Co-evolved Partners of Immunity: A Trait-Based Map of Human Keystone Organisms

Amir Asiaee[1], Natalie Mallal[2], Elizabeth Phillips[3,4], Simon Mallal[3,4*]

[1]Department of Biostatistics, Vanderbilt University Medical Center, TN, USA.
[2]College of Arts and Science, Vanderbilt University, TN, USA.
[3*]Department of Medicine, Vanderbilt University Medical Center, TN, USA.
[4]Institute for Immunology and Infectious Diseases, Murdoch University, Murdoch, Western Australia.

*Corresponding author(s). E-mail(s): s.mallal@vumc.org;
Contributing authors: amir.asiaeetaheri@vumc.org; natalie.f.mallal@vanderbilt.edu; elizabeth.j.phillips@vumc.org;

Abstract

Humans remain subject to ancient ecological forces in which persistent, human-adapted microbes act as partners and threats. We operationalize "keystone organisms" as pathogens whose control requires coordinated engagement of multiple immune arms and whose residence is structured across tissue niches. Using 18 curated immunological and evolutionary traits across 43 organisms, unsupervised analyses resolved four archetypes; keystones score high on multi-arm coordination, co-evolutionary history and structured tropism, whereas non-human-specific pathogens and opportunists do not (Fig.1). Independent pairwise matrices (pathogen×immune-perturbation; pathogen×niche) yielded an interpretable systems readout—diagnostic breadth—which, together with niche use, separated keystones with 93% cross-validated accuracy and showed significant group differences (Kruskal-Wallis $p = 0.026$) (Fig. 2). At the micro level, focusing of memory on keystone-patterned epitopes creates niche-matched vulnerability to modified-self ligands, offering a mechanism for autoimmune and drug-hypersensitivity syndromes. In the same way, fast-evolving RNA viruses exploit keystone-established hierarchies by inflating immunodominant but mutable decoy epitopes that re-engage familiar memory while offering little control; unless immunogens exclude such decoys and force recognition of conserved,

functionally constrained sites, vaccine responses are diverted and outpaced by host-specific adaptation [1–6].

Keywords: coevolution; keystone organisms; immunodominance; tissue-resident memory; diagnostic breadth; molecular mimicry; RNA virus decoy; vaccine design; drug hypersensitivity; autoimmunity

1 Introduction

Throughout evolutionary history, organismal complexity has been tightly linked to the surrounding microbial environment. One of the earliest and most consequential examples was the incorporation of an oxygen-fixing prokaryote into the primordial eukaryotic cell, a mutualistic event that gave rise to mitochondria and enabled aerobic multicellular life to flourish in an oxygen-rich atmosphere more than two billion years ago [7]. This foundational instance of endosymbiosis illustrates a broader principle: successful organisms, including humans, have co-evolved in dynamic, reciprocal relationships with microbial partners.

A dynamic macro-evolutionary backdrop.

Over deep time, host-microbe relationships have been highly dynamic. Some lineages co-diverged with vertebrates and became near-obligate, tissue-structured companions, whereas others crossed species barriers, replaced prior residents, or entered host genomes as endogenous elements. Today's human keystones therefore represent only some outcomes of this churn, and intermediate groups in our framework likely capture organisms in transit, either tightening toward mutualistic persistence or exploiting transient ecological opportunity. This view reconciles strong host-virus co-divergence in some families with episodic turnovers and host shifts in others, and it frames immunological hierarchies as products of on-going dynamism rather than fixed properties [3, 8].

While bacteria have long dominated discussions of host-microbe mutualism, viruses are increasingly recognized as contributors to host fitness across plants, animals, and humans [8, 9]. Viral elements can facilitate development, shape ecological and immunological adaptability, and calibrate immune function. These observations challenge a purely defensive view of immunity and instead support a model of continuous, calibrated co-regulation between host and infectious organisms [9].

Building on this ecological and evolutionary framing, Virgin, Wherry, and Ahmed argued that chronic viral infections, especially the human herpesviruses, are stable and dynamic components of our metagenome that imprint the immune system and calibrate future responses [3]. They emphasized co-evolutionary depth, the capacity to establish metastable latency, and subtle, periodic reactivation that demands coordinated immune engagement without immunopathology. These hallmarks overlap with what we term ***keystone properties***.

We propose that a subset of persistent, human-adapted organisms function as *immunological keystones*. By occupying structured tissue niches and demanding coordinated $CD8^+$, $CD4^+$, B cell, NK, and innate control, they imprint postnatal immune hierarchies. We refer to this postnatal focusing of immunodominance as a third stage of selection, *keystone epitope focusing*. This perspective helps explain why immunodominance is adaptive, why it can become vulnerable to mimicry, and why adaptable RNA viruses evolve strategies that decoy these learned priorities. Evidence for niche-structured coordination spans barrier tissues and internal compartments [e.g., 10–12].

The term *keystone* originates in architecture, where the central wedge-shaped stone at the apex of an arch bears and distributes weight to hold the structure in balance. Ecologists later adopted the concept to describe *keystone species*, organisms whose presence exerts an outsized influence on ecosystem stability [13]. By analogy, we use *keystone organism* to describe a class of co-evolved human-specific microbes that not only persist, but anchor immune architecture through multi-arm coordination and lifelong imprinting. Just as the loss of a keystone species disrupts ecological integrity, the absence or delayed acquisition of keystone organism infection to later life may have cascading effects on immune compartmentalization and resilience.

Not all persistent organisms meet this definition. Microbes exist along a continuum. At one end are acute, external threats that trigger transient, compartmentalized immune responses. At the other are deeply integrated symbionts that not only require containment, but also persist through sustained, multi-arm coordination. For these organisms, coordinated immune control is not merely host defense, it is a prerequisite for stable persistence within specific cellular niches or tissue compartments. This compartmental equilibrium represents a preferred ecological state. The relationship is reciprocal as well: such organisms may shape and calibrate the immune system to favor host survival, indirectly protecting against more virulent competitors. Between these poles lies a spectrum whose degree of keystoneness reflects persistence, specificity, evolutionary intimacy, and immunological impact. The lack of a systematic framework to classify these relationships has hindered our understanding of how particular microbes organize immune hierarchies, not only in protection, but also in priority setting, redundancy, and baseline calibration.

A macro to meso to micro funnel.

To bridge concept and measurement, we adopt a macro to meso to micro funnel. At the macro level, persistent human-adapted infections shape durable immune hierarchies [3]. At the meso level, the organisms we call keystones require, and thereby coordinate, redundant containment by $CD8^+$, $CD4^+$, B cells, NK cells, and innate effectors while occupying structured tissue niches [14, 15]. At the micro level, such focusing of immunodominance creates a vulnerability: modified self ligands that resemble keystone epitopes in the same niche can recruit pre-existing memory and bypass peripheral regulation, producing autoimmunity or T cell mediated drug hypersensitivity [1, 5, 16]. In parallel, fast-evolving RNA viruses exploit these hierarchies: if an immunogen re-stimulates keystone-imprinted responses that favor the pathogen,

immunity may be decoyed toward mutable epitopes rather than conserved, functionally constrained sites—at least long enough for viral adaptation to enable persistence [2, 4, 6].

In this study, we develop a rigorous trait-based scoring system to evaluate 43 candidate organisms across 18 immunological traits spanning species specificity, early-life acquisition, lifelong persistence/latency, co-evolutionary duration, multi-arm immune coordination, structured compartmental tropism, immune modulation/evasion, mimicry, dynamic equilibrium and non-cytolytic control, as well as epitope-centric and receptor-selection features (Table 1). We then validate these organism-level signals against independent pathogen×immune-perturbation and pathogen×niche matrices to derive two interpretable diagnostics—diagnostic breadth and tissue tropism—that predict keystoneness and explain clinical patterns under immunosuppression.

We address the resulting gap by proposing measurable criteria for evaluating organism keystoneness. Specifically, we combine trait-based clustering to resolve pathogen archetypes with pairwise immune-perturbation × niche matrices that quantify diagnostic breadth and compartmental tropism. While our analysis focuses on pathogens, this framework lays the groundwork for deeper examination of the specific epitopes that mediate multi-arm coordination. Such a shift, from organism level to epitope level, sets the stage for mechanistic studies of mimicry risk and the development of mechanism-specific diagnostics and therapeutics.

2 Results

2.1 Trait Curation and Rationale

To quantify pathogen keystoneness, we curated a panel of 18 immunological traits capturing features hypothesized to drive multi-arm immune coordination, structured tissue residence, and postnatal immune imprinting. Traits were selected based on mechanistic links to immune calibration, co-evolutionary intimacy, compartmental persistence, and capacity to influence host immune priorities. Our goal was to define an operational space for "keystone-like" behavior that could be scored across diverse pathogens.

Scoring followed a rubric-anchored system applied to structured literature reviews, prioritizing human evidence in the following order: (i) genetic, interventional, or large epidemiologic data; (ii) mechanistically plausible human clinical evidence; and (iii) supporting animal or in vitro results only when necessary. Each trait was scored using an allowed integer set defined a priori (e.g., {0,1} or {0,1,2,3}), with higher values reflecting stronger or more specific evidence for the trait. For each pathogen–trait pair, evidence was extracted from the literature using consistent synonyms and selection criteria (see Methods). Table 1 summarizes the scoring rules for each trait.

Table 1 Trait definitions and score interpretations used to assess keystoneness.

Trait	Allowed Scores	Score Interpretation
Human-specific	{0,1}	1: Sustained human-only transmission; 0: Multi-host or zoonotic cycle.
Latency and Lifelong Infection	{0,1,2,3}	3: Lifelong latency with immune calibration; 2: Common persistence; 1: Limited or sporadic persistence; 0: None.
Length of Co-evolution with Humans	{0,1,2,3}	3: Deep-time co-divergence; 2: Long-standing association; 1: Recent emergence; 0: Recent zoonosis.
Global Prevalence	{0,1,2,3}	3: High global burden; 2: Intermediate; 1: Low or regional; 0: Rare.
Genetic Polymorphisms Reflecting Human Migration	{0,1,2,3}	3: Pathogen lineages tightly track human migration; 2: Partial coupling; 1: Contextual; 0: None.
Prevalence in Modern Indigenous Populations	{0,1,2,3}	3: High prevalence in multiple isolated populations; 2: Replicated across groups; 1: Limited; 0: None.
Consequences of Delayed Acquisition	{0,1,2,3}	3: Delayed infection causes severe disease; 2: Common age-related effects; 1: Context-specific; 0: None.
Association with Transplant Rejection	{0,1,2,3}	3: Robust, replicated clinical link; 2: Plausible but mixed; 1: Limited; 0: None.
Cross-reactivity with Adaptable RNA Viruses	{0,1}	1: Documented heterologous immunity or cross-reactivity; 0: None.
Perturbation by RNA Viruses	{0,1,2,3}	3: Frequent reactivation/modulation by RNA viruses; 2: Repeated in defined contexts; 1: Sporadic reports; 0: None.
Hypersensitivity and Autoimmune Flares	{0,1,2,3}	3: Reproducible links to autoimmune flares or immune complex disease; 2: Multiple plausible reports; 1: Sparse; 0: None.
Structured Immune Receptor Selection	{0,1,2,3}	3: Public or stereotyped TCR/BCRs with function; 2: Convergent motifs with partial function; 1: Scattered hints; 0: None.
Proximity and Persistence of Immune Cells	{0,1,2,3}	3: TRM or TLS consistently present with systemic effects; 2: Frequent persistence with limited scope; 1: Sparse; 0: None.
Epitope Structure and Immunodominance	{0,1,2,3}	3: Reproducible dominant epitopes across cohorts; 2: Replicated patterns; 1: Limited; 0: None.
Absence of Peripheral Tolerance	{0,1,2,3}	3: Established mimicry with autoimmune outcomes; 2: Strong evidence of tolerance break; 1: Suggestive hints; 0: None.
Epitope Presentation Dynamics	{0,1,2,3}	3: Strong HLA associations or MHC modulation; 2: Replicated effects; 1: Limited; 0: None.
Multi-Arm Coordination	{0,1}	1: Coordinated innate–adaptive containment with redundancy; 0: Partial or local coordination only.
Structured Tropism	{0,1}	1: Multi-niche or systemic tissue persistence; 0: Localized or stochastic tropism.

2.2 Immunological trait structure resolves four pathogen archetypes (Fig. 1)

PCA reveals organized dispersion with interpretable loadings.

Principal component analysis of 18 immunological traits across 43 organisms produced a structured biplot in which the first two components explain 52.9% of total variance (PC1 42.2%, PC2 10.7%). Organisms spread primarily along PC1. Trait

Table 2 Keystone organisms and their immunological and evolutionary characteristics. These pathogens exhibit multi-arm immune coordination, structured tissue tropism, and deep co-evolution with humans. Reactivation in immunosuppressed hosts reflects systems-level failure. Abbreviations in parentheses match those used in Figures and trait score tables.

Organism (Abbr.)	Type	Infection Pattern	Geo. Dist.	Key Immune Deps.	Lifelong?	Tissues/Cell Involved	Immunosuppression Impact	Co-evo. Time
Cytomegalovirus (CMV)	Herpesvirus	Chronic latency	Global	NK, $CD8^+$, $CD4^+$, Abs	Yes	Monocytes/Macrophages, Endothelium	Post-transplant reactivation; severe disease with T/NK suppression	5–25 Mya
Epstein–Barr Virus (EBV)	Herpesvirus	Chronic latency	Global	$CD8^+$, NK; $CD4^+$ help	Yes	Memory B cells, Oropharyngeal epithelium	PTLD with B-cell depletion; risk with $CD8^+$ defects	12–14 Mya
Herpes Simplex Virus 1 (HSV-1)	Herpesvirus	Chronic latency	Global	$CD8^+$ (TRM), NK; $CD4^+$ help	Yes	Mucosal epithelium, Sensory neurons	Severe mucocutaneous/CNS disease in NK or $CD8^+$ defects	∼6 Mya
Herpes Simplex Virus 2 (HSV-2)	Herpesvirus	Chronic latency	Global	$CD8^+$, NK; $CD4^+$ help	Yes	Genital mucosa, Sensory neurons	Severe genital/CNS disease in NK or $CD8^+$ defects	<1.6 Mya
Human Herpesvirus 6 (HHV-6)	Herpesvirus	Chronic latency	Global	$CD4^+$, $CD8^+$, NK	Yes	$CD4^+$ T cells, CNS	Encephalitis/reactivation post-transplant; risk with $CD4^+$ loss	∼30 Mya
Mycobacterium tuberculosis (Mtb)	Bacteria	Intracellular persistence	Global	$CD4^+$, TNF, IFN-γ	Yes	Alveolar macrophages, Granulomas	Reactivation with TNF blockade or $CD4^+$ loss	0.07–3 Mya
Plasmodium falciparum (P.fal)	Protozoan	Chronic infection	Trop/Subtrop	$CD4^+$, IFN-γ, Abs	No	Hepatocytes, RBCs, Microvasculature	Severe disease with impaired $CD4^+$ responses; pregnancy risk	50–100 kya
Varicella–Zoster Virus (VZV)	Herpesvirus	Chronic latency	Global	$CD4^+$, $CD8^+$ (TRM)	Yes	Dorsal root ganglia, Skin	Zoster with aging or T-cell suppression; dissemination on high-dose steroids	25–30 Mya

Table 3 Specialist organisms and their immunological and evolutionary characteristics. These pathogens show focused tropism and epitope-centric strategies with moderate coordination, often maintaining chronic infection without broad systems-level imprinting. Abbreviations in parentheses match those used in Figures and trait score tables.

Organism (Abbr.)	Type	Infection Pattern	Geo. Dist.	Key Immune Deps.	Lifelong?	Tissues/Cell Involved	Immunosuppression Impact	Co-evo. Time
BK Virus (BKV)	Polyomavirus	Chronic latency	Global	$CD8^+$, $CD4^+$, Abs	Yes	Renal tubular epithelium	BK nephropathy post-transplant; risk with T-cell dysfunction	>400 kya
Candida albicans (C.alb)	Fungus	Colonization/overgrowth	Global	Th17, Neutrophils, Macrophages	Yes	Oral/genital mucosa, Skin	Mucocutaneous disease with IL-17 blockade or steroids	>20 Mya
Helicobacter pylori (H.pyl)	Bacteria	Chronic colonization	Global	Th1/Th17; B cells	Yes	Gastric epithelium	Persistent colonization; activity shaped by T-cell/IL-17 pathways	60–100 kya
Human Papillomavirus (HPV)	Papillomavirus	Chronic latency	Global	$CD8^+$, $CD4^+$, NK	Yes	Basal epithelial cells (anogenital/oropharyngeal)	High-grade lesions with T-cell defects; transplant risk	∼500 kya
JC Virus (JCV)	Polyomavirus	Chronic latency	Global	$CD4^+$, $CD8^+$, IFN-γ	Yes	Oligodendrocytes, Astrocytes	PML with $CD4^+$ loss or integrin blockade	400–500 kya
Kaposi's Sarcoma-Assoc. Herpesvirus (KSHV)	Herpesvirus	Chronic latency	Focal	$CD8^+$, NK; $CD4^+$ help	Yes	Endothelium, B cells	KS/viremia in HIV/transplant; T-cell suppression	20–30 Mya
Merkel Cell Polyomavirus (MCPyV)	Polyomavirus	Chronic latency	Global	$CD8^+$, $CD4^+$, IFN-γ	Yes	Merkel cells, Skin	Merkel cell carcinoma risk with T-cell suppression	>500 kya
Neisseria meningitidis (N.men)	Bacteria	Colonization + invasion	Global	Complement (C5–9), B cells	No	Nasopharynx, Meninges	Invasive disease with complement deficiency/asplenia	500–1000 kya
Parvovirus B19 (PVB19)	Virus	Seasonal infection	Global	Neutralizing Abs; $CD8^+$	No	Erythroid progenitors (marrow)	Chronic anemia in immunodeficiency/transplant	10–20 kya
Pneumocystis jirovecii (P.jir)	Fungus	Chronic colonization	Global	$CD4^+$, Macrophages	No	Alveoli (type I/II interface)	PCP with $CD4^+$ loss (AIDS, steroids)	100–200 kya
Treponema pallidum (T.pal)	Bacteria	Mucosal contact	Global	$CD4^+$, Macrophages	No	Skin/mucosa, Endothelium, CNS	Neurosyphilis/relapses with T-cell suppression	2.5–15 Mya

Table 4 Multi-hostorganisms with broad tissue engagement but lacking deep co-evolution, human specificity, or stable imprinting. They can trigger strong immune activation or mimicry yet remain evolutionarily unsettled. Abbreviations in parentheses match those used in Figures and trait score tables.

Organism (Abbr.)	Type	Infection Pattern	Geo. Dist.	Key Immune Deps.	Lifelong?	Tissues/Cell Involved	Immunosuppression Impact	Co-evo. Time
Aspergillus spp. (Asper)	Fungus	Airborne inhalation	Ubiquitous	Neutrophils, Macrophages, $CD4^+$	No	Lungs, Paranasal sinuses; CNS (disseminated)	Invasive disease with neutropenia or high-dose steroids	<10 kya
Cryptococcus neoformans (C.neo)	Fungus	Environmental exposure	Global	$CD4^+$, Macrophages	Yes	Lungs, CNS (meningitis)	Severe disease in $CD4^+$ loss; transplant risk	Millions (environmental)
Histoplasma capsulatum (H.cap)	Fungus	Inhalation	Endemic (Americas)	Macrophages, $CD4^+$, IFN-γ	Yes	Lungs, RES (liver, spleen)	Dissemination with $CD4^+$/TNF blockade	<10 kya
Leishmania spp. (Leish)	Protozoan	Vector-borne	Trop/Subtrop	Macrophages, IFN-γ, TNF	Yes	Skin, Liver/Spleen (RES)	Reactivation with TNF/IFN-γ suppression	>15 Mya (clade)
Mycobacterium avium complex (MAC)	Bacteria	Opportunistic	Global	Macrophages, $CD4^+$	No	Lungs, RES	Dissemination with $CD4^+$ loss (AIDS)	<10 kya
Schistosoma spp. (Schis)	Helminth	Chronic infection	Trop/Subtrop	Th2, Eosinophils, Macrophages	Yes	Mesenteric/vesical veins, Liver, GI	Granulomas; immune remodeling; steroid effects	1–4 Mya
Strongyloides stercoralis (S.ste)	Helminth	Auto-infection	Trop/Subtrop	Th2, Eosinophils	Yes	GI tract, Lungs	Hyperinfection on steroids; HTLV-1 co-risk	20–50 kya
Toxoplasma gondii (T.gon)	Protozoan	Food/zoonotic	Global	$CD8^+$, $CD4^+$, IFN-γ	Yes	CNS, Retina, Muscle	Encephalitis/retinitis with $CD4^+$ loss	~10 kya (domestication)
Trypanosoma cruzi (T.cru)	Protozoan	Vector-borne	Latin America	$CD8^+$, $CD4^+$, IFN-γ	Yes	Cardiac muscle, GI, CNS	Reactivation in AIDS/transplant	~9 kya

Table 5 Opportunistic organisms characterized by low coordination, environmental ubiquity, and narrow immune dependencies. Their emergence under immunosuppression flags pathway-specific failure rather than systemic collapse. Abbreviations in parentheses match those used in Figures and trait score tables.

Organism (Abbr.)	Type	Infection Pattern	Geo. Dist.	Key Immune Deps.	Lifelong?	Tissues/Cell Involved	Immunosuppression Impact	Co-evo. Time
Acanthamoeba spp. (Acan)	Amoeba	Contact/Inhalation	Global (soil, water)	$CD4^+$, Macrophages	No	Cornea, CNS	Keratitis/CNS disease with T-cell defects	Unknown
Balamuthia mandrillaris (B.man)	Amoeba	Contact/Inhalation	Soil, water	$CD4^+$	No	CNS, Skin	Granulomatous amebic encephalitis in T-cell suppression	Unknown
Bartonella spp. (Barto)	Bacteria	Vector-borne	Ubiquitous (cats, lice)	$CD4^+$, Macrophages	Yes	Endothelium, Erythrocytes	Bacillary angiomatosis with T-cell suppression	∼10 kya
Brucella spp. (Bruce)	Bacteria	Zoonotic	Livestock-associated	$CD4^+$, IFN-γ, Macrophages	Yes	RES, Reproductive tract	Chronic infection with TNF/IFN-γ axis defects	<10 kya
Burkholderia cepacia complex (Bcc)	Bacteria	Contact/Inhalation	Water, hospitals	Neutrophils; oxidative burst	No	Lungs, Bloodstream	Severe in CGD/CF; neutropenia risk	<10 kya
Coccidioides spp. (Cocci)	Fungus	Inhalation	Endemic (US SW)	$CD4^+$, Macrophages	No	Lungs, CNS	Dissemination with $CD4^+$ loss	<10 kya
Coxiella burnetii (C.bur)	Bacteria	Inhalation	Global (dust/aerosol)	IFN-γ, TNF, Macrophages	No	Lungs, Liver, Endothelium	Severe disease with TNF/IFN-γ blockade	<10 kya
Fusarium spp. (Fusar)	Fungus	Inhalation/Direct	Global (vegetation)	Neutrophils, Macrophages	No	Skin, Lungs, Blood	Dissemination with neutropenia/steroids	<10 kya
Human Herpesvirus 7 (HHV-7)	Herpesvirus	Chronic latency	Global	$CD4^+$, NK	Yes	T cells ($CD4^+$)	Reactivation with T-cell suppression	25–30 Mya
Legionella pneumophila (L.pne)	Bacteria	Inhalation	Water systems	IFN-γ, Macrophages	No	Lungs, Macrophages	Severe pneumonia with corticosteroids or TNF blockade	<100 years
Leptospira spp. (Lepto)	Bacteria	Zoonotic contact	Trop/Subtrop	Neutrophils, Macrophages, Abs	No	Kidney, Liver, Endothelium	Severe leptospirosis with innate/B-cell defects	5–10 kya
Nocardia spp. (Nocar)	Bacteria	Inhalation	Global (soil)	$CD4^+$, Macrophages	No	Lungs, Brain, Skin	Dissemination with T-cell suppression	<10 kya
Scedosporium spp. (Scedo)	Fungus	Inhalation	Global (stagnant water)	Neutrophils, Macrophages	No	Lungs, CNS	Severe disease with neutropenia; transplant risk	<10 kya
Torque Teno Virus (TTV)	Virus	Ubiquitous viremia	Near-universal	Tracks T-cell status (biomarker)	No	Blood, Lymphocytes	Higher loads with T-cell suppression; biomarker of net state	>500 kya
Vibrio spp. (Vibrio)	Bacteria	Ingestion/Wound	Coastal/Estuarine	Neutrophils, Complement	No	GI tract, Skin; Liver disease risk	Severe disease with liver disease or complement defects	<10 kya

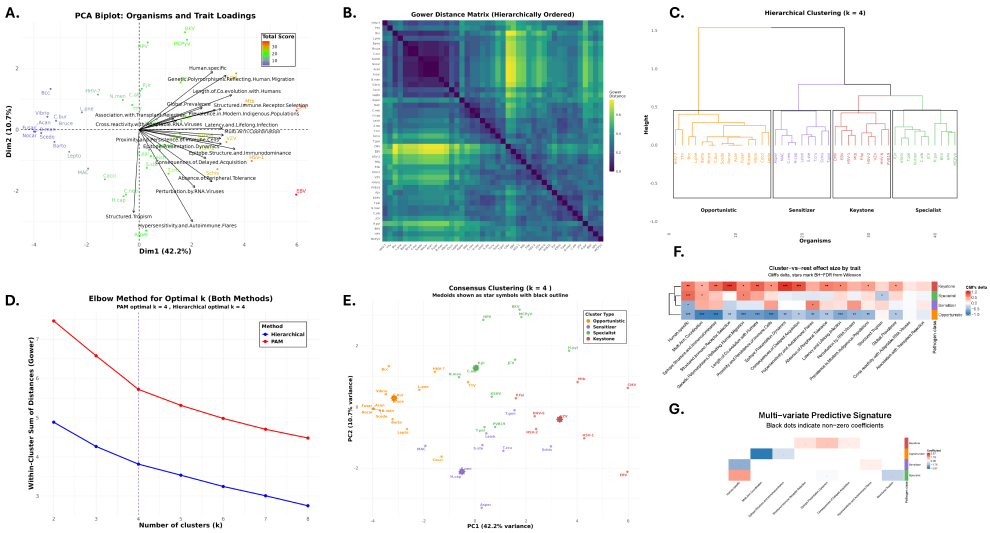

Fig. 1 Unsupervised trait clustering resolves four pathogen immunological archetypes. **(A)** PCA biplot of 18 traits across 43 organisms. The first two components explain 52.9% of variance (PC1 42.2%, PC2 10.7%). Points are colored by total score, and black arrows show trait loadings. **(B)** Gower distance matrix, hierarchically ordered, shows block-diagonal structure that supports discrete groupings. **(C)** Agglomerative hierarchical clustering on Gower dissimilarities, cut at $k = 4$, reveals four clades that correspond to the archetypes. **(D)** Elbow curves of the within-cluster sum of dissimilarities for PAM and hierarchical methods nominate $k = 4$ for both, as indicated by the vertical dashed line. **(E)** Consensus cluster assignments visualized in PCA space. Medoids are star symbols with black outlines. The two algorithms showed good agreement but disagreed for four organisms. We applied a conservative consensus rule that assigns discordant organisms to the lower-scoring of the two candidate clusters. Under this rule Parvovirus B19 moves from *Keystone* to *Specialist*, and three others were downgraded as *Specialist* → *Opportunist*. **(F)** Cluster-versus-rest effect sizes by trait using Cliff's delta. Stars mark traits with Benjamini–Hochberg FDR-adjusted significance from Wilcoxon tests. **(G)** Multivariate predictive signature from penalized multinomial regression. Tiles show signed coefficients for selected traits by class, with black dots marking non-zero coefficients. Keystone organisms are enriched for multi-system and history traits, specialists are epitope-centric with the lowest structured tropism, Multi-hostshow hypersensitivity signals, and opportunists score low across most traits. The keystone pole (high multi-arm, structured tropism) matches postnatal imprinting of immunodominance predicted by Keystone Epitope Theory.

loading vectors separate multi-system and history traits *(length of co-evolution with humans, global prevalence, genetic polymorphisms tracking human migration, proximity and persistence of immune cells, structured tropism, multi-arm coordination)* from epitope-centric traits *(epitope presentation dynamics, epitope structure and immunodominance, structured immune receptor selection)*. This pattern indicates non-random organization in the trait space (Fig. 1A).

Pairwise distances display block structure consistent with discrete groups. The hierarchically ordered Gower distance matrix shows clear block-diagonal structure, with dark blocks along the diagonal and lighter off-block regions. This pattern supports the presence of discrete clusters rather than a single continuum (Fig. 1B).

Four clusters capture the dominant structure.

Both clustering strategies, partitioning around medoids (PAM) and agglomerative hierarchical clustering on Gower dissimilarities, were evaluated for $k = 2, \ldots, 8$. The elbow and angle heuristics applied to the within-cluster sum of distances nominate $k = 4$ for both methods, after which additional clusters offer diminishing returns (Fig. 1D).

Method concordance and conservative consensus labeling.

Hierarchical clustering yields four coherent clades that align closely with PAM clusters (Fig. 1C,E). Agreement, quantified by the Adjusted Rand Index, was good but not perfect, with four organisms receiving different labels across the two methods. We therefore adopted a conservative consensus rule: for each discordant organism, we compared the average total trait score of the two candidate clusters and assigned the organism to the cluster with the smaller average score. This rule is conservative because higher total scores indicate a higher probability of keystone-like behavior. Under this consensus, Parvovirus B19 (PVB19) moved from *Keystone* in the hierarchical solution to *Specialist*. The remaining three discordant organisms fell on *Specialist* versus *Opportunistic*, and were assigned to the lower-scoring *Opportunistic* group. All downstream analyses use these consensus labels.

Cluster semantics and biological interpretation.

We name the four archetypes by their dominant trait signatures. *Keystone* organisms have high scores on multi-arm coordination, structured tropism, length of co-evolution, and global prevalence. *Specialists* emphasize epitope-centric strategies, with high epitope presentation dynamics and epitope structure or immunodominance, and they show the lowest structured tropism among the four groups. *Multi-host* are enriched for hypersensitivity and autoimmune flares, consistent with their name. *Opportunists* show low values across most traits, consistent with limited specialization. These interpretations match the distribution of organisms in PCA space and their relative total scores (Fig. 1A,E). High scores on multi-arm coordination and structured tropism in the keystone group are the trait-level signature of postnatal imprinting: these organisms plausibly coordinate Class I and II targeting and sustain tissue-resident memory in defined niches across life.

Trait signatures, univariate and multivariate.

Cluster-versus-rest effect sizes computed with Cliff's delta highlight distinctive single-trait signatures and identify traits with statistically significant separation after Benjamini–Hochberg false discovery rate control (Fig. 1F). Multivariate modeling with a sparse multinomial logistic regression yields a compact, predictive signature that agrees with the univariate patterns: keystone membership is driven by multi-system and history traits, specialists by epitope-centric traits, and Multi-hostby hypersensitivity, with opportunists characterized by negative coefficients across most features (non-zero coefficients marked with black dots, Fig. 1G).

2.3 Pathogen Archetypes as Immunological Diagnostic Signatures

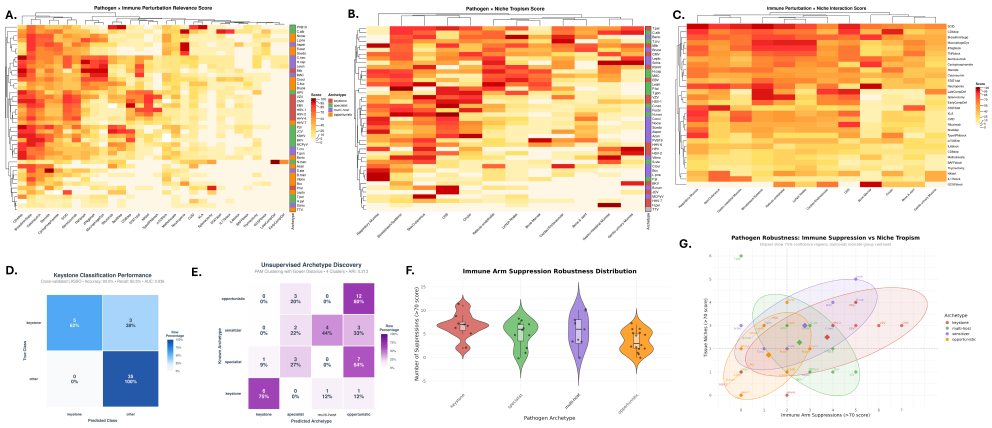

Fig. 2 Pathogen archetypes reveal distinct patterns of immune vulnerability and tissue tropism. (**A**) Pathogen × Immune Perturbation matrix showing revelation scores - how strongly each immune deficiency reveals specific pathogens (score >70 indicates strong diagnostic value). (**B**) Pathogen × Niche tropism scores indicating tissue-specific pathogen presence. (**C**) Immune Perturbation × Niche interaction matrix revealing tissue-specific immune dependencies. (**D**) Supervised classification (elastic net, $\alpha = 0.5$) successfully distinguishes keystone pathogens from others (accuracy: 93%, AUC: 0.936) using combined immune and niche features. (**E**) Unsupervised clustering (PAM with Gower distance) recovers four pathogen archetypes without prior labels (ARI: 0.213). (**F**) Distribution of diagnostic breadth - number of immune perturbations that reveal each pathogen archetype (Kruskal-Wallis $p = 0.026$). (**G**) Strategic landscape plotting diagnostic breadth (x-axis) against tissue tropism (y-axis), with 75% confidence ellipses showing distinct ecological niches for each archetype. Panel F–G formalize diagnostic breadth as the systems signature that makes keystone reactivation a multi-arm failure signal under immunosuppression.

Immune Perturbation Patterns Define Pathogen Archetypes

The pathogen–immune perturbation matrix (Fig. 2A) revealed striking differences in how pathogens respond to immune deficiencies. Keystone pathogens (red annotation), including CMV, EBV, and other herpesviruses, showed high revelation scores (>70) across 4–7 different immune perturbations, indicating they remain controlled by multiple redundant immune mechanisms in healthy hosts. Their reactivation serves as a diagnostic indicator of broad immunosuppression. In contrast, opportunistic pathogens (orange) showed limited revelation patterns (0–2 perturbations), suggesting they exploit specific, narrow immune vulnerabilities rather than requiring systemic immune failure. Because immunosuppression increases the incidence of all pathogen groups, it is unsurprising that the distinct signals of keystone reactivation have been difficult to detect in clinical practice. The key distinction is that keystone reactivation reflects multi-arm failure, while isolated opportunists often indicate pathway-specific compromise.

Tissue Tropism Reveals Complementary Strategies

The pathogen-niche matrix (Fig. 2B) demonstrated that tissue tropism operates independently of immune revelation patterns. Multi-hostpathogens (purple), particularly *Leishmania*, showed the broadest tissue distribution (4-5 niches), while maintaining moderate immune revelation scores. This orthogonal relationship between immune vulnerability and tissue colonization suggests pathogens face evolutionary trade-offs between investing in immune evasion versus tissue adaptation capabilities. The orthogonality of immune revelation and tropism supports compartment-specific imprinting and predicts niche-specific mimicry risk when modified self resembles a keystone epitope in that compartment.

Machine Learning Validates Archetype Distinctions

Supervised classification using combined immune and niche features (Fig. 2D) achieved 93% accuracy in distinguishing keystone pathogens from others (AUC: 0.936, recall: 62.5%), confirming that these archetypes represent biologically meaningful categories. Remarkably, unsupervised clustering (Fig. 2E) recovered similar groupings without prior labels (ARI: 0.213), with keystones and opportunistic correctly separated from the other pathogens in 75% and 80% of cases, respectively.

Diagnostic Breadth Differentiates Pathogen Strategies

Quantitative analysis of diagnostic breadth (Fig. 2F) confirmed significant differences between archetypes (Kruskal-Wallis $p = 0.026$). Keystone pathogens showed significantly higher diagnostic breadth than opportunistic pathogens (median 4.5 vs 1; Wilcoxon $p = 0.010$, Cliff's delta=0.59), indicating they are revealed by diverse immune failures. This pattern was specific to immune perturbations, as tissue tropism showed no significant differences between groups ($p = 0.12$).

Strategic Landscape Reveals Evolutionary Trade-offs

The two-dimensional strategic landscape (Fig. 2G) integrated diagnostic breadth with tissue tropism, revealing four distinct ecological niches. Keystone pathogens occupied the high diagnostic breadth region (x-axis: 4-7), forming a horizontal distribution suggesting specialization in immune-dependent persistence. Opportunistic pathogens clustered tightly in the low-low quadrant (0-2 on both axes), representing a minimalist strategy. Notably, specialist pathogens showed bimodal distribution: *T. pallidum* achieved maximum tissue tropism (6 niches) with zero immune revelation, while HPV and JCV showed the opposite pattern. The minimal overlap between archetype confidence ellipses (75% CI) and significant MANOVA results (Pillai's trace $p = 0.031$) confirm these represent distinct evolutionary strategies rather than continuous variation. Keystones occupy the high diagnostic breadth region—consistent with immune-dependent persistence. This prioritization may be exploited by RNA viruses, which mimic keystone epitopes at mutable sites to decoy immunodominance and evade durable immune control.

3 Discussion

Macro co-evolution and dynamism.

Across vertebrate history, host–microbe relationships have been dynamic. Some lineages show deep host–pathogen co-divergence and extreme species specificity, others undergo repeated host switches, replacements, or endogenization. The organisms we label keystones today are contingent end-points of this churn. Intermediate groups in our hierarchy likely capture organisms in transit, either tightening toward durable mutualism with structured niches or exploiting transient ecological opportunity. This framing explains why persistent DNA viruses such as herpesviruses and polyomaviruses often appear as deeply adapted companions, while waves of retroviruses and adaptable RNA viruses reveal episodic turnover and redirection of immune priorities (Roossinck, Nat Rev Microbiol, 2011; McGeoch, Virus Research, 2006; Virgin, Cell, 2009; Dominguez-Andres, Trends Immunol, 2019).

A quantitative framework and evolutionary snapshots.

We introduced a quantitative system to identify human-specific keystone organisms and to arrange pathogens along an immunological hierarchy. Unsupervised trait clustering across 18 curated traits resolved four archetypes with clear semantics in coordination, evolutionary history, and tropism, and a reduced representation built from pathogen-by-immune and pathogen-by-niche scores reproduced the signal with predictive separation. Within trait space, the primary axis separated multi-arm coordination and structured tropism from epitope-centric features, with keystones scoring high on the former. This aligns with the view that keystones require coordinated containment across multiple arms and occupy stable tissue niches. In cases of discordant cluster assignment, we used a conservative consensus rule to favor specificity of the keystone label. Viewed through a co-evolutionary lens, these archetypes are evolutionary snapshots along persistent–replacement trajectories: keystones reflect entrenched, co-diverged lineages, while some specialists and Multi-hostmay include host-shifted or turnover taxa that show greater heterogeneity across hosts and regions (Roossinck, Nat Rev Microbiol, 2011; McGeoch, Virus Research, 2006; Virgin, Cell, 2009).

Limitations.

Trait and pairwise scores were literature-curated with rubric-anchored integer scales, which introduces subjectivity and potential reporting bias. Missing information was treated as zeros in pairwise matrices, a conservative choice that may undercount true associations. Sample sizes for some archetypes remain modest, widening uncertainty. The elastic net classifier reflects internal cross-validation and requires external validation. Our analysis is static and cross-sectional, whereas pathogen–immune interactions evolve with age, exposure, and clinical context. Finally, the current framework operates at the organism and trait level. It does not yet resolve epitope-specific distinctions between decoy and coordinating targets, nor does it fully capture niche-specific tissue-resident memory circuits. These limitations motivate the Keystone Protein Atlas and future molecular studies that resolve tissue-specific coordination and mimicry risk.

Postnatal focusing of immunodominance.

Beyond thymic selection, postnatal imprinting by keystone organisms focuses immunodominance onto coordinating class I and class II epitopes, aligning CD4, CD8, B cell, NK, and innate programs and seeding tissue-resident memory in anatomical niches. This mechanism explains why trait-defined keystones anchor the strategic landscape at high diagnostic breadth and why their reactivation is particularly informative for systems-level integrity (Zinkernagel and Doherty, Nature, 1974; Hislop, Annu Rev Immunol, 2007; Malouli, J Clin Invest, 2014).

Hierarchy, redundancy, and the strategic landscape.

Pairwise analyses revealed that keystone pathogens are broadly unmasked across immune perturbations, consistent with redundant control in healthy hosts. Summarizing this as diagnostic breadth yielded a simple, testable statistic that separates keystones from opportunists. Combining breadth with tissue tropism placed the four archetypes into distinct ecological positions, suggesting trade-offs between immune evasion, tissue adaptation, and long-term persistence. Because breadth and tropism are partially orthogonal, the framework captures both systems-level dependence and compartmental imprinting.

Impact of immunosuppression on differnt classes of organisms.

The breadth signature proposes a practical readout for depth of immunosuppression in transplantation, hematologic malignancy, and targeted blockade. Concurrent reactivation or overgrowth of keystone organisms should be interpreted as a multi-arm failure signal, while isolated opportunists more often mark pathway-specific defects. The immune-perturbation by niche matrix also highlights site-specific vulnerabilities that can guide surveillance and prophylaxis. Because immunosuppression increases incidence across all pathogen groups, it is unsurprising that distinctive keystone signals have been hard to discern; the present framework separates pattern from prevalence without implying prior oversight [3].

Box 1. Archetype-aware clinical rules

- **Interpret keystone reactivation as multi-arm failure.** When one or more keystone-family infections reactivate together, or recur within a short interval, treat this as a systems-level alarm rather than a single-pathway breach [**?**].
- **Monitor depth with a *Keystone Reactivation Index*.** Track the count of keystone-family reactivation events per patient-time, optionally weighted by severity and adjusted by the implicated niche. Use changes in the *Keystone Reactivation Index* over time to titrate immunosuppression, escalate diagnostics, and guide prophylaxis [**?**].
- **Map syndrome to niche, then surveil the matching keystone.** Align presenting signs and symptoms to tissue compartments and resident memory circuits, then focus surveillance on the keystone organisms most relevant

to that niche, for example cutaneous-mucosal surfaces, lymphoid tissue, and vascular-myeloid reservoirs [14, 15].

- **Treat isolated opportunists as pathway-specific alarms.** Single-organism events outside the keystone set usually indicate a narrow deficit, for example a specific cellular or cytokine pathway, rather than a global system failure. Investigate the corresponding arm of immunity [**?**].
- **Avoid decoy-biased immunogens for RNA viruses.** Prioritize functionally constrained epitopes that are co-presented on class I and class II in the correct anatomical context. Avoid highly mutable decoy epitopes that divert immunodominance and weaken protection [4, 6, 17].

How to calculate the Keystone Reactivation Index: choose a window W (for example 30 to 90 days). Let E be the set of confirmed reactivation events from the keystone set. Define ***Keystone Reactivation Index***

$$\mathrm{KRI}(W) = \sum_{e \in E} w_e \,,$$

where $w_e = 1$ for simple counts or can scale with severity, viral burden, or site relevance. Track the change in *Keystone Reactivation Index* between adjacent windows to detect worsening multi-arm compromise.

Keystone mimicry and divergent immune vulnerabilities.

Because keystone epitopes are not subject to peripheral tolerance, any modified self that mimics a keystone epitope, particularly in the same niche, risks reactivating pre-positioned memory T cells. This can trigger autoimmunity or T cell-mediated drug hypersensitivity such as Stevens-Johnson syndrome or toxic epidermal necrolysis. These outcomes are rare and depend on multiple converging factors, including structural mimicry, HLA restriction, and the presence of tissue-resident memory capable of immediate response. This layering explains the clinical paradox of high negative predictive value and low positive predictive value for many HLA risk alleles. Genetic risk alone is insufficient without contextual antigenic mimicry in the right compartment (Phillips and Mallal, J Clin Invest, 2018; Pavlos, Annu Rev Med, 2015; Gibson, Nat Commun, 2024; Ostrov, PNAS, 2012; Metushi, PLoS One, 2015).

By contrast, IgE-mediated allergy follows a different logic. It is influenced by route of exposure, timing, and barrier status, tends to be less durable, and is not shaped by keystone-trained T cell memory. This difference helps explain why T cell diseases often persist once triggered, while many IgE-mediated allergies wane over time.

RNA viruses as decoy exploiters and design implications.

RNA viruses evolved within keystone-shaped hierarchies and often redirect responses toward mutable epitopes, drawing immunity away from functionally constrained targets. This decoying can frustrate vaccines that emphasize immunogenicity over

coordination. Rational design should prioritize constrained epitopes, ensure coordinated class I and class II co-presentation in the correct anatomical context, and avoid decoys that soak up T cell help without improving clearance. Where possible, immunogens should reinforce keystone-aligned hierarchies rather than compete with them (Moore, Science, 2002; Almeida, J Immunol, 2011; Hertz, J Virol, 2011; Kiepiela, Nature, 2004; Files, J Virol, 2022).

Future directions: Keystone Protein Atlas, ancestry mismatch, and tests of dynamism.

A near-term priority is a Keystone Protein Atlas by niche. For each compartment, define sentinel proteins that coordinate class I and class II presentation across arms, such as CMV pp65 or immediate-early proteins, EBV latent and early proteins, and HSV glycoproteins. Integrating epitope mapping with public TCR mining and single-cell readouts in the relevant compartments will test co-presentation, tissue-resident memory formation, and cross-reactivity (Hislop, Annu Rev Immunol, 2007; Malouli, J Clin Invest, 2014). The framework also predicts population structure shaped by migration and admixture, with measurable discordance when host and pathogen ancestries are mismatched. Overlaying pathogen phylogeography and host–pathogen ancestry on the breadth by tropism plane should reveal whether intermediate groups are enriched for host shifts or replacements, for example EBV or CMV clade structure and Helicobacter pylori mismatch analyses (Correia, J Virol, 2018; Charles, PNAS, 2023; Kodaman, PNAS, 2014; Munoz-Ramirez, ISME J, 2021; Dominguez-Andres, Trends Immunol, 2019).

Partners not Pathogens

The organisms that have hitched their survival to ours now help coordinate the very defenses that protect us. By treating keystones not as incidental passengers but as structural partners that calibrate immune priorities, we can read their reactivation under immunosuppression as a systems-level signal, anticipate where modified self will be pathogenic, and design vaccines and immunotherapies that avoid decoys and target what evolution has made hard to change. Understanding these relationships—partners, not just pathogens—is indispensable to diagnosing and treating the modern diseases of our time.

4 Material and Method

4.1 Methods for Fig. 1

4.1.1 Trait curation and scoring

We constructed a pathogen-by-trait matrix for 43 organisms across 18 immunological traits by AI-assisted literature review. The curation protocol prioritized human genetic, clinical, and epidemiologic evidence and assigned integer scores from prespecified sets for each trait *(for example* $\{0,1\}$*,* $\{0,1,2\}$*, or* $\{0,1,2,3\}$*).* The full prompt and instructions used for evidence search and scoring are provided in the Supplement. For organism i and trait t, the recorded value is x_{it}. A simple total score was computed

as

$$S_i = \sum_{t=1}^{18} x_{it},$$

with equal weights by design, since the primary goal was structure discovery rather than optimized prediction.

4.1.2 Preprocessing and distance metric

All analyses were performed on the raw integer trait scores. Pairwise dissimilarities were computed with the Gower metric, which accommodates mixed data types and different ranges. For organisms i and j, the Gower dissimilarity is

$$d_{ij} = \frac{\sum_{t=1}^{p} w_t \, s_{ijt} \, \delta_{ijt}}{\sum_{t=1}^{p} w_t \, s_{ijt}}, \quad \delta_{ijt} = \begin{cases} \dfrac{|x_{it} - x_{jt}|}{R_t}, & \text{numeric or ordinal,} \\ \mathbb{1}\{x_{it} \neq x_{jt}\}, & \text{binary,} \end{cases}$$

where R_t is the observed range for trait t, w_t is a trait weight set to 1, and $s_{ijt} \in \{0, 1\}$ indicates non-missingness. Missing entries are handled by s_{ijt} in the denominator.

4.1.3 Unsupervised clustering

We applied two clustering strategies to the Gower dissimilarities.

Partitioning around medoids (PAM).

PAM minimizes the within-cluster sum of dissimilarities to representative medoids. For a partition $\mathcal{C} = \{C_1, \ldots, C_k\}$ and medoids $m_c \in C_c$,

$$\mathrm{WCSD}(\mathcal{C}) = \sum_{c=1}^{k} \sum_{i \in C_c} d_{i m_c}.$$

We explored $k = 2, \ldots, 8$.

Agglomerative hierarchical clustering.

We performed hierarchical clustering using the complete-link criterion on the same Gower dissimilarities. The inter-cluster dissimilarity for clusters A and B is

$$D(A, B) = \max_{i \in A,\, j \in B} d_{ij}.$$

For display, a horizontal cut selected $k = 4$ groups.

Choosing the number of clusters

For each method we computed the within-cluster sum of dissimilarities $\mathrm{WCSD}(k)$ and inspected the elbow and the discrete-angle heuristics. Both methods agreed on $k = 4$, which we report as the parsimonious solution.

Consensus clustering and conservative reassignment

Method agreement was quantified with the Adjusted Rand Index (ARI). Given two partitions of n items with contingency counts n_{uv} for clusters u and v,

$$\mathrm{ARI} = \frac{\displaystyle\sum_{u,v}\binom{n_{uv}}{2} - \frac{\displaystyle\sum_u\binom{a_u}{2}\sum_v\binom{b_v}{2}}{\binom{n}{2}}}{\displaystyle\frac{1}{2}\left[\sum_u\binom{a_u}{2}+\sum_v\binom{b_v}{2}\right] - \frac{\displaystyle\sum_u\binom{a_u}{2}\sum_v\binom{b_v}{2}}{\binom{n}{2}}}, \qquad a_u = \sum_v n_{uv},\ b_v = \sum_u n_{uv}.$$

When PAM and hierarchical disagreed for an organism, we applied a conservative rule motivated by the keystone-search objective. Between the two candidate clusters, we computed the cluster means of the total score, $\bar{S}_c$. The organism was assigned to the cluster with the smaller $\bar{S}_c$. This procedure downgraded four pathogens, including a change for Parvovirus B19 from *Keystone* $\rightarrow$ *Specialist*, and three others were downgraded *Specialist* $\rightarrow$ *Opportunist*. All labels reported are these consensus labels.

4.1.4 Dimensionality reduction and visualization

PCA was performed on the centered and scaled trait matrix $X \in \mathbb{R}^{n\times p}$. With the singular value decomposition $X = USV^\top$, principal component scores are $Z = US$ and loadings are V. The biplot overlays organism scores and trait loading vectors. Cluster medoids from the consensus solution are shown in PCA space for visualization only.

4.1.5 Trait signature

Single-trait signatures

For each cluster c and trait t, we compared values in C_c against the rest with Cliff's delta,

$$\delta_{c,t} = \frac{\#\{x_{it} > x_{jt}\} - \#\{x_{it} < x_{jt}\}}{|C_c|\,|C_c^{\complement}|},$$

and tested with a two-sided Wilcoxon rank-sum test. p-values were corrected across traits within cluster by Benjamini–Hochberg FDR at $\alpha = 0.05$. Heatmaps display $\delta_{c,t}$ with stars for FDR-significant tests.

Multivariate predictive signature

To summarize discriminative patterns across traits, we fit a multinomial logistic regression with an ℓ_1 penalty,

$$\Pr(Y = g \mid x) = \frac{\exp(\beta_g^\top x)}{\sum_{h=1}^{k} \exp(\beta_h^\top x)}, \qquad \hat{\beta} = \arg\min_{\{\beta_g\}} \Big\{ -\sum_{i=1}^{n} \log \Pr(Y_i \mid x_i) + \lambda \sum_{g=1}^{k} \|\beta_g\|_1 \Big\},$$

using standardized traits and 5-fold cross-validation to select $\lambda = \lambda_{\min}$. The coefficient heatmap reports signed effects by class, with black dots marking non-zero coefficients. The model was used for interpretation rather than for defining clusters.

4.2 Methods for Fig. 2.

4.2.1 Definition and Rationale for Immune Perturbations

To systematically evaluate immune dependencies across pathogens, we defined 30 immune perturbations based on established clinical and immunological literature, encompassing genetic immunodeficiencies and pharmacologically induced impairments [18]. These perturbations were selected for their relevance to assessing pathogen-specific vulnerabilities, particularly for keystone pathogens that require coordinated, multi-arm immune responses for containment and immune calibration [3]. Each perturbation targets distinct immune components, enabling quantitative scoring of susceptibility, severity, or dissemination across diverse pathogens.

1. **CD4dep**: CD4+ T-cell deficiency, as in idiopathic CD4 lymphocytopenia, impairs helper T-cell functions critical for pathogens like *Cytomegalovirus* [19]. It is essential for studying adaptive immune coordination in keystone infections.
2. **CD8dep**: CD8+ T-cell deficiency disrupts cytotoxic responses, increasing susceptibility to viruses like *Epstein-Barr Virus* [18]. This perturbation is key for evaluating cytotoxic T-cell dependency in viral control.
3. **Thymectomy**: Thymectomy reduces naive T-cell output, impacting pathogens requiring robust T-cell priming, such as *Mycobacterium tuberculosis* [20]. It is relevant for assessing thymic-dependent immune reconstitution.
4. **SCID**: Severe combined immunodeficiency (SCID) abrogates T- and B-cell function, leading to severe infections by pathogens like *Pneumocystis jirovecii* [19]. Its inclusion captures profound adaptive immune defects.
5. **STAT1def**: STAT1 deficiency impairs IFN signaling, increasing susceptibility to mycobacteria and viruses [21]. It is critical for studying IFN-dependent control of intracellular pathogens.
6. **STAT3def**: STAT3 deficiency, as in hyper-IgE syndrome, disrupts Th17 and follicular helper T-cell responses, affecting pathogens like *Candida albicans* [22]. It is vital for evaluating Th17-mediated immunity.
7. **CalcineurInh**: Calcineurin inhibitors (e.g., tacrolimus) suppress T-cell activation, increasing risk for *CMV* reactivation [23]. This perturbation assesses pharmacologic T-cell suppression.

8. **mTORinh**: mTOR inhibitors (e.g., sirolimus) modulate T-cell and innate responses, impacting pathogens like *BK Virus* [23]. It is relevant for studying metabolic regulation of immunity.
9. **Alemtuzumab**: Alemtuzumab (anti-CD52) depletes lymphocytes, increasing susceptibility to herpesviruses [24]. It is essential for assessing broad lymphoid depletion effects.
10. **Steroids**: Glucocorticoids suppress multiple immune arms, exacerbating infections like *Aspergillus spp.* [25]. This perturbation evaluates broad immunosuppressive effects.
11. **NKdef**: NK-cell deficiency, as in GATA2 mutations, impairs antiviral and antitumor immunity, critical for pathogens like *Human Papillomavirus* [26]. It is key for studying innate cytotoxic responses.
12. **MacrophageDys**: Macrophage dysfunction, as in IFN-γ/IL-12 axis defects, increases susceptibility to *Mycobacterium avium complex* [21]. It is crucial for evaluating innate phagocytic control.
13. **TNFblock**: TNF blockers (e.g., infliximab) increase risk for *Mycobacterium tuberculosis* reactivation [27]. This perturbation assesses TNF-dependent immune containment.
14. **IL17block**: IL-17 inhibitors (e.g., secukinumab) impair mucosal immunity, increasing *Candida albicans* infections [28]. It is vital for studying Th17-driven mucosal defense.
15. **IFNgblock**: IFN-γ blockade, via autoantibodies or therapeutics, exacerbates mycobacterial infections [21]. It is essential for assessing IFN-γ-mediated immunity.
16. **IL6block**: IL-6 blockers (e.g., tocilizumab) modulate inflammation, impacting pathogens like *Staphylococcus aureus* [29]. This perturbation evaluates IL-6-driven immune regulation.
17. **TypeIIFNblock**: Type I IFN blockade (e.g., IFN-α/β defects) increases viral susceptibility, as seen with *Varicella-Zoster Virus* [30]. It is critical for studying antiviral innate immunity.
18. **GCSFblock**: G-CSF blockade disrupts neutrophil production, increasing risk for bacterial infections like *Vibrio spp.* [25]. It is relevant for assessing neutrophil-dependent defense.
19. **Bcelldep**: B-cell depletion impairs humoral immunity, affecting pathogens like *Streptococcus pneumoniae* [31]. It is key for evaluating antibody-mediated protection.
20. **Rituximab**: Rituximab (anti-CD20) depletes B cells, increasing risk for *Hepatitis B Virus* reactivation [31]. This perturbation assesses targeted B-cell suppression.
21. **XLA**: X-linked agammaglobulinemia (BTK deficiency) eliminates antibody production, increasing susceptibility to *Neisseria meningitidis* [32]. It is essential for studying humoral deficiency.
22. **CVID**: Common variable immunodeficiency impairs antibody and T-cell responses, affecting pathogens like *Helicobacter pylori* [32]. It is vital for evaluating combined humoral defects.
23. **BAFFblock**: BAFF blockade (e.g., belimumab) reduces B-cell survival, impacting humoral responses to pathogens [31]. It is relevant for assessing B-cell maturation defects.

24. **EarlyCompDef**: Early complement deficiencies (C1-C4) impair opsonization, increasing susceptibility to *Neisseria spp.* [33]. This perturbation evaluates early complement roles.
25. **LateCompDef**: Late complement deficiencies (C5-C9) disrupt membrane attack complex formation, critical for *Neisseria meningitidis* [33]. It is key for studying terminal complement function.
26. **BroadImmSupp**: Broad immunosuppression, as in post-transplant settings, increases risk for *CMV* and *Pneumocystis jirovecii* [23]. It is essential for assessing global immune suppression.
27. **Splenectomy**: Splenectomy impairs phagocytic and humoral immunity, increasing *Streptococcus pneumoniae* risk [32]. It is vital for studying splenic immune functions.
28. **Neutropenia**: Neutropenia increases susceptibility to *Aspergillus spp.* and *Pseudomonas aeruginosa* [25]. It is critical for evaluating neutrophil-dependent defense.
29. **Cyclophosphamide**: Cyclophosphamide, an alkylating agent, causes broad lymphopenia, increasing *CMV* risk [23]. It is relevant for assessing chemotherapeutic immunosuppression.
30. **Methotrexate**: Methotrexate, an antimetabolite, suppresses lymphocyte proliferation, impacting *Pneumocystis jirovecii* control [23]. It is key for studying antimetabolite effects on immunity.

4.2.2 Definition and Rationale for Anatomical Immune Niches

To model immune coordination and niche-specific vulnerabilities in human infections, we defined 12 anatomical immune niches based on established immunological literature and their relevance to quantitative pathogen tropism scoring. These niches reflect distinct immunological environments shaped by resident immune cell populations, anatomical barriers, and pathogen-specific interactions, enabling the study of both localized and systemic immune calibration, particularly for keystone pathogens that exert broad influence on immune architecture [3]. Each niche was selected for its clinical and mechanistic significance in pathogen persistence, dissemination, or disease manifestation, facilitating robust scoring of tropism across diverse pathogens.

1. **Bloodstream/Systemic**: This niche encompasses systemic infections characterized by bacteremia, viremia, fungemia, or parasitemia, often involving dissemination to multiple organs. It is critical for studying pathogens like *Cytomegalovirus*, which drive systemic immune calibration through broad dissemination and require coordinated innate and adaptive responses [3, 10].
2. **Central Nervous System (CNS)**: The CNS, protected by the blood-brain barrier, hosts unique immune regulation via microglia and limited peripheral immune infiltration, making it a key niche for pathogens causing meningitis or encephalitis, such as *Neisseria meningitidis* [12]. Its distinct immunological constraints are essential for assessing neurotropic pathogen effects on immune architecture.
3. **Respiratory Mucosa**: Encompassing the upper and lower respiratory tract, this niche includes bronchus-associated lymphoid tissue (BALT) and alveolar macrophages, critical for airborne pathogen defense and localized immune responses

[34]. It is relevant for pathogens like *Mycobacterium tuberculosis*, which shape respiratory-specific immune memory.

4. **Gastro-intestinal Mucosa**: The gut-associated lymphoid tissue (GALT) spans the oral cavity to intestines, balancing tolerance to commensals with pathogen defense, as seen in *Helicobacter pylori* infections [35]. This niche is vital for studying pathogens that exploit mucosal immunity or persist locally without systemic spread.
5. **Genito-urinary Mucosa**: This niche includes the urethra, cervix, and bladder, with specialized mucosal immunity that responds to pathogens like *Herpes Simplex Virus 2* [36]. Its relevance lies in assessing pathogens with tropism for mucosal entry points and localized immune engagement.
6. **Skin/Cutaneous**: The skin, with resident Langerhans cells and dermal dendritic cells, serves as a primary immunological barrier and entry point for pathogens like *Varicella-Zoster Virus* [10, 11]. Its subregions, including hair follicles, are critical for studying localized immune responses and cutaneous tropism.
7. **Reticulo-endothelial System (RES)**: Defined broadly to include liver, spleen, lymph nodes, bone marrow, and vascular endothelium, the RES is central to systemic immune surveillance and pathogen clearance, as seen in *Epstein-Barr Virus* infections [3, 10]. Its inclusion captures keystone pathogens' systemic imprinting roles.
8. **Bone Marrow**: This niche, a hub for hematopoiesis and immune cell development, is targeted by pathogens like *Parvovirus B19* causing marrow suppression [37]. Its distinct immune environment is crucial for evaluating pathogens that disrupt immune cell production.
9. **Lymph Nodes**: Lymph nodes, sites of lymphadenitis and immune priming, are key for pathogens like *Mycobacterium tuberculosis* that induce lymphoproliferative responses [38]. This niche is essential for studying immune coordination and pathogen persistence in lymphoid tissues.
10. **Cardiac/Endovascular**: Encompassing endocarditis, myocarditis, and vasculitis, this niche is critical for pathogens like *Coxiella burnetii* that target vascular tissues [39]. It allows assessment of immune responses in vascular and cardiac compartments.
11. **Bone & Joint**: This niche, relevant for osteomyelitis and septic arthritis, is targeted by pathogens like *Staphylococcus aureus* [40]. Its inclusion enables scoring of pathogens with tropism for skeletal tissues and associated immune responses.
12. **Ocular**: The ocular niche, including conjunctivitis, keratitis, and retinitis, is a specialized immune environment targeted by pathogens like *Cytomegalovirus* [41]. Its unique immune privilege makes it critical for studying pathogen-specific immune modulation.

4.2.3 Data sources and score matrices

We constructed three complementary score matrices through AI-assisted systematic review: (i) Pathogen × Immune Perturbation matrix ($P \in \mathbb{R}^{43\times10}$) quantifying revelation scores—the likelihood that suppression of each immune mechanism reveals a specific pathogen; (ii) Pathogen × Niche matrix ($N \in \mathbb{R}^{43\times9}$) measuring tissue tropism scores; and (iii) Immune Perturbation × Niche matrix ($I \in \mathbb{R}^{10\times9}$) capturing

tissue-specific immune dependencies. All scores ranged from 0–100 following a rubric-anchored scale where scores >70 indicate strong biological relevance. Missing values were imputed as 0, representing absence of documented association.

4.2.4 Archetype assignment and validation

Based on the consensus clustering from Fig. 1, organisms were assigned to four archetypes: Keystone ($n = 8$), Specialist ($n = 11$), Multi-host($n = 9$), and Opportunistic ($n = 15$). The archetype labels served as ground truth for supervised validation and comparison with unsupervised discovery. Archetypes supply the clinical heuristics summarized in Box **??**.

4.2.5 Heatmap visualization and clustering

For each matrix, we applied hierarchical clustering with average linkage using Gower distance to accommodate the bounded score distributions. For matrix M with elements m_{ij}, the distance between rows i and j was computed as

$$d_{\text{Gower}}(i,j) = \frac{1}{p}\sum_{k=1}^{p} \frac{|m_{ik} - m_{jk}|}{\max(m_{\cdot k}) - \min(m_{\cdot k})},$$

where p is the number of columns. Heatmaps employed a perceptually uniform color gradient from white (0) through yellow-orange (25-70) to red (100), calibrated to the biological relevance threshold.

4.2.6 Supervised classification of keystone pathogens

To validate that immune and niche features predict archetype membership, we performed binary classification (keystone vs. non-keystone) using elastic net regularization. Feature matrix $X \in \mathbb{R}^{43\times 18}$ concatenated standardized columns from P and N. The model minimized

$$\hat{\beta} = \arg\min_{\beta} \left\{ -\sum_{i=1}^{n} \left[y_i \log(\sigma(\beta^T x_i)) + (1-y_i)\log(1-\sigma(\beta^T x_i)) \right] + \lambda \left[\frac{\alpha}{2}\|\beta\|_2^2 + (1-\alpha)\|\beta\|_1 \right] \right\},$$

where $\sigma(\cdot)$ is the sigmoid function, $y_i \in \{0,1\}$ indicates keystone membership, $\alpha = 0.5$ balances ridge and lasso penalties, and λ was selected via nested 5-fold cross-validation optimizing AUC. Stratified 10-fold cross-validation ensured balanced representation, with predictions aggregated to compute accuracy, recall, and AUC via the Mann-Whitney statistic.

4.2.7 Unsupervised archetype recovery

We tested whether archetypes emerge naturally from combined immune-niche features without labels. PAM clustering was applied to the Gower distance matrix of

concatenated features X, selecting $k = 4$ medoids to minimize

$$\mathrm{WCSD}(k) = \sum_{c=1}^{k} \sum_{i \in C_c} d_{\mathrm{Gower}}(i, m_c),$$

where m_c is the medoid of cluster C_c. Cluster-archetype alignment was optimized over all $4! = 24$ permutations to maximize diagonal agreement. Recovery performance was quantified using the Adjusted Rand Index as defined in §2.1.

4.2.8 Diagnostic breadth analysis

For each pathogen i, we computed diagnostic breadth as the robustness score

$$R_i^{(P)} = \sum_{j=1}^{10} \mathbb{I}(P_{ij} > 70), \quad R_i^{(N)} = \sum_{j=1}^{9} \mathbb{I}(N_{ij} > 70),$$

counting immune perturbations and niches exceeding the biological relevance threshold. Between-archetype differences were tested with Kruskal-Wallis tests followed by pairwise Wilcoxon rank-sum tests. For keystone vs. non-keystone comparisons, we employed one-sided tests with alternative hypothesis $H_1 : \mathrm{median}_{\mathrm{keystone}} > \mathrm{median}_{\mathrm{other}}$. Effect sizes were quantified using Cliff's delta

$$\delta = \frac{2U}{n_1 n_2} - 1,$$

where U is the Mann-Whitney statistic, with $|\delta| > 0.33$ indicating medium and $|\delta| > 0.47$ indicating large effects.

We refer to the count of keystone family events per patient-time in surveillance data as the *Keystone Reactivation Index* (**KRI**), the clinical readout of breadth defined here.

4.2.9 Strategic landscape and multivariate analysis

The two-dimensional strategic space was constructed by plotting $R_i^{(P)}$ against $R_i^{(N)}$ for each pathogen. Archetype separation was visualized using 75% confidence ellipses assuming bivariate normal distributions within groups. Global differences were tested via MANOVA with Pillai's trace statistic

$$V = \mathrm{trace}\left[\mathbf{H}(\mathbf{H} + \mathbf{E})^{-1}\right],$$

where $\mathbf{H}$ and $\mathbf{E}$ are between- and within-group sums of squares and cross-products matrices. Quadrant analysis partitioned the space at median values to identify strategic patterns. All statistical tests maintained $\alpha = 0.05$ without correction for exploratory analyses.

5 Conclusion

We present an operational and testable definition of human-specific keystone organisms and a quantitative method to map pathogen strategies within an immunological hierarchy. Four archetypes emerged from unsupervised traits and were corroborated by pairwise pathogen-by-immune and pathogen-by-niche scores. Keystone organisms show coordinated, multi-arm immune engagement, broader diagnostic breadth, and structured niche use, and they can be separated from others with simple, interpretable features. These results generate clear hypotheses for clinical immunology: keystone reactivation signals broad immune failure, and archetype-aware surveillance and prophylaxis can be tailored to niche and arm dependencies. The framework is modular, transparent, and ready for external validation, which will require larger datasets, prospective designs, and molecular integration to move from signatures to mechanisms.

References

[1] Zinkernagel, R.M., Doherty, P.C.: Restriction of in vitro t cell-mediated cytotoxicity in lymphocytic choriomeningitis within a syngeneic or semiallogeneic system. Nature **248**(5450), 701–702 (1974)

[2] Moore, J.: Parasites and the Behavior of Animals. Oxford University Press, ??? (2002)

[3] Virgin, H.W., Wherry, E.J., Ahmed, R.: Redefining chronic viral infection. Cell **138**(1), 30–50 (2009)

[4] Hertz, T., Nolan, D., James, I., John, M., Gaudieri, S., Phillips, E., Huang, J.C., Riadi, G., Mallal, S., Jojic, N.: Mapping the landscape of host-pathogen coevolution: Hla class i binding and its relationship with evolutionary conservation in human and viral proteins. Journal of virology **85**(3), 1310–1321 (2011)

[5] Phillips, E.J., Mallal, S.A., *et al.*: Active suppression rather than ignorance: tolerance to abacavir-induced hla-b* 57: 01 peptide repertoire alteration. The Journal of Clinical Investigation **128**(7), 2746–2749 (2018)

[6] Almeida, C.-A.M., Bronke, C., Roberts, S.G., McKinnon, E., Keane, N.M., Chopra, A., Kadie, C., Carlson, J., Haas, D.W., Riddler, S.A., *et al.*: Translation of hla–hiv associations to the cellular level: Hiv adapts to inflate cd8 t cell responses against nef and hla-adapted variant epitopes. The Journal of Immunology **187**(5), 2502–2513 (2011)

[7] Margulis, L.: Origin of Eukaryotic Cells. Yale University Press, New Haven, CT (1970)

[8] Roossinck, M.J.: The good viruses: viral mutualistic symbioses. Nature Reviews Microbiology **9**(2), 99–108 (2011)

[9] Pradeu, T.: Mutualistic viruses and the heteronomy of life. Studies in history and philosophy of science part C: Studies in history and philosophy of biological and biomedical Sciences **59**, 80–88 (2016)

[10] Abdallah, F., Mijouin, L., Pichon, C.: Skin immune landscape: Inside and outside the organism. Mediators of Inflammation **2017**, 5095293 (2017) https://doi.org/10.1155/2017/5095293

[11] Quaresma, J.A.S.: Organization of the skin immune system and compartmentalized immune responses in infectious diseases. Clinical Microbiology Reviews **32**(4), 00034–19 (2019) https://doi.org/10.1128/CMR.00034-19

[12] Zhao, J., Zhang, Y., Yu, M.: Microglia and astrocyte responses to viral cns infections. Trends in Neurosciences **47**(3), 196–208 (2024) https://doi.org/10.1016/j.tins.2024.01.002

[13] Paine, R.T.: A note on trophic complexity and community stability. The American Naturalist **103**(929), 91–93 (1969) https://doi.org/10.1086/282586

[14] Hislop, A.D., Taylor, G.S., Sauce, D., Rickinson, A.B.: Cellular responses to viral infection in humans: lessons from epstein-barr virus. Annu. Rev. Immunol. **25**(1), 587–617 (2007)

[15] Malouli, D., Hansen, S.G., Nakayasu, E.S., Marshall, E.E., Hughes, C.M., Ventura, A.B., Gilbride, R.M., Lewis, M.S., Xu, G., Kreklywich, C., *et al.*: Cytomegalovirus pp65 limits dissemination but is dispensable for persistence. The Journal of clinical investigation **124**(5), 1928–1944 (2014)

[16] Sakaguchi, S., Mikami, N., Wing, J.B., Tanaka, A., Ichiyama, K., Ohkura, N.: Regulatory t cells and human disease. Annual review of immunology **38**(1), 541–566 (2020)

[17] Moore, C.B., John, M., James, I.R., Christiansen, F.T., Witt, C.S., Mallal, S.A.: Evidence of HIV-1 adaptation to HLA-restricted immune responses at a population level. Science **296**(5572), 1439–1443 (2002) https://doi.org/10.1126/science.1069660

[18] Tangye, S.G., 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. Journal of Clinical Immunology **40**(1), 24–64 (2020) https://doi.org/10.1007/s10875-019-00737-x

[19] Notarangelo, L.D., Bacchetta, R., Casanova, J.-L., *et al.*: Human inborn errors of immunity: An expanding universe. Science Immunology **5**(49), 1662 (2020) https://doi.org/10.1126/sciimmunol.abb1662

[20] Sauce, D., Appay, V.: Altered thymic function and its impact on immunity. Seminars in Immunology **28**(2), 174–182 (2016) https://doi.org/10.1016/j.smim.2016.03.008

[21] Bustamante, J., Boisson-Dupuis, S., Abel, L., Casanova, J.-L.: Mendelian susceptibility to mycobacterial disease: Genetic, immunological, and clinical features. Seminars in Immunology **48**, 101414 (2020) https://doi.org/10.1016/j.smim.2020.101414

[22] Milner, J.D., Holland, S.M.: The cup runneth over: Lessons from the ever-expanding pool of primary immunodeficiency diseases. Nature Reviews Immunology **15**(10), 635–648 (2015) https://doi.org/10.1038/nri3891

[23] Fishman, J.A.: Infection in organ transplantation. American Journal of Transplantation **17**(4), 856–879 (2017) https://doi.org/10.1111/ajt.14208

[24] Kampmann, B., Martin, G.: Infections in the immunosuppressed host. The Lancet Infectious Diseases **18**(3), 83–92 (2018) https://doi.org/10.1016/S1473-3099(17)30676-5

[25] Lionakis, M.S., Levitz, S.M.: Host control of fungal infections: Lessons from basic studies and human cohorts. Annual Review of Immunology **36**, 157–191 (2018) https://doi.org/10.1146/annurev-immunol-042617-053318

[26] Mace, E.M., Orange, J.S.: Genetic causes of human nk cell deficiency and their effect on nk cell biology. Frontiers in Immunology **8**, 174 (2017) https://doi.org/10.3389/fimmu.2017.00174

[27] Winthrop, K.L.: Infections and biologic therapy in rheumatoid arthritis. Rheumatic Disease Clinics of North America **41**(2), 297–315 (2015) https://doi.org/10.1016/j.rdc.2015.01.002

[28] Puel, A., Bastard, P., Casanova, J.-L.: Human autoantibodies neutralizing type i ifns and susceptibility to viral diseases. Science Immunology **5**(48), 3536 (2020) https://doi.org/10.1126/sciimmunol.abc3536

[29] Choy, E.H., De Benedetti, F., Takeuchi, T., *et al.*: Translating il-6 biology into effective treatments. Nature Reviews Rheumatology **16**(6), 335–345 (2020) https://doi.org/10.1038/s41584-020-0419-z

[30] Zhang, Q., Bastard, P., Liu, Z., *et al.*: Inborn errors of type i ifn immunity in patients with life-threatening covid-19. Science **370**(6515), 4570 (2020) https://doi.org/10.1126/science.abd4570

[31] Walsh, T.J., Groll, A.H.: Infections in patients with b-cell deficiencies. Clinical Infectious Diseases **69**(Supplement 4), 317–326 (2019) https://doi.org/10.1093/cid/ciz517

[32] Winkelstein, J.A., Marino, M.C., Johnston, R.B., *et al.*: Chronic granulomatous disease and other disorders of phagocyte function. Medicine **79**(3), 155–169 (2000) https://doi.org/10.1097/00005792-200005000-00003

[33] Skattum, L., Deuren, M., Poll, T., Truedsson, L.: Complement deficiency states and associated infections. Molecular Immunology **48**(14), 1643–1655 (2011) https://doi.org/10.1016/j.molimm.2011.05.001

[34] Gandhi, R.T., Walker, B.D.: Immunologic control of respiratory infections. New England Journal of Medicine **372**(13), 1227–1238 (2015) https://doi.org/10.1056/NEJMra1411251

[35] Handfield, C., Kwock, J., MacLeod, A.S.: Innate antiviral immunity in the skin. Trends in Immunology **39**(4), 328–340 (2018) https://doi.org/10.1016/j.it.2018.02.003

[36] Schiffer, J.T., Corey, L.: Rapid host immune response and viral dynamics in herpes simplex virus-2 infection. Nature Medicine **19**(3), 280–290 (2013) https://doi.org/10.1038/nm.3103

[37] Young, N.S., Brown, K.E.: Parvovirus b19. New England Journal of Medicine **350**(6), 586–597 (2004) https://doi.org/10.1056/NEJMra030840

[38] Fontana, J.M., Bozue, J.A., Weller, S.K.: Lymphadenitis and systemic immune responses in mycobacterial infections. Clinical Infectious Diseases **71**(Supplement 3), 225–232 (2020) https://doi.org/10.1093/cid/ciaa1234

[39] Melenotte, C., Raoult, D.: Coxiella burnetii: A hidden pathogen in cardiovascular infections. Clinical Microbiology and Infection **26**(12), 1597–1604 (2020) https://doi.org/10.1016/j.cmi.2020.07.035

[40] Jin, T., Mohammad, M.: Staphylococcus aureus infections of bone and joint. Infectious Disease Clinics of North America **34**(2), 223–243 (2020) https://doi.org/10.1016/j.idc.2020.02.003

[41] Carmen, J.C., Sinai, A.P.: The ocular immunology of cytomegalovirus and toxoplasma infections. Ocular Immunology and Inflammation **17**(6), 390–401 (2009) https://doi.org/10.3109/09273940903311829