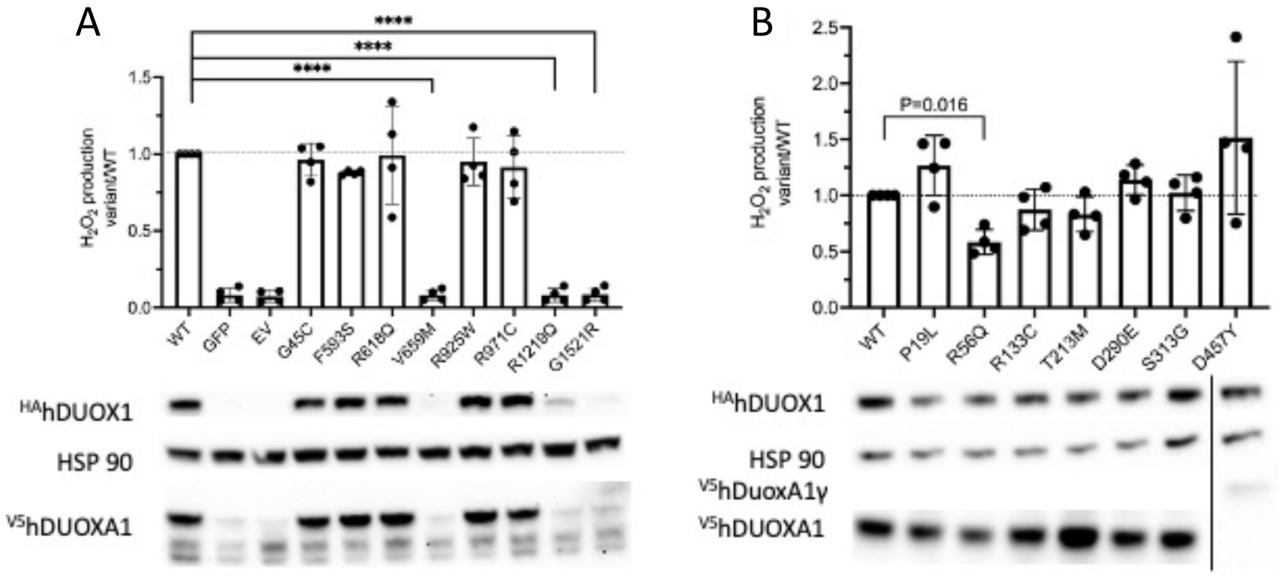


Figure 4.14. Functional assessment of identified DUOX1/DUOXA1 variants. A. DUOX1 variants identified in DCM patients were transfected into HEK Flp-In cells stably expressing wild-type DUOXA1. Cells were stimulated with ionomycin and H_2O_2 production was measured for 60 minutes. Results are the average of triplicate wells presented as the ratio of H_2O_2 production by variant/wild-type (WT), each dot represents the average of triplicate wells from a unique experiment. **** P<0.0001 using Ordinary one-way ANOVA and Dunnett's multiple comparisons test. Western blot showing decreased protein abundance of several DUOX1 variants following transfection. B. DUOXA1 variants identified in DCM patients transfected into HEK Flp-In cells stably expressing DUOX1. Measured as in A, P=0.016 using Ordinary one-way ANOVA and Dunnett's multiple comparisons test.

Given the expression of DECTIN-1 by bronchial epithelial cells[60], we hypothesized that DECTIN-1 activation by fungal components and subsequent PLCγ2-dependent intracellular Ca^{++} increase might activate DUOX1 through its Ca^{++}-sensing EF-hand domains, causing H_2O_2 production. HEK-293 cells co-transfected with WT *DUOX1*, *DUOXA1*, DECTIN-1 and *PLCG2* constructs were stimulated with ionomycin, bypassing DECTIN-1, or depleted zymosan, a β-

glucan preparation from *Saccharomyces cerevisiae* which is a DECTIN-1 agonist that does not activate Toll-like receptors, resulting in measurable H_2O_2 production (Figure 4.15).

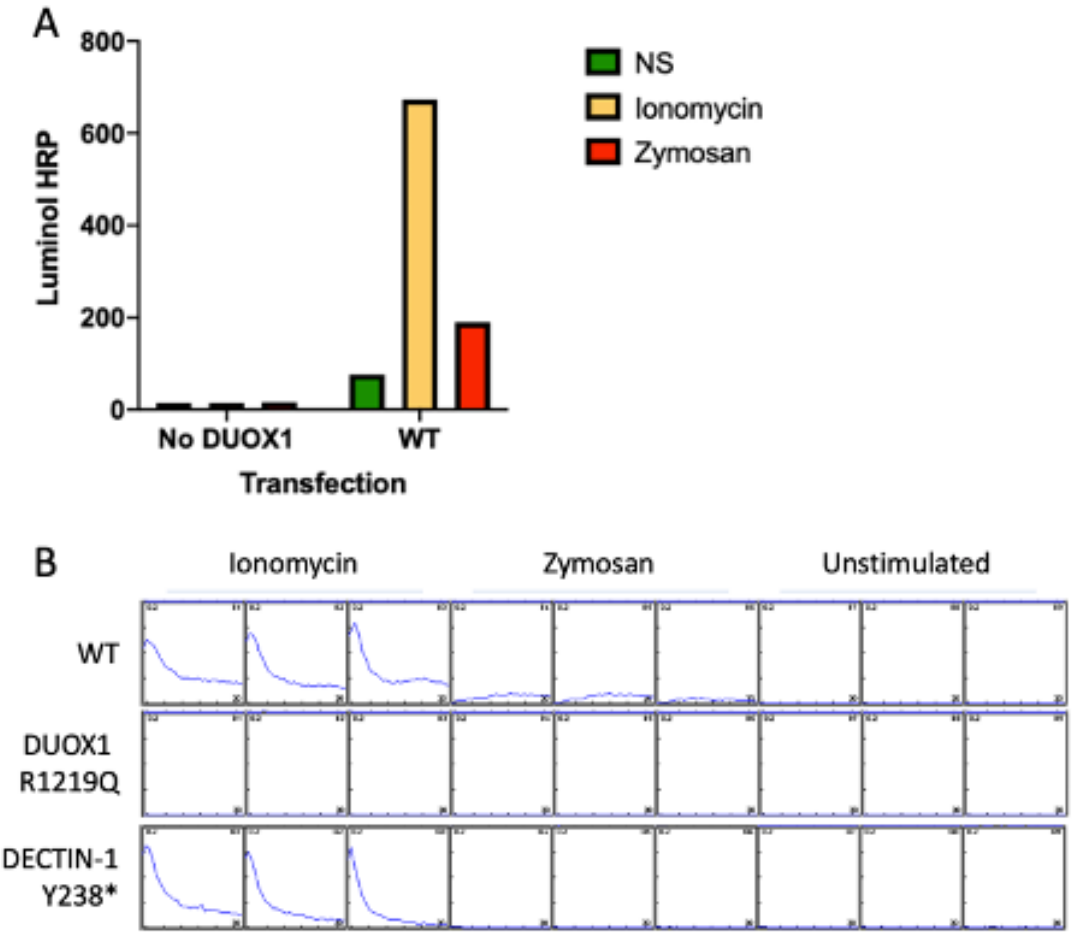


Figure 4.15. DUOX1/DUOXA1 produces H_2O_2 after DECTIN-1 engagement. HEK-293 cells were transiently transfected with DUOX1, DUOXA1, DECTIN-1 and PLCG2 expression constructs. 48 hours later cells were stimulated or not with Ionomycin (100 ng) or depleted Zymosan (100 μg). H_2O_2 production was measured for 60 minutes. A. Cumulative H_2O_2 production from transfected cells. Data is average of triplicate wells from one representative experiment. B. Kinetics of H_2O_2 production. Graphs show kinetics of H_2O_2 production in triplicate wells from one representative experiment.

Substituting WT *DUOX1*, *DUOXA1*, or DECTIN-1 with patient variants or replacing *PLCG2* with GFP abrogated zymosan-induced H_2O_2 production (Figure 4.16). *PLCG2* p.R268W supported H_2O_2 production in the overexpressed transfection system (data not shown), but primary PBMCs from patients with that variant did not upregulate TNF after β-glucan stimulation, suggesting a more complex underlying mechanism for this variant.

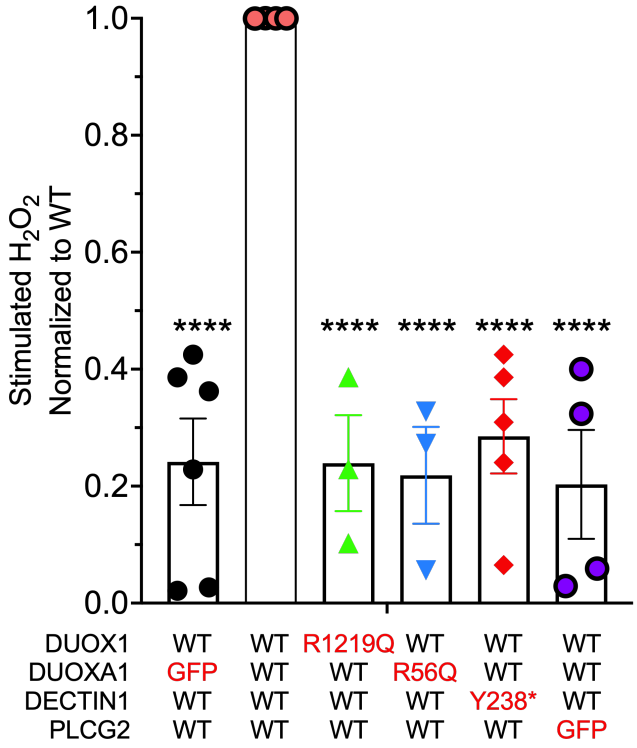


Figure 4.16. Hydrogen peroxide production in HEK cells transfected with WT or patient variant DUOX1, DUOXA1, DECTIN1 or lacking PLCG2 constructs. Results are the average of triplicate wells presented as ratio of H_2O_2 production by variant/wild-type (WT), each dot represents a unique experiment. **** P<0.0001 using Ordinary one-way ANOVA and Dunnett's multiple comparisons test.

While cells transfected with DECTIN-1 Y238* failed to produce DUOX1/DUOXA1-derived H_2O_2, individuals carrying DECTIN-1 Y238* had normal neutrophil superoxide production in

both the homozygous and heterozygous states (Figure 4.17) indicating intact ROS production by the myeloid NADPH oxidase, NOX2.

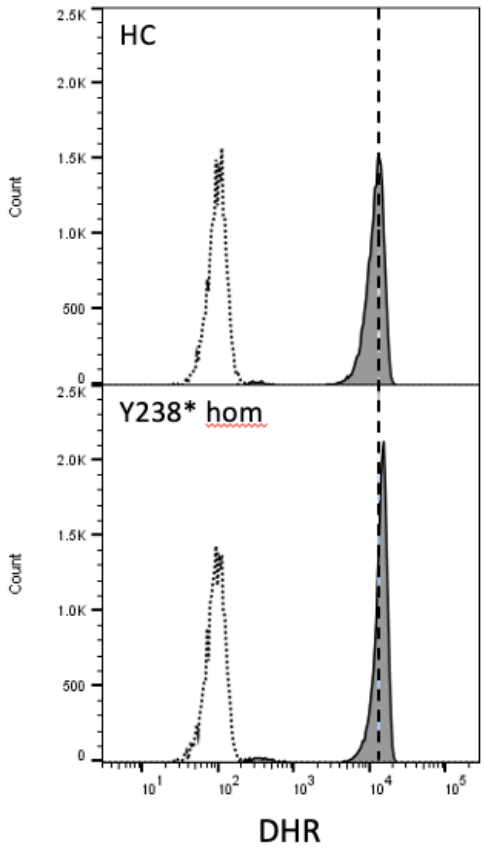


Figure 4.17. Normal neutrophil ROS in DECTIN-1 Y238* individual. Neutrophil dihydrorhodamine oxidation (DHR) from a healthy control (HC) or an individual homozygous for DECTIN-1 Y238*. Neutrophil populations were gated using forward and right-angle light scattering. Open histograms – neutrophil reactive oxygen species (ROS) production with buffer under basal conditions; solid histograms – neutrophil ROS production in response to phorbol 12-myristate 13-acetate (PMA) (400 ng/mL). The vertical dashed line represents the peak of PMA-treated neutrophils. Representative histogram from one of five DECTIN-1 Y238* patients evaluated. Data courtesy of Douglas Kuhns, PhD. NIH

4.3.9 Coccidioides infection of Duox1$^{-/-}$ mice

Given the in vitro effects observed, we infected *Duox1*$^{-/-}$ mice with *C. posadasii* strain 1038[147]. Compared to C57BL/6 controls, *Duox1*$^{-/-}$ mice had increased weight loss and mortality (Figure 4.18). Together with the in vitro functional assays, these data demonstrate that DECTIN-1 signaling is sufficient to support PLCγ2-dependent activation of DUOX1 and that DUOX1/DUOXA1 are relevant for *Coccidioides* infection.

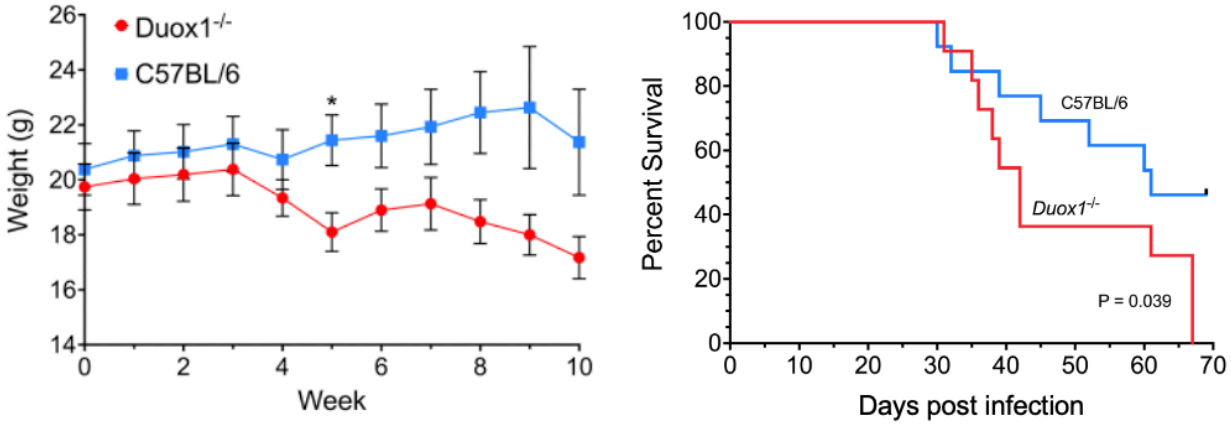


Figure 4.18. *Duox1*$^{-/-}$ mice have increased morbidity and mortality after intranasal Cp1038 infection. Weekly weights (left) of C57BL/6 and *Duox1*$^{-/-}$ mice infected with Cp1038 (median +/- SEM) and survival curves (right). * P = 0.024 using Mann-Whitney U, survival curve, P = 0.039 using Log-rank (Mantel-Cox). Median survival C57BL/6 61 days, *Duox1*$^{-/-}$ 42 days. Data courtesy of Lisa Shubitz, DVM. University of AZ.

4.3.10 Regulation of DUOX1/DUOXA1 by STAT3

It was not immediately clear whether STAT3 mediated separate pathways from DECTIN-1/PLCG2/DUOX or overlapped with them. Examination of the Encyclopedia of DNA Elements (ENCODE) (encodeproject.org) track on the UCSC genome browser (genome.ucsc.edu)

identified STAT3 binding *cis*-regulatory elements within *DUOX1*, which loop to the transcription start sites of *DUOX1* and *DUOXA1* (Figure 4.19A). In addition to the two patients in our cohort with haploinsufficient STAT3 mutations, AD-HIES patients with heterozygous, dominant-negative STAT3 mutations and reduced, but not absent, STAT3 signaling have recurrent fungal pneumonias[168]. To address whether STAT3 signaling regulates *DUOX1* and *DUOXA1*, we performed RNA sequencing on nasal or bronchial epithelial cells from AD-HIES patients. This showed decreased *DUOX1* and *DUOXA1* transcript levels compared to healthy controls (Figure 4.19B). Therefore, *DUOX1* and *DUOXA1* appear to be transcriptionally regulated by STAT3.

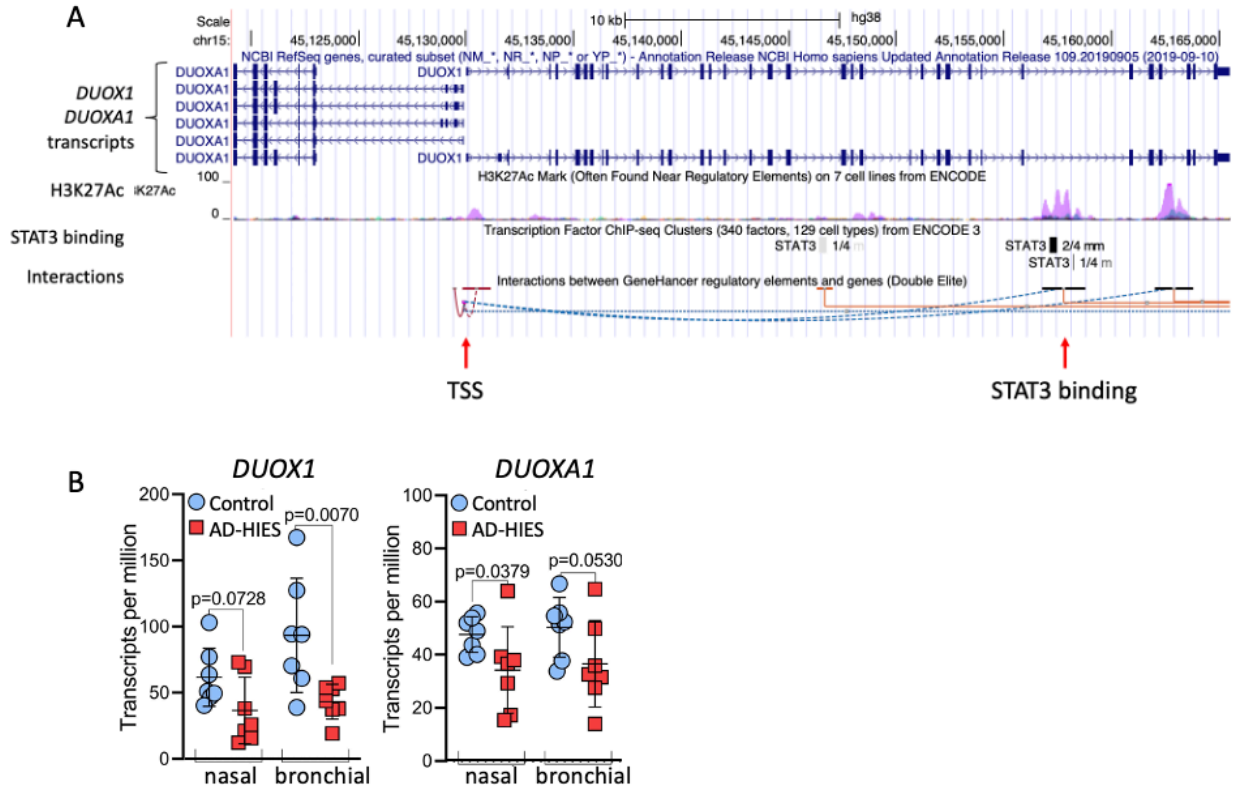


Figure 4.19. STAT3 is a transcriptional regulator of *DUOX1* and *DUOXA1*. A. *DUOXA1*, *DUOX1* locus on chromosome 15. There is minimal distance between transcription start sites (TSS) of the two genes. A cis-regulatory element located within intron 29 has H3K27-acetylation sites characteristic of active enhancers and STAT3 binding sites. This region interacts with the transcription start sites of *DUOX1* and *DUOXA1,* acting as a transcription enhancer. (Data from UCSC browser http://genome.ucsc.edu accessed 1/2022). B. Nasal and bronchial epithelial *DUOX1* and *DUOXA1* transcript levels from healthy controls or AD-HIES patients. RNA-Seq data courtesy of Gang Chen, PhD. UNC Chapel Hill. P values calculated by unpaired, 2-tailed *t*-test.

4.3.11 Comparison between disseminated and pulmonary coccidioidomycosis patients

Next we wanted to assess whether these variants were specific risk factors for dissemination or infection. Within the validation cohort, there were 59 patients with chronic pulmonary coccidioidomycosis defined as having treatment refractory (>1 yr) disease without

extra-pulmonary dissemination. Repeating ancestry matching between patients and 1000G we selected 4 controls for each patient with either DCM or chronic disease (1:4 case:control). *CLEC7A* remained significantly overrepresented (P=0.0041, OR 2.058, 95% CI 1.26 - 3.37) with p.I223S (P=0.0128, OR 3.712, 95% CI 1.32 - 10.426) and p.Y238* (P=0.0564, OR 1.735, 95% CI 0.985 - 3.058) at or near significance. At the gene level, neither *PLCG2* nor *DUOX1/DUOXA1* were significant (P = 0.596, 0.805 respectively) while at the variant level, PLCG2 p.R268W no longer reached significance (P = 0.0859). We next expanded the analysis to include individuals with primary pulmonary disease defined as those patients who sought treatment and cleared the infection with no further disease more than 1 yr after treatment. We had 65 primary pulmonary patients recruited from Arizona as well as 298 patients from the validation cohort. Including these individuals to look at infection risk rather than dissemination, we performed a 1:2 case:control analysis. *PLCG2* was significantly overrepresented at the gene level in the Discovery cohort (P = 0.0133) with p.R268W and p.N571S reaching variant level significance (P = 0.025 and 0.0024 respectively). In contrast, no gene reached significance in the validation cohort and only PLCG2 p.N571S reached significance at the variant level (P = 0.036) as under-represented. Taken together, these data suggest variants in the *CLEC7A* / *PLCG2* / *DUOX1/DUOXA1* pathway are not risk factors for primary infection but rather for control of infection (*CLEC7A*) and dissemination (*PLCG2/DUOX1/DUOXA1*).

In total, STAT3 haploinsufficient mutations, variants in the DECTIN-1 fungal recognition pathway and H_2O_2-producing pathways involving DUOX1/DUOXA1 were found in 34/67 (50.7%)

of DCM patients (Figure 4.20), spanning fungal recognition and response in both hematopoietic and non-hematopoietic compartments.

Patient	54	79	80	69	71	91	57	67	84	29	39
STAT3	R423*	R84*									
DECTIN-1			Y238* Y238*	Y238* Y238*	Y238* Y238*	Y238*	Y238*	Y238*	Y238*	I223S	Y238*
PLCG2			R268W	R268W		R268W N571S					
DUOX1							F593S	R1219Q	V1139I		
DUOXA1			F216C							P19L	

Patient	45	50	52	60	93	61	77	33	63	70	73
STAT3											
DECTIN-1	Y238*	Y238*	Y238*	Y238*	Y238*						
PLCG2		N571S	N571S			P317L E721K	R268W K775R	R268W	R268W	R268W	R268W
DUOX1											
DUOXA1											

Patient	86	94	36	72	41	25	22	37	42	66	76	65
STAT3												
DECTIN-1												
PLCG2	R268W	M28L	N571S	N571S								
DUOX1			R618Q		G45C R971C	R1219Q						
DUOXA1			P19L				D457Y	S313G	R56Q	T213M	R133C	P19L

Figure 4.20. Identified genetic variants in exploratory cohort. 34 DCM patients with identified fungal pattern recognition pathway or DUOX1 / DUOXA1 variants.

4.4 Discussion

level variants in fungal recognition and response genes affect both innate immune cells and pulmonary epithelia, the first responders to *Coccidioides*. These variants cause decreased TNF in response to fungal stimuli as well as decreased H_2O_2 production. DECTIN-1-dependent production of H_2O_2 by DUOX1/DUOXA1 is previously unrecognized.

In the <1% of *Coccidioides* infected patients who develop disseminated disease, exogenous immunosuppression is a major risk factor, as highlighted by the warnings for TNF biologics as a class. Consistent with this, *Tnf*$^{-/-}$ mice died more rapidly than wild-type mice after *Coccidioides* infection, with a median survival of 22.5 vs 70 days, and failed to form granulomata[147]. Additionally, B6D2F1 mice, which are intrinsically resistant to *Coccidioides* infection, when treated with αTNF antibodies for only the first 14 days after infection, had decreased survival and increased lung and spleen fungal burdens compared to isotype treated controls[121]. Lastly, mice with stable, controlled infection treated with αTNF antibodies began dying as soon as 2 weeks after treatment and treated mice exhibited significantly higher extrapulmonary fungal burdens[121]. These data highlight the importance of TNF production and response both early in infection and for ongoing control of *Coccidioides*. Therefore, genetically impaired PBMC TNF production following β-glucan stimulation in DCM is highly likely to be biologically relevant based on human and mouse experience.

Development of mouse models for *Coccidioides* infection has been hampered by rapid lethality, limiting the ability to discriminate individual immune responses[147]. Despite this,

increased dissemination has been demonstrated for *Clec7a*$^{-/-}$ [170] as well as the TLR downstream signaling adaptor, *Myd88*$^{-/-}$ mice[170]. Differences between susceptible C57BL/6 and resistant DBA mice have, in part, been attributed to a splice difference in *Clec7a* leading to a shortened DECTIN-1 surface receptor in C57BL/6 mice[28]. Replacement of the short form with the complete long form confers resistance in C57BL/6 mice[28]. These murine studies emphasize the importance of pattern recognition receptors in fungal response and control.

Several variants identified in our study have been previously implicated in fungal disease or immune dysregulation. Homozygous DECTIN-1 p.Y238* was reported in a family with mucocutaneous fungal infections[39] while heterozygosity in either donors or recipients was implicated in susceptibility to invasive aspergillosis after HSCT[22]. Heterozygosity for DECTIN-1 p.I223S is associated with oropharyngeal candidiasis and reduced IFN-γ production after stimulation with heat-killed *C. albicans*[119]. A patient with DCM and homozygous for p.I223S has previously been reported[81]. Further, a patient with compound heterozygous DECTIN-1 variants, p.I223S and p.Y238*, had severe, treatment-refractory infection with the far less virulent mold, *Corynespora cassiicola*[32]. *PLCG2* p.R268W was identified as a likely causal variant in a large IBD genome-wide association study[26], suggesting that this change may alter response to microbial antigens.

It is noteworthy that within our cohort, 26 of 40 patients with additional infections (Table 4.1) carried damaging variants in *CLEC7A, PLCG2* or *DUOX1/DUOXA1* compared to 6 out of 27 patients without additional serious infections (Table 4.2) (26/40 vs 6/27, P=0.001).

Further, 28 patients had viral infections, 19 of whom carried either *CLEC7A* or *PLCG2* variants, compared to 9 without variants (19/34 vs 9/33, P=0.0258, Fisher's Exact). Quintin et al.[126] demonstrated that β -glucan induced training of monocytes from a DECTIN-1 deficient patient failed to support increased cytokine release after subsequent challenge with various agonists. The recognition that β-glucan induced epigenetic changes can lead to trained immunity[103,135], implies that the dampening of host response to β-glucan by these variants may impact patient immune response beyond acute fungal infections.

While DUOX1/DUOXA1 variants have not previously been implicated in human disease, airway-epithelial DUOX1-derived H_2O_2 is sufficient to kill several oxidant-sensitive organisms including *Pseudomonas aeruginosa*, *S. aureus*, *B. cepacia* and *Hemophilus influenzae*[128]. Hydrogen peroxide is also directly inhibitory to the growth of spherules[45]. Beyond the antimicrobial effects of H_2O_2, alveolar macrophages respond to locally elevated H_2O_2 via the H_2O_2-conducting aquaglyceroporin, AQP3[71]. Highlighting the signaling role of H_2O_2, pretreatment of alveolar macrophages with catalase suppressed chemokine production[71]. Blocking transit of H_2O_2 using *Aqp3*$^{-/-}$ mice reduced cytokine or chemokine production after immune activation of keratinocytes[53] or macrophages[71]. Further, *Aqp3*$^{-/-}$ and *Duox1*$^{-/-}$ mice had decreased leukocyte recruitment in models of allergic asthma[52,71] and *Duox1*$^{-/-}$ mice have increased susceptibility to influenza[137]. More recently, Morris and colleagues[105] demonstrated a requirement for pulmonary-epithelial DUOX1 in acute house dust mite infection with significantly decreased IL-33 production 1 hour post-infection. In a chronic infection model,

macrophage-intrinsic DUOX1 was required for macrophage recruitment and activation. These studies reinforce the importance of H_2O_2 generated by DUOX1/DUOXA1 at both the pulmonary epithelial surface and in alveolar macrophages. We have demonstrated that fungal recognition through DECTIN-1 leads to production of H_2O_2. In the lung this may be pulmonary epithelia-derived but available to nearby alveolar macrophages via AQP3, macrophage intrinsic, or both, leading to enhanced signaling and release of cytokines and chemokines, all of which have been demonstrated in other systems[53,55,71,105,183]. Despite the ability of DECTIN-1 to activate NADPH oxidase activity in both DUOX1/DUOXA1 and NOX2/gp91^{phox} systems, patients with chronic granulomatous disease due to defects in the NOX2 complex do not have increased risk of infection from *Coccidioides*[62] and gp91^{phox} mice have normal resistance to intranasal *Coccidioides* infection[48,98].

Genetic variants identified by GWAS have been implicated in numerous studies in broad health risks such as heart disease, obesity and dementia. Specific variants present in the population that confer increased risk of tuberculosis (TYK2 P1104A[8]), West Nile virus (CCR5Δ32[47]), schizophrenia (loss-of-function *SETD1A* mutations[149]), or protection from HIV (CCR5Δ32[70]) have been previously reported. Similar to these studies, we demonstrate the roles of common population variants which are benign unless the carrier is infected with a geographically-isolated, pathogenic organism, *Coccidioides*.

In summary, early defects in innate fungal recognition and response are critical to the host control of the pathogen *Coccidioides;* defects in these pathways contribute to the

development of severe coccidioidomycosis, a predisposition only apparent in the corresponding locale. These mostly population-based variants define a critical set of previously unrecognized gene/environment interactions.

Chapter 5: Discussion and future directions

5.1 Discussion

In the preceding chapters, I have discussed the ecology of *Coccidioides spp*. and the economic impact of coccidioidomycosis. The search for genetic susceptibility has been underway for decades with limited success. I provide a background of known genetic susceptibilities and how to evaluate variants identified in large scale genetic sequencing as possible causative mutations. Within the cohort of disseminated coccidioidomycosis patients we collected, the search for Mendelian mutations resulted in identification of only two individuals, both of whom carried premature stop codons in STAT3 leading to haploinsufficiency. Finally, I present an alternate analysis of the genetic data, demonstrating population variants predispose individuals to severe disease.

Coccidioides spp infections are not uncommon in the endemic regions, however on a global scale they are rare with an estimated ~150,000 infections / year in the United States[46]. Considering 2/3 of infected individuals will not come to medical attention, and the majority who do will have minor or treatable symptoms, this project was approached with the hypothesis that individuals with the most severe infections, disseminated coccidioidomycosis, would carry pathogenic mutations in genes critical for antifungal innate immunity. I collected a cohort of patients with DCM and proceeded to search for rare, pathogenic mutations in known innate immune genes. When that yielded only two patients, I extended the search to genes

outside of the known innate pathways, considering all deleterious variants identified by WES using the techniques outlined in Chapter 2. Despite scrutiny of the exomes from the patients, concrete findings only emerged when a novel approach was taken, considering more common population based variants with enrichment in the affected cohort.

In Chapter 4, I demonstrated that patients with disseminated disease had significantly increased variants in *CLEC7A*, encoding DECTIN-1, and *PLCG2*[68]. This was confirmed with a validation cohort of 111 DCM patients from California[68]. In addition, I demonstrated that signaling via DECTIN-1 was sufficient to activate DUOX1/DUOXA1, the NADPH oxidase present on pulmonary epithelial cells. I proposed a novel pathway whereby recognition of Coccidioides, via DECTIN-1, is sufficient to activate DUOX1/DUOXA1 leading to release of H_2O_2. Recent work by Drummond et al.[32] strengthens the case for the importance of DECTIN-1, examining an index patient with disseminated *Corynespora cassiicola* infection and bi-allelic *CLEC7A* mutations. In addition, 11/16 phaeohyphomycosis patients carried DECTIN-1 mutations suggesting fungal recognition after invasive inoculation is crucial for cellular response and control of an otherwise benign fungal organism.

Functionally, I demonstrated patients with *CLEC7A* or *PLCG2* mutations have impaired TNF production after stimulation with purified b-glucan, a component of the *Coccidioides* cell wall[20]. TNF signaling is crucial to control of *Coccidioides* infection as evidenced by the FDA black-box warnings on TNF inhibitors. Early production of TNF is crucial for long-term control of *Coccidioides*. This was recently demonstrated by Powell et al[121] in which mice were treated with

α-TNF antibodies starting two days before infection with Cp1038. Antibody treated mice had 100% mortality compared to 100% survival in isotype treated mice[121]. Additionally, mice treated under the same regimen but receiving α-TNF for only 14 days, rather than the duration of infection, exhibited 90% mortality by 10 weeks with significantly higher lung and spleen CFUs versus isotype treated mice which had 100% survival[121]. These studies reinforce the significance of early TNF production in control of *Coccidioides* infection.

While logic would suggest a common variant should not be associated with such severe infection, I have proposed and demonstrated an alternate hypothesis. Unlike patients with bi-allelic *IFNGR1* or *IL12RB1* mutations exposed to non-pathogenic organisms and infected early in life, individuals carrying *DECTIN-1/PLCG2* or *DUOX1/DUOXA1* variants may never be exposed to *Coccidioides* if they do not enter the endemic areas. Even for individuals living in the endemic region, risk of infection is estimated to be ~3% per year[46]. If an individual is not infected until becoming an adult they may already have children meaning there is not a purifying selection against these variants even within the endemic region. Given the narrow geographic location of *Coccidioides*, it is understandable that these variants do not undergo selective pressure since the phenotype of DCM only appears in the context of mutation + exposure (Figure 5.1).

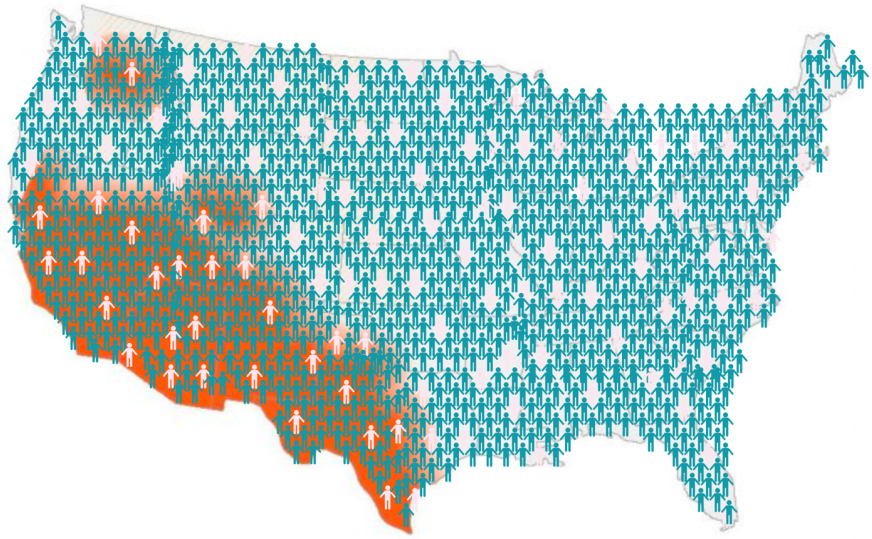


Figure 5.1. A combination of genetic susceptibility and environment is required for severe coccidioidomycosis. Identified DCM genetic variants are not uncommon in the general population (light color figures) but express only in the setting of exposure and infection within endemic regions (orange coloration on map). Figure made in BioRender.

5.2 Future work

5.2.1 Animal models of DECTIN-1 deficiency

The findings presented suggest at least two further lines of investigation. First, an exploration of DECTIN-1 variants in other mammals susceptible to DCM. C57BL/6 mice are known to be highly susceptible to Coccidioides infection with the susceptibility mapped to a splice mutation producing an altered DECTIN-1 protein[28]. Comparative genomics across different mouse strains (informatics.jax.org) identifies a single nucleotide polymorphism, rs30117984, occurring at the canonical splice acceptor for intron 2. C57BL/6 and 129X1/SvJ mouse strains carry the allele leading to exon 3 skipping resulting in the in-frame deletion of

the extracellular stalk region between the transmembrane and C-type lectin domains. In contrast, DBA/2J and A/J strains carry the alternate allele allowing proper splicing and production of a long-stalk DECTIN-1 protein and leading, in part, to their relative resistance to *Coccidioides* infections (Figure 5.2).

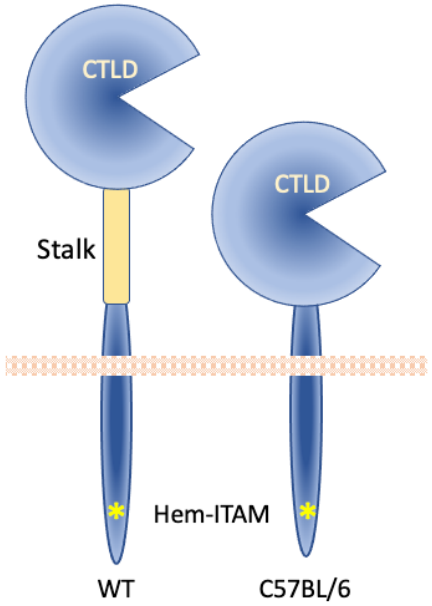


Figure 5.2. Alternative splicing of *Clec7a* across mouse strains leads to altered DECTIN-1 protein. C57BL/6 mice are homozygous for a point mutation affecting the splicing of *Clec7a* exon 3[28]. Skipping of this exon results in the in-frame deletion of the extracellular stalk region (shown in yellow) and causing the C-type lectin domain (CTLD) to be closer to the cell surface than in wild-type mice. The intracellular Hem-ITAM signaling domain is shown as a yellow asterisk.

Other mammals living in the endemic area are also susceptible to *Coccidioides* infections, including llamas[13], horses (reviewed in Brilhante[9]), cats[4], and dogs[24]. Similar to mouse strains differing in susceptibility, it has been reported that specific breeds of horses and

dogs are more or less susceptible to severe infections[14] (LF Shubitz personal communication). Given my findings in humans and the evidence of the importance of a point mutation in *Clec7a* in mouse strains, I am evaluating variants identified by whole genome sequencing of 722 Canids including wolves, dingos and domestic dogs (NCBI BioProject PRJNA448733) to determine if predicted pathogenic variants exist among the sequenced animals. After extracting the variants in the *Clec7a* locus, I used the human reference DECTIN-1 protein sequence (NP_922938.1) to query *Canis* proteins using BLAST (blast.ncbi.nlm.nih.gov). Strikingly, while the predicted dingo protein aligned closely with the human, the domestic dog, *Canis lupus familiaris*, displayed an internal deletion (Figure 5.3). Closer examination revealed deletion of the exon 3 encoded stalk region, similar to the protein seen in the C57BL/6 mouse.

A

B

Score	Expect	Method	Identities	Positives	Gaps
327 bits(839)	9e-114	Compositional matrix adjust.	183/246(74%)	201/246(81%)	0/246(0%)

```
Query  1    MEYHPDLENLDEDGYTQLHFDSQSNTRIAVVSEKGSCAASPPWRliavilgilclvilvi  60
            MEYH  +ENLDEDGYTQL+F SQ  T   VV EK +CA SP WR IAV LGILCL++LVI
Sbjct  1    MEYHSGVENLDEDGYTQLNFHSQGITGRPVVLEKVTCATSPRWRPIAVTLGILCLLMLVI  60

Query  61   avvlgTMAIWRSNSGSNTLENGYFLSRNKENHSQPTQSSLEDSVTPTKAVKTTGVLSSPC  120
            AV+LGT A WR NSGSN L+N  F SRNKENHSQPTQSSLED V PTKA+ TTG  SS C
Sbjct  61   AVILGTTAFWRFNSGSNPLKNDNFPSRNKENHSQPTQSSLEDHVAPTKALTTTGAFSSSC  120

Query  121  PPNWIIYEKSCYLFSMSLNSWDGSKRQCWQLGSNLLKIDSSNELGFIVKQVSSQPDNSFW  180
            PPNWI ++ +CYLFS SL SW+ SKR C QL SNLLKID++ EL FIV+QVSSQPDNSFW
Sbjct  121  PPNWITHKNNCYLFSTSLASWNRSKRHCSQLHSNLLKIDTAEELEFIVRQVSSQPDNSFW  180

Query  181  IGLSRPQTEVPWLWEDGSTFSSNLFQIRTTATQENPSPNCVWIHVSVIYDQLCSVPSYSI  240
            IGLSR QTE P LWEDGS FSSNLFQIR+T TQEN S NCVWIH+S+IYDQLCSVPSYSI
Sbjct  181  IGLSRHQTEGPLLWEDGSVFSSNLFQIRSTDTQENSSHNCVWIHLSIIYDQLCSVPSYSI  240

Query  241  CEKKFS  246
            CEKK S
Sbjct  241  CEKKMS  246
```

C

Score	Expect	Method	Identities	Positives	Gaps
281 bits(719)	4e-96	Compositional matrix adjust.	149/246(61%)	165/246(67%)	46/246(18%)

```
Query  1    MEYHPDLENLDEDGYTQLHFDSQSNTRIAVVSEKGSCAASPPWRLIAVILGILCLVILVI  60
            MEYH  +ENLDEDGYTQL+F SQ  T   VV EK +CA SP WR IAV LGILCL++LVI
Sbjct  1    MEYHSGVENLDEDGYTQLNFHSQGITGRPVVLEKVTCATSPRWRPIAVTLGILCLLMLVI  60

Query  61   AVVLGTMAIWRSNSGSNTLENGYFLSRNKENHSQPTQSSLEDSVTPTKAVKTTGVLSSPC  120
            AV+LGT                                          TG  SS C
Sbjct  61   AVILGT----------------------------------------------TGAFSSSC  74

Query  121  PPNWIIYEKSCYLFSMSLNSWDGSKRQCWQLGSNLLKIDSSNELGFIVKQVSSQPDNSFW  180
            PPNWI ++ +CYLFS SL SW+ SKR C QL SNLLKID++ EL FIV+QVSSQPDNSFW
Sbjct  75   PPNWITHKNNCYLFSTSLASWNRSKRHCSQLHSNLLKIDTAEELEFIVRQVSSQPDNSFW  134

Query  181  IGLSRPQTEVPWLWEDGSTFSSNLFQIRTTATQENPSPNCVWIHVSVIYDQLCSVPSYSI  240
            IGLSR QTE P LWEDGS FSSNLFQIR+T TQEN S NCVWIH+S+IYDQLCSVPSYSI
Sbjct  135  IGLSRHQTEGPLLWEDGSVFSSNLFQIRSTDTQENSSHNCVWIHLSIIYDQLCSVPSYSI  194

Query  241  CEKKFS  246
            CEKK S
Sbjct  195  CEKKMS  200
```

Figure 5.3. Domestic dogs have similar DECTIN-1 protein as C57BL/6 mice. A. Alignment of *Homo sapiens* DECTIN-1 protein against predicted DECTIN-1 protein sequence from *Canis lupus dingo* (dingos) and *Canis lupus familiaris* (domestic dog, Boxer breed). B. Amino acid alignment between *Homo sapiens* and predicted *Canis lupus dingo* DECTIN-1 protein showing high

homology. C. Amino acid alignment between *Homo sapiens* and predicted *Canis lupus familiaris* DECTIN-1 protein showing high homology but skipping the 46 amino acid stalk region (dashed line).

Since the dog proteins are bioinformatically predicted from the reference genome sequence, the actual dog DECTIN-1 protein will need to be experimentally determined. In collaboration with Lisa Shubitz, DVM at the University of Arizona I have obtained blood from 5 dogs to screen *CLEC7A* genomic and cDNA sequence as well as perform Western blots to determine protein size. Given the dog reference genome is from a Boxer and Boxers as a breed are more susceptible to severe coccidioidomycosis (LF Shubitz, personal communication), evaluating several dogs of various breeds may show differences across breeds, reinforcing the human and mouse data showing DECTIN-1 variability.

5.2.2 The role of H_2O_2 in cell-cell signaling in the lung

The second line of investigation to pursue is the role of H_2O_2 in the pulmonary epithelium and the role of peroxide signaling to alveolar macrophages or other cell types present in the lung, as well as recruited leukocytes. Identification of the DECTIN-1/PLCG2/DUOX1/DUOXA1 pathway is a novel finding from my work. Signaling via DECTIN-1 is leading to H_2O_2 production via DOUX1 yet literature on the role of H_2O_2 signaling in the pulmonary environment is limited opening the question of roles for H_2O_2 . In the past two decades the lactoperoxidase (LPO), hydrogen peroxide (H_2O_2), thiocyanate (SCN^-) system has been recognized as an important component of innate immunity in the airways[106,127]. In this

system, thiocyanate is secreted into the airway surface liquid by CFTR while H_2O_2 is produced by the dual oxidase enzyme complex, DUOX1/DUOXA1, present on the apical surface of the pulmonary epithelia[85]. *CFTR* mutations diminish the SCN^- secretion leading to decreased hypothiocyanite ($OSCN^-$) killing of *S. aureus*[106] and *Pseudomonas aeruginosa*[128]. If the primary role for H_2O_2 in the lung is the LPO/ H_2O_2 /SCN^- system, then cystic fibrosis (CF) patients would be expected to have higher rates of Coccidioides infection. Recently it was demonstrated CF patients in Southern Arizona have a lower prevalence of coccidioidomycosis than the general population[38]. Alternatively, H_2O_2 has been demonstrated to act as a signaling molecule, entering the cell through the aquaglyceroporin, Aquaporin-3 (AQP3)[7], and activating NFkB signaling[53].

Aqp3 knockout mice are available and were originally characterized in renal function[180] demonstrating failure of $Aqp3^{-/-}$ mice to concentrate urine. Subsequently, the role for AQP3 in NFkB signaling has been demonstrated in keratinocytes in a psoriasis model[53], and as a driver of macrophage NFkB signaling during CCl_4 induced liver injury[55]. Additionally, knockdown of AQP3 in a cell model led to attenuated cell growth and migration after EGF stimulation[54]. Work from Albert van der Vliet's lab has demonstrated DUOX1-mediated EGFR activation in airway epithelia[52,59]. Together, these studies provide a model for pulmonary epithelial DUOX1-produced H_2O_2 to enter nearby cells such as alveolar macrophages via AQP3 and activate or enhance NFkB signaling. The role of DUOX1 in innate immunity was highlighted by recent work demonstrating increased morbidity and mortality of influenza challenged $Duox1^{-/-}$ mice[137].

Innate chemokine production and immune cell recruitment were DUOX1 dependent[137]. While I have demonstrated increased morbidity and mortality in *Duox1*$^{-/-}$ mice challenged with Coccidioides, the role of AQP3 in *Coccidioides* infections remains unknown. Survival, dissemination, and histopathology studies of *Coccidioides* infected *Aqp3*$^{-/-}$ and *Duox1*$^{-/-}$ mice will provide a foundation for further studies.

Specific cells involved in H_2O_2 driven signaling within the complex lung environment have not been carefully identified. One approach to this question would focus on alveolar macrophages as the bridge between pulmonary epithelial H_2O_2 and immune cell signaling and recruitment. Two recent works highlight differential roles of the alveolar macrophage and type II alveolar epithelial cells (AECII). Work performed by Liu et al.[89] demonstrated IL-1α and IL-1β produced by *Legionella* infected alveolar macrophages signals to AECII driving the production of GM-CSF from AECII and amplifying the inflammatory response of monocytes. Morris and colleagues defined a macrophage-intrinsic role for DUOX1 in a mouse model of allergy to house dust mite using conditional deletion of DUOX1 in either myeloid lineages or AECII[105]. Alveolar macrophages can be obtained using bronchoalveolar lavage (BAL). RNA-Seq and cytokine analysis of BAL-macrophages at baseline, after intratracheal administration of depleted zymosan[157], and after *ex vivo* treatment with depleted zymosan can be used to compare cellular responses. With this groundwork, similar experiments using reciprocal bone marrow chimeras among C57BL/6, *Aqp3*$^{-/-}$ and *Duox1*$^{-/-}$ will enable elucidation of the role of hydrogen peroxide production and activation of hematopoietic cells. The reciprocal bone marrow

chimeric mice can also be used in survival, dissemination and histopathology studies as well to study the macro-level effects of the pulmonary hydrogen peroxide system. Similarly, the utilization of floxed-*Duox1* mice will allow ablation of DUOX1 in specific cell lineages such as alveolar macrophages, all myeloid cells or AECII.

5.3 Perspectives

While *Coccidioides* infection was identified more than a century ago and has actively been studied for decades, several basic questions remain unanswered. Why do most infected individuals remain asymptomatic without disease progression while others require medical attention? Why do a fraction of infected individual fail to clear the fungus from their lungs despite the best antifungal therapy available? And lastly, why do a fraction of individuals lose control of the infection and have extra-pulmonary dissemination? In this work I have begun to address some of the underlying population genetics across the range of infectious presentations. I have proposed a novel signaling mechanism for *Coccidioides* recognition in the lung. Lastly, based on my work, I have suggested at least two major lines of investigation proceeding from these findings including specific experiments to address these questions. This work, in the midst of the SARS-COV-2 epidemic, strives to answer some of the same questions facing the world today. Why do only a fraction of individuals infected with SARS-COV-2 become severely ill? Why do some clear the infection, regardless of severity, and experience few or no long-term sequelae? I hope that I have successfully laid a groundwork of population genetics

not only for study of disseminated coccidioidomycosis but for other infectious diseases around the globe.

Bibliography

1. Adzhubei IA, Schmidt S, Peshkin L, et al. Nat Methods. 2010;7(4):248-249.
2. Aggor FEY, Break TJ, Trevejo-Nunez G, et al. Oral epithelial IL-22/STAT3 signaling licenses IL-17 mediated immunity to oral mucosal candidiasis. Sci Immunol. 2020;5:eaba0570.
3. Alcorn JF. IL-22 plays a critical role in maintaining epithelial integrity during pulmonary infection. Front Immunol. 2020;11:1160
4. Arbona N, Butkiewicz CD, Keyes M, Shubitz LF. Clinical features of cats diagnosed with coccidioidomycosis in Arizona, 2004-2018. J Feline Med Surg. 2020;22(2):129-137.
5. Arita K, South AP, Hans-Filho G, et al. Oncostatin M receptor-beta mutations underlie familial primary localized cutaneous amyloidosis. Am J Hum Genet. 2008;82(1):73-80.
6. Beziat V, Tavernier SJ, Chen YH, et al. A recessive form of hyper-IgE syndrome by disruption of ZNF341-dependent STAT3 transcription. Sci Immunol. 2018;3(24):eaat4956.
7. Bienert GP, Moller AL, Kristiansen KA, et al. Specific aquaporins facilitate the diffusion of hydrogen peroxide across membranes. J Biol Chem. 2007;282(2):1183-92.
8. Boisson-Dupuis S, Ramirez-Alejo N, Li Z, et al. Tuberculosis and impaired IL-23-dependent IFN-g immunity in humans homozygous for a common TYK2 missense variant. *Sci Immunol*. 2018;3(30):eaau8714.

9. Brilhante RSN, Bittencourt PV, Chaves Lima RA, et al. Coccidioidomycosis and Histoplasmosis in equines: An overview to support the accurate diagnosis. J Eq Vet Sci. 2016;40:62-73.
10. Brown GD, Herre J, Williams DL, et al. Dectin-1 mediates biological effects of b-glucans. *J Exp Med*. 2003; 197(9):1119-1124.
11. Brown J, Benedict K, Park BJ, Thompson GR 3rd. Coccidioiodomycosis: epidemiology. *Clin Epidemiol*. 2013:5:185-197.
12. Brown J, O'Callaghan CA, Marshall AS, et al. Structure of the fungal b-glucan-binding immune receptor dectin-1: Implications for function. *Protein Sci*. 2007;16(6):1042-1052.
13. Butkiewicz CD, Shubitz LF. Coccidioidomycosis in alpacas in the southwestern United States. Transbound Emerg Dis. 2019;66(2):807-812.
14. Cafarchia C, Figueredo LA, Otranto D. Fungal diseases of horses. Vet Microbiol 2013;167:215–34.
15. Cagdas D, Mayr D, Baris S, et al. Genomic spectrum and phenotypic heterogeneity of human IL-21 receptor deficiency. J Clin Immunol. 2021;41:1272-1290.
16. Calabrese DR, Wang P, Chong T, et al. Dectin-1 genetic deficiency predicts chronic lung allograft dysfunction and death. *JCI Insight*. 2019;4(22):e133083.

17. Caruso R, Fina D, Peluso I, et al. IL-21 is highly produced in Helicobacter pylori-infected gastric mucosa and promotes gelatinases synthesis. J Immunol. 2007;178(9):5957–5965.
18. Chen YH, Grigelioniene G, Newton PT, et al. Absence of GP130 cytokine receptor signaling causes extended Stüve-Wiedemann syndrome. J Exp Med. 2020;217(3):e20191306.
19. Chu CS, Trapnell BC, Curristin S, et al. Genetic basis of variable exon 9 skipping in cystic fibrosis transmembrane conductance regulator mRNA. Nat Genet. 1993;3:151–156.
20. Cole GT, Hung C-Y. The parasitic cell wall of Coccidioides immitis. Med Mycol. 2001;39(Supplement 1):31-40.
21. Conti HR, Shen F, Nayyar N, et al. Th17 cells and IL-17 receptor signaling are essential for mucosal host defense against oral candidiasis. J Exp Med. 2009;206(2):299-311.
22. Cunha C, Di Ianni M, Bozza S, et al. Dectin-1 Y238X polymorphism associates with susceptibility to invasive aspergillosis in hematopoietic transplantation through impairment of both recipient- and donor-dependent mechanisms of antifungal immunity. *Blood*. 2010;116(24):5394-402.
23. Dagoneau N, Scheffer D, Huber C, et al. Null leukemia inhibitory factor receptor (LIFR) mutations in Stuve-Wiedemann/Schwartz-Jampel type 2 syndrome. Am J Hum Genet. 2004;74(2)"298-305.

24. Davidson AP, Shubitz LF, Alcott CJ, Sykes JE. Selected clinical features of coccidioidomycosis in dogs. Med Mycol. 2019;57(Supplement_1):S67-S75.
25. Davydov EV, Goode DL, Sirota M, et al. Identifying a high fraction of the human genome to be under selective contraint using GERP++. PLoS Comput Biol. 2010;6(12): e1001025.
26. de Lange KM, Moutsianas L, Lee JC, et al. Genome-wide association study implicates immune activation of multiple integrin genes in inflammatory bowel disease. *Nat Genet*. 2017;49(2):256-261.
27. Deerhake ME, Danzaki K, Inoue M, et al. Dectin-1 limits autoimmune neuroinflammation and promotes myeloid cell-astrocyte crosstalk via Card9-independent expression of Oncostatin M. *Immunity*. 2021;54:1-15
28. del Pilar Jimenez AM, Viriyakosol S, Walls L, et al. Susceptibility to Coccidioides species in C57BL/6 mice is associated with expression of a truncated splice variant of Dectin-1 (*Clec7a*). *Genes Immun*. 2008;9(4):338-48.
29. Dickinson RE, Griffin H, Bigley V, et al. Exome sequencing identifies GATA-2 mutation as the cause of dendritic cell, monocyte, B and NK lymphoid deficiency. Blood. 2011;118(10):2656–8.
30. Distler JH, Jungel A, Kowal-Bielecka O. Expression of interleukin-21 receptor in epidermis from patients with systemic sclerosis. Arthritis Rheum. 2005; 52(3):856-864.

31. Dobos RR, Benedict K, Jackson BR, McCotter OZ. Using soil survey data to model potential Coccidioides soil habitat and inform Valley fever epidemiology. PLoS One. 2021;16(2):e0247263.
32. Drummond RA, Desai JV, Hsu AP, et al. Human Dectin-1 deficiency impairs macrophage-mediated defense against phaeohyphomycosis. *J Clin Invest*. 2022;(*in press*).
33. Dulal HP, Adachi Y, Ohno N, Yamaguchi Y. β-glucan-induced cooperative oligomerization of Dectin-1 C-type lectin-like domain. *Glycobiology*. 2018;28(8):612-623.
34. Egeberg RO, Elconin AE, Egeberg MC. Effect of salinity and temperature on Coccidioides immitis and three antagonistic soil saprophytes. J Bacteriol. 1964;88(2):473-6.
35. Egeberg RO, Ely AF. Coccidioides immitis in the soil of the southern San Joaquin Valley. Am J Med Sci. 1956;231:151–154.
36. Elconin AF, Egeberg RO, Egeberg MC. Significance of soil salinity on the ecology of Coccidioides immitis. J Bacteriol. 1964;87(3):500-3.
37. Elsegeiny W, Zheng M, Eddens T, et al. Murine models of Pneumocystis infection recaptiulate human primary immune disorders. JCI Insight. 2018;3(12):e91894.
38. Estrada L, Autust J, Ojo T, Campion J. Prevalence of coccidioidomycosis in cystic fibrosis patients residing in Southern Arizona. Med Mycol. 2021;59(3):309-312.

39. Ferwerda B, Ferwerda G, Plantinga TS, et al. Human Dectin-1 deficiency and mucocutaneous fungal infections. *N Engl J Med*. 2009;361(18):1760-1767.
40. Fisher FS, Bultman MW, Johnson SM, et al. Coccidioides niches and habitat parameters in the southwestern United States: a matter of scale. Ann N Y Acad Sci. 2007;1111:47–72.
41. Flynn NM, Hoeprich PD, Kawachi MM, et al. An unusual outbreak of windborne coccidioidomycosis. N Engl J Med. 1979;301(7):358-361.
42. Freeman AF, Olivier KN. Hyper IgE syndromes and the lung. Clin Chest Med. 2016;37(3):557-567.
43. Freeman AF, Renner ED, Henderson C, et al. Lung parenchyma surgery in autosomal dominant hyper-IgE syndrome. J Clin Immunol. 2013;33(5):896-902.
44. Frey-Jakobs S, Hartberger JM, Fliegauf M, et al. ZNF341 controls STAT3 expression and thereby immunocompetence. Sci Immunol. 2018;3(24):eaat4941.
45. Galgiani JN. Inhibition of different phases of Coccidioides immitis by human neutrophils or hydrogen peroxide. *J Infect Dis*. 1986;153:217-22.
46. Galgiani JN, Ampel NM, Blair JE et al. 2016 Infectious Diseases Society of America (IDSA) Clinical Practice Guideline for the Treatment of Coccidioidomyocosis. *Clin Infec Dis*. 2016:63(6):e112-46.
47. Glass WG, McDermott DH, Lim JK, et al. CCR5 deficiency increases risk of symptomatic West Nile virus infection. *J Exp Med*. 2006;203(1):35-40.

48. Gonzalez A, Hung CY, Cole GT. Absence of phagocyte NADPH oxidase 2 leads to severe inflammatory response in lungs of mice infected with Coccidioides. *Microb Pathog*. 2011;51(6):432-41.

49. Goodridge HS, Simmons RM, Underhill DM. Dectin-1 stimulation by Candida albicans yeast or zymosan triggers NFAT activation in macrophages and dendritic cells. *J Immunol*. 2007;178: 3107–3115.

50. Grasberger H, Magis AT, Sheng E, et al. DUOX2 variants associate with preclinical disturbances in microbiota-immune homeostasis and increased inflammatory bowel disease risk. *J Clin Invest*. 2021;131(9):e141676.

51. Grizzle AJ, Wilson L, Nix DE, Galgiani JN. Clinical and economic burden of valley fever in Arizona: An incidence-based cost-of-illness analysis. Open Forum Infect Dis. 2021;8(2):ofaa623.

52. Habibovic A, Hristova M, Heppner DE, et al. DUOX1 mediates persistent epithelial EGFR activation, mucous cell metaplasia, and airway remodeling during allergic asthma. *JCI Insight*. 2016;1(18): e88811.

53. Hara-Chikuma M, Satooka H, Watanabe S, et al. Aquaporin-3-mediated hydrogen peroxide transport is required for NF-kappaB signalling in keratinocytes and development of psoriasis. Nat Commun. 2015;6:7454.

54. Hara-Chikuma M, Watanabe S, Satooka H. Involvement of aquaporin-3 in epidermal growth factor receptor signaling via hydrogen peroxide transport in cancer cells. Biochem Biophys Res Commun. 2016;471(4):603-609.
55. Hara-Chikuma M, Tanaka M, Verkman AS, Yasui M. Inhibition of aquaporin-3 in macrophages by a monoclonal antibody as a potential therapy for liver injury. *Nat Commun*. 2020;11(1):5666.
56. Harrington LE, Hatton RD, Mangan PR, et al. Interleukin 17-producing CD4+ effector T cells develop via a lineage distinct from the T helper type 1 and 2 lineages. Nat Immunol. 2005;6:1123-1132.
57. Hawn TR, Verbon A, Lettinga KD, et al. A common dominant TLR5 stop codon polymorphism abolishes flagellin signaling and is associated with susceptibility to legionnaires' disease. J Exp Med. 2003;198(10):1563-72.
58. Hawn TR, Wu H, Grossman JM, et al. A stop codon polymorphism of Toll-like receptor 5 is associated with resistance to systemic lupus erythematosus. Proc Natl Acad Sci U S A. 2005;102(30):10593-7.
59. Heppner DE, Hristova M, Dustin CM, et al. The NADPH oxidases DUOX1 and NOX2 play distinct roles in redox regulation of epidermal growth factor receptor signaling. J
60. Heyl KA, Klassert TE, Heinrich A, et al. Dectin-1 is expressed in human lung and mediates the proinflammatory immune response to nontypeable *Haemophilus influenzae*. *mBio*. 2014;5(5):e01492-14.

61. Holland SM, DeLeo FR, Elloumi HZ, et al. STAT3 mutations in the hyper-IgE syndrome. N Engl J Med. 2007;357(16):1608-19.

62. Holland SM. Chronic granulomatous disease. *Hematol Oncol Clin North Am*. 2013;27(1):89–99–viii.

63. Hsu AP, Sampaio EP, Khan J, et al. Mutations in GATA2 are associated with the autosomal dominant and sporadic monocytopenia and mycobacterial infection (MonoMAC) syndrome. Blood. 2011;118(10):2653-2655.

64. Hsu AP, Johnson KD, Falcone EL, et al. GATA2 haploinsufficiency caused by mutations in a conserved intronic element leads to MonoMAC syndrome. Blood. 2013;121(19):3830-7.

65. Hsu AP, Sowerwine KJ, Lawrence MG, et al. Intermediate phenotypes in patients with autosomal dominant hyper-IgE syndrome caused by somatic mosaicism. J Allergy Clin Immunol. 2013;131(6):1586-93.

66. Hsu AP, Zerbe CS, Foruraghi L, et al. IKBKG (NEMO) 5' untranslated splice mutations lead to severe, chronic disseminated Mycobacterial infections. Clin Infect Dis. 2018;67(3):456-459.

67. Hsu AP, Holland SM. Host genetics of innate immune system in infection. Curr Opin Immunol. 2022;74:140-149.

68. Hsu AP, Korzeniowska A, Aguilar CC, et al. Immunogenetics associated with severe coccidioidomycosis. JCI Insight. 2022; in press.

69. Huang JY, Bristow B, Shafir S, Sorvillo F. Coccidioidomycosis-associated deaths, United States, 1990-2008. *Emerg Infect Dis.* 2012;18:1723-1728.

70. Huang Y, Paxton WA, Wolinsky SM, et al. The role of a mutant CCR5 allele in HIV-1 transmission and disease progression. *Nat Med*. 1996;2(11):1240-3.

71. Ikezoe K, Oga T, Honda T, et al. Aquaporin-3 potentiates allergic airway inflammation in ovalbumin-induced murine asthma. *Sci Rep*. 2016;6:25781.

72. Imran T, Cui C. GATA2 transcription factor deficiency predisposing to severe disseminated Coccidioidomycosis. Front. Immunol. 2013. Conference Abstract: 15th International Congress of Immunology (ICI). doi: 10.3389/conf.fimmu.2013.02.00408

73. Itan Y, Zhang SY, Vogt G, et al. The human gene connectome as a map of short cuts for morbid allele discovery. Proc Natl Acad Sci U S A. 2013;114(14):5558-63.

74. Jin J, Xie X, Xiao Y, et al. Epigenetic regulation of the expression of IL12 and IL23 and autoimmune inflammation by the deubiquitinase Trabid. Nat Immunol. 2016;17(3):259-68.

75. Karczewski KJ, Weisburd B, Thomas B, et al. The ExAC browser: displaying reference data information from over 60,000 exomes. Nucleic Acids Res. 2017;45(Database issue):D840-D845.

76. Karczewski KJ, Francioli LC, Tiao G, et al. The mutational constraint spectrum quantified from variation in 141,456 humans. Nature. 2020; 581:434-443.

77. Kennedy AD, Willment JA, Dorward DW, et al. Dectin-1 promotes fungicidal activity of human neutrophils. *Eur J Immunol*. 2007;37:467-478.

78. Khader SA, Gaffen SL, Kolls JK. Th17 cells at the crossroads of innate and adaptive immunity against infectious diseases at the mucosa. Mucosal Immunol. 2009;2:403-411.

79. Kircher M, Witten DM, Jain P, et al. A general framework for estimating the relative pathogenicity of human genetic variants. Nat Genet. 2014;46(3):310-315.

80. Kollath DR, Teixeira MM, Funke A, et al. Investigating the role of animal burrows in the ecology and distribution of Coccidioides spp. In Arizona soils. Mycopathologia. 2020;185(1):145-159.

81. Krogstad P, Johnson R, Garcia-Lloret MI, et al. Host-pathogen interactions in Coccioidomycosis: Prognostic clues and opportunities for novel therapies. *Clin Ther*. 2019;41(10):1939-1954.

82. Kupferwasser D. Miller LG. Sociodemographic factors associated with patients hospitalised for coccidioidomycosis in California and Arizona, State inpatient database 2005-2011. Epidemiol Infect. 2020;149:e127.

83. Lacy GH, Swatek FE. Soil ecology of Coccidioides immitis at Amerindian middens in California. Appl Microbiol. 1974;27(2):379–388.

84. Langrish CL, Chen Y, Blumenschein WM, et al. IL-23 drives a pathogenic T cell population that induces autoimmune inflammation. J Exp Med. 2005;201:233-240.

85. Leto TL, Morand S, Hurt D, Ueyama T. Targeting and regulation of reactive oxygen species generation by Nox family NADPH oxidases. Antioxid Redox Signal. 2009 Oct; 11(10): 2607-2619.

86. Lindsay ME, Schepers D, Bolar NA, et al. Loss-of-function mutations in TGFB2 cause a syndromic presentation of thoracic aortic aneurysm. Nature Genet. 2012;44(8):922-927.

87. Lisco A, Hsu AP, Dimitrova D, et al. Treatment of relapsing HPV diseases by restored function of natural killer cells. N Engl J Med. 2021;385(10):921-929.

88. Liu S, Han W, Zang Y, et al. Identification of two missense mutations in DUOX1 (p.R1307Q) and DUOXA1 (p.R56W) that can cause congenital hypothyroidism through impairing H_2O_2 generation. *Front Endocrinol (Lausanne).* 2019;10:526.

89. Liu X, Boyer MA, Holmgren AM, Shin S. Legionella-infected macrophages engage the alveolar epithelium to metabolically reprogram myeloid cells and promote antibacterial inflammation. Cell Host Microbe. 2020;28(5):683-698.

90. Loeys BL, Chen J, Neptune ER, et al. A syndrome of altered cardiovascular, craniofacial, neurocognitive and skeletal development caused by mutations in TGFBR1 or TGFBR2. Nat Genet. 2005;37(3):275-281.

91. Loeys BL, Schwarze U, Holm T, et al. Aneurysm syndromes caused by mutations in the TGF-beta receptor. New Eng J Med. 2006;355(8):788-798.

92. Lyons JJ, Liu Y, Ma CA, et al. ERBIN deficiency links STAT3 and TGF-b pathway defects with atopy in humans. J Exp Med. 2017;214(3):669-680.

93. Ma CS, Chew GYJ, Simpson N, et al. Deficiency of Th17 cells in hyper IgE syndrome due to mutations in STAT3. J Exp Med. 2008;205(7):1551-1557.

94. Maddy KT, Coccozza J. The probable geographic distribution of *Coccidioides immitis* in Mexico. Bol Oficina Sanit Panam. 1964;57:44-54.

95. Maddy KT, Crecelius HG. Establishment of *Coccidioides immitis* in negative soil following burial of infected animals and animal tissues. In: Ajello L, ed. *Coccidioidomycosis*. Tucson, Arizona: University of Arizona Press, 1967: 309–312.

96. Malca H, Shomron N, Ast G. The U1 snRNP base pairs with the 5' splice site with a penta-snRNP complex. Mol Cell Biol. 2003;23(10):3442-3455.

97. Manel N, Unutmaz D, Littman DR. The differentiation of human TH-17 cells requires transforming growth factor-β and induction of the nuclear receptor RORγt. Nat Immunol. 2008;9(6):641-649.

98. Margolis DA, Viriyakosol S, Fierer J, Kirkland TN. The role of reactive oxygen intermediates in experimental coccidioidomycosis in mice. *BMC Microbiology*. 2011;11:71.

99. Mccurdy SA, Portillo-Silva C, Sipan CL, et al. Risk for Coccidioidomycosis among hispanic farm workers, California, USA, 2018. Emerg Infect Dis. 2020;26(7):1430-1437.

100. Miller AC, Comellas AP, Hornick DB, et al. Cystic fibrosis carriers are at increased risk for a wide range of cystic fibrosis-related conditions. Proc Natl Acad Sci U S A. 2020;117(3):1621-1627.

101. Milner JD, Brenchley JM, Laurence A, et al. Impaired Th17 cell differentiation in subjects with autosomal dominant hyper-IgE syndrome. Nature. 2008;452(10):773-777.

102. Minegishi Y, Saito M, Tsuchiya S, et al. Dominant-negative mutations in the DNA-binding domain of STAT3 cause hyper-IgE syndrome. Nature. 2007;448(7157):1058-62.

103. Moerings BGJ, de Graaff P, Furber M, et al. Continuous exposure to non-soluble b-glucans induces trained immunity in M-CSF-differentiated macrophages. *Front Immunol*. 2021;12:672796.

104. Morand S, Ueyama T, Tsujibe S, et al. Duox maturation factors form cell surface complexes with Duox affecting the specificity of reactive oxygen species generation. *FASEB J*. 2009;23(4):1205-18.

105. Morris CR, Habibovic A, Dustin CM, et al. Macrophage-intrinsic DUOX1 contributes to type 2 inflammation and mucus metaplasia during allergic airway disease. *Mucosal Immunol.* 2022;15(5):977-989.

106. Moskwa P, Lorentzen D, Excoffon KJ, et al. A novel host defense system of airways is defective in cystic fibrosis. A J Respir Crit Care Med. 2007;175(2):174-183.

107. Munoz-Hernandez B, Martinez-Rivera MA, Palma Cortes G, et al. Mycelial forms of Coccidioides spp. In the parasitic phase associated to pulmonary coccidioidomycosis with type 2 diabetes mellitus. Eur J Clin Microbiol Infect Dis. 2008;27(9):813-20.

108. Nahum A, Sharfe N, Broides A, et al. Defining the biological responses of IL-6 by the study of a novel IL-6 receptor chain immunodeficiency. J Allergy Clin Immunol. 2020;145(3):1011-1015.

109. Natarajan M, Hsu AP, Weinreich MA, et al. Aspergillosis, eosinophilic esophagitis, and allergic rhinitis in signal transducer and activator of transcription 3 haploinsufficiency. *J Allergy Clin Immunol*. 2018 Sep;142(3):993-997.e3.

110. Nelson KK and Green MR. Mammalian U2 snRNP has a sequence-specific RNA-binding activity. Genes & Dev. 1989;3:1562-1571.

111. Ng PC, Henikoff S. SIFT: predicting amino acid changes that affect protein function. Nucleic Acids Res. 2003;31:3812–3814.

112. Nieminen P, Morgan NV, Fenwick AL, et al. Inactivation of IL11 signaling causes craniosynostosis, delatyed tooth eruption, and supernumerary teeth. Am J Hum Genet. 2011;89(1):67-81.

113. Odio CD, Milligan KL, McGowan K, et al. Endemic mycoses in patients with STAT3-mutated hyper-IgE (Job) syndrome. J Allergy Clin Immunol. 2015;136(5):1411-3.e1-2.

114. Odio CD, Marciano BE, Galgiani JN, Holland SM. Risk factors for disseminated Coccidioidomycosis, United States. Emerg Infect Dis. 2017 Feb;23(2):308-311.

115. Ohno N, Uchiyama M, Tsuzuki A, et al. Solubilization of yeast cell-wall β-(1→3)-d-glucan by sodium hypochlorite oxidation and dimethyl sulfoxide extraction. *Carbohydr Res*. 1999;316:161-172.

116. Ophüls W. Further observations on a pathogenic mould formerly described as a protozoan (Coccidioides immitis, Coccidioides pyogenes), J Exp Med, 1905, vol. 6 (pg. 443-86)

117. Pappagianis D, Einstein H. Tempest from Tehachapi takes toll or Coccidioides conveyed aloft and afar. West J Med. 1978;129(6):527-530.

118. Park H, Li Z, Yang XO, et al. A distinct lineage of CD4 T cells regulates tissue inflammation by producing interleukin 17. Nat Immunol. 2005;6:1133-1141.

119. Plantinga TS, van der Velden WJ, Ferwerda B, et al. Genetic variation of innate immune genes in HIV-infected african patients with or without oropharyngeal candidiasis. *J Acquir Immune Defic Syndr*. 2010;55(10):87-94.

120. Powell DA, Shubitz LF, Hsu A, et al. STAT4 mutation in three generations with disseminated coccidioidomycosis (DCM) also exhibits increased susceptibility to coccidioidal infection in transfected mice. Open forum Infect Dis 2019 Oct; 6(Suppl 2): S77-S78.

121. Powell DA, Shubitz LF, Butkiewicz CD, et al. TNFa blockade inhibits both initial and continued control of pulmonary *Coccidioides*. *Front Cell Infect Microbiol*. 2021;11:796114.

122. Powell DA, Hsu AP, Shubitz LF, et al. Mouse model of a human STAT4 point mutation that predisposes to disseminated coccidioidomycosis. Immunohorizons. 2022;6(2):130-143.

123. Powers AE, Bender JM, Kumánovics A, et al. Coccidioides immitis meningitis in a patient with hyperimmunoglobulin E syndrome due to a novel mutation in signal transducer and activator of transcription. Pediatr Infect Dis J. 2009;28:664–666.

124. Pu J, Donovan FM, Ellingson K, et al. Clinician practice patterns that result in the diagnosis of coccidioidomycosis before or during hospitalization. *Clin Infect Dis*. 2021; 73(7):e1587-e1593.

125. Quintin J, Saeed S, Martens JHA, et al. Candida albicans infection affords protection against reinfection via functional reprogramming of monocytes. *Cell Host Microbe*. 2021;12(2):223-32.

126. Rada B, Lekstrom K, Damian S, et al. The Pseudomonas toxin pyocyanin inhibits the dual oxidase-based antimicrobial system as it imposes oxidative stress on airway epithelial cells. J Immunol. 2008; 181: 4883-4893.

127. Rada B, Leto TL. Redox warfare between airway epithelial cells and Pseudomonas: dual oxidase versus pyocyanin. *Immunol Res*. 2009;43(1-3):198-209.

128. Rentzsch P,Schubach M, Shendure J, Kircher M. CADD-Splice-improving genome-wide variant effect prediction using deep learning-derived splice scores. *Genome Med*. 2021;13(1):31.

129. Rienhoff HY, Yeo CY, Morissette R, et al. A mutation in TGFB3 associated with a syndrome of low muscle mass, growth retardation, distal arthrogryposis and clinical features overlapping with Marfan and Loeys-Dietz syndrome. Am J Med Genet. 2013; 161A(8):2040-2046.

130. Rixford E, Gilchrist TC. Two cases of protozoan (coccidioidal) infection of the skin and other organs, Johns Hopkins Hosp Rep, 1896, vol. 10 (pg. 209-68)

131. Roos D, van Leeuwen K, Hsu AP, et al. Hematologically important mutations: X-linked chronic granulomatous disease (fourth update). Blood Cells Mol Dis. 2021;92:102587.

132. Rosenstein NE, Emery KW, Werner SB, et al. Risk factors for severe pulmonary and disseminated coccidioidomycosis: Kern County, California, 1995-1996. Clin Infect Dis. 2001;32(5):708-15.

133. Ruddy BE, Mayer AP, Ko MG, et al. Coccidioidomycosis in African Americans. *Mayo Clin Proc*. 2011;86(1):63-69.

134. Saeed S, Quintin J, Kerstens HHD, et al. Epigenetic programming of monocyte-to-macrophage differentiation and trained immunity. *Science*. 2014;345(6204):1251086.

135. Sampaio EP, Hsu AP, Pechacek J, et al. Signal transducer and activator of transcription 1 (STAT1) gain-of-function mutations and disseminated coccidioidomycosis and histoplasmosis. J Allergy Clin Immunol. 2013;131:1624–34.

136. Sarr D, Gingerich AD, Asthiwi NM, et al. Dual oxidase 1 promotes antiviral innate immunity. *Proc Natl Acad Sci U S A*. 2021;118(26):e2017130118.

137. Sawyer TK. Src homology-2 domains: Structure, mechanisms, and drug discovery. Biopolymers. 1998;47(3):243-61.

138. Schreiber F, Arasteh JM, Lawley TD. Pathogen resistance mediated by IL-22 signaling at the epithelial-microbiota interface. J Mol Biol. 2015;427:3676-3682.

139. Schuetz AN, Pisapia D, Yan J, Hoda RS. An atypical morphologic presentation of Coccidioides spp. In fine-needle aspitation of lung. Diagn Cytopathol. 2012;40(2):163-7.

140. Schwerd T, Twigg SRF, Aschenbrenner D, et al. A biallelic mutation in IL6ST encoding the GP-130 co-receptor causes immunodeficiency and craniosynostosis.J Exp Med. 2017;214(9):2547-2562.

141. Schwerd T, Krause F, Twigg SRF, et al. A variant in IL6ST with a selective IL-11 signaling defect in human and mouse. Bone Res. 2020;8(1):24.

142. Seitz AE, Prevots DR, Holland SM. Hospitalizations associated with disseminated coccidioidomycosis, Arizona and California, USA. Emerg Infect Dis. 2012;18(9):1476-9.

143. Shahin T, Aschenrenner D, Cagdas D, et al. Selective loss of function variants in IL6ST cause Hyper-IgE syndroe with distinct impairments of T-cell phenotype and function. Haematologica. 2019;104(3):609-621

144. Sharpton TJ, Stajich JE, Rounsley SD et al. Comparative genomic analyses of the human fungal pathogens Coccidioides and their relatives. Genome Res. 2009;19:1722–1731.

145. Shubitz LF. Comparative aspects of coccidioidomycosis in animals and humans. Anna NY Acad Sci. 2007;1111:395-403.

146. Shubitz LF, Powell DA, Butkiewicz CD, et al. A chronic murine disease model of coccidioidomycosis using Coccidioides posadasii, strain 1038. J Infect Dis. 2021;223(1):166-173.

147. Sies H. Hydrogen peroxide as a central redox signaling molecule in physiological oxidative stress: Oxidative eustress. Redox Biol. 2017;11:613-619.

148. Singh T, Kurki MI, Curtis D, et al. Rare loss-of-function variants in SETD1A are associated with schizophrenia and developmental disorders. Nat Neurosci. 2016;19(4):571-7.

149. Smith CE, Beard RR, Whiting EG, Rosenberger HG. Varieties of coccidioidal infection in relation to the epidemiology and control of the diseases. Am J Public Health Nations Health. 1946;36(12):1394-1402.

150. Sondermeyer Cooksey GL, Nguyen A, Vugia D, Jain S. Regional Analysis of Coccidioidomycosis Incidence — California, 2000–2018. MMWR Morb Mortal Wkly Rep 2020;69:1817–1821.

151. Spencer S, Kostel Bal S, Egner W, et al. Loss of the interleukin-6 receptor causes immunodeficiency, atopy, and abnormal inflammatory responses. J Exp Med. 2019;216(9):1986-1998.

152. Stanga SD, Dajud MV. Visual changes in a 4-year-old. Clinical Pediatrics. 2008;47(9):959-961.

153. Stelzer G, Plaschkes I, Oz-Levi D, et al. VarElect: the phenotype-based variation prioritizer of the GeneCards Suite. BMC Genomics. 2016;17(Suppl 2):444

154. Swatek FE, Omieczynski DT, Plunkett OA. Coccidioides immitis in California. Paper presented at: The Second Symposium on Coccidiodomycosis 1965; Phoenix, Arizona.

155. Takeda K, Noguchi K, Shi W, et al. Targeted disruption of the mouse Stat3 gene leads to early embryonic lethality. Proc Natl Acad Sci U S A. 1997;94(8):3801-4.

156. Takeuchi K, Umeki Y, Matsumoto N, et al. Severe neutrophil-mediated lung inflammation in myelopoeroxidase-deficient mice exposed to zymosan. Inflamm Res. 2012;61:197-2012.

157. Tangye SG, Al Herz W, Bousfiha A, et al. Human inborn errors of immunity: 2019 update on the classification from the International Union of Immunological Societies expert committee. *J Clin Immunol*. 2020;40(1):24-64.

158. Taylor JW, Barker BM. The endozoan, small-mammal reservoir hypothesis and the life cycle of Coccidioides species. Med Mycol. 2019;57:S16-S20.

159. The 1000 Genomes Project Consortium. A global reference for human genetic variation. Nature. 2015;526:68–74.

160. Tsang CA, Tabnak F, Vugia DJ et al. Increase in reported coccidioidomycosis – United States, 1998-2011. MMWR Morb Mortal Wkly Rep. 2013;62(12);217-221.

161. Tsilifis C, Freeman AF, Gennery AR. STAT3 Hyper-IgE Syndrome - an update and unanswered questions. J Clin Immunol. 2021;41(5):864-880.

162. Twarog M, Thompson GR III. Coccidioidomycosis: Recent updates. *Semin Respir Crit Care Med*. 2015;36(5):746-55.

163. van de Laar IM, Oldenburg RA, Pals G, et al. Mutations in SMAD3 cause a syndromic form of aortic aneurysms and dissections with early-onset osteoarthritis. Nat. Genet. 2011;43(2):121–126.

164. Van der Auwera GA, Carneiro MO, Hartl C, et al. From FastQ data to high confidence variant calls: the Genome Analysis Toolkit best practices pipeline. *Curr Protoc Bioinformatics*. 2013;43(1110):11.10.1-11.10.33.

165. Veal E, Day A. Hydrogen peroxide as a signaling molecule. Antioxid Redox Signal. 2011;15(1):147-51.

166. Vinh DC, Masannat F, Dzioba RB, Galgiani JN, Holland SM. Refractory disseminated coccidioidomycosis and mycobacteriosis in interferon-gamma receptor 1 deficiency. Clin Infect Dis. 2009;49:e62–5.

167. Vinh DC, Sugui JA, Hsu AP, Freeman AF, Holland SM. Invasive fungal disease in autosomal-dominant Hyper-IgE syndrome. J Allergy Clin Immunol
. 2010 Jun;125(6):1389-90.

168. Vinh DC, Schwartz B, Hsu AP, et al. Interleukin-12 receptor β1 deficiency predisposing to disseminated coccidioidomycosis. Clin Infect Dis. 2011;52:e99– 102.

169. Viriyakosol S, Jimenez M del P, Gurney MA, et al. Dectin-1 is required for resistance to coccidioidomycosis in mice. *mBio*. 2013;4(1):e00597-12.

170. Volpe E, Servant N, Zollinger R, et al. A critical function for transforming growth factor-β, interleukin 23 and proinflammatory cytokines in driving and modulating human Th-17 responses. Nat Immunol. 2008;9(6):650-657.

171. Wahl MC, Will CL, Luhrmann R. The spliceosome: Design principles of a dynamic RNP machine. Cell. 2009;136(4):701-718.

172. West TE, Chantratita N, Chierakul W, et al. Impaired TLR5 functionality is associated with survival in melioidosis. J Immunol. 2013;190(7):3373-9.

173. Will CL, Luhrmann R. Spliceosome structure and function. Cold Spring Harb Perspect Biol. 2011;3(7):a003707.

174. Willett FM, Weiss A. Coccidioidomycosis in Southern California: report of a new endemic area with a review of 100 cases. Ann Intern Med. 1945;23:349-375.

175. Williams PL, Sable DL, Mendez P, Smyth LT. Symptomatic coccidioidomycosis following a severe natural dust storm: an outbreak at the Naval Air Station, Lemoore, Calif. Chest. 1979;76(5):566-570.

176. Willis BC, Borok Z. TGF-beta-induced EMT: mechanisms and implications for fibrotic lung disease. Am J Physiol Lung Cell Mol Physiol. 2007;293(3):L525-34.

177. Wilson L, Ting J, Lin H, et al. The rise of valley fever: Prevalence and cost burden of coccidioidomycosis infection in California. *Int J Environ Res Public Health*. 2019 ;16(7):1113.

178. Xu S, Huo J, Lee KG, et al. Phospholipase Cg2 is critical for Dectin-1-mediated Ca2+ flux and cytokine production in dendritic cells. *J Biol Chem*. 2009;284(11):7038-7046.

179. Yang B, Ma T, Verkman AS. Erythrocyte water permeability and renal function in double knockout mice lacking aquaporin-1 and aquaporin-3. J Biol Chem. 2001;276(1):624-8.

180. Yeste A, Mascanfroni ID, Nadeau M, et al. IL-21 induces IL-22 production in CD4+ T cells. Nat Comm. 2014. 5:3753

181. Zepeda MR, Kobayashi GK, Appleman MD, Navarro A. Coccidioides immitis presenting as a hyphal form in cerebrospinal fluid. J Natl Med Assoc. 1998;90(7):435-6.

182. Zhu N, Feng X, He C, et al. Defective macrophage function in aquaporin-3 deficiency. *FASEB J*. 2011;25(12):4233-9.