

Review

How colonization bottlenecks, tissue niches, and transmission strategies shape protozoan infections

Dana A. May ^{1,3} Fatima Taha ^{2,3} Matthew A. Child ^{2,*} and Sarah E. Ewald ^{1,*}

Protozoan pathogens such as *Plasmodium* spp., *Leishmania* spp., *Toxoplasma gondii*, and *Trypanosoma* spp. are often associated with high-mortality, acute and chronic diseases of global health concern. For transmission and immune evasion, protozoans have evolved diverse strategies to interact with a range of host tissue environments. These interactions are linked to disease pathology, yet our understanding of the association between parasite colonization and host homeostatic disruption is limited. Recently developed techniques for cellular barcoding have the potential to uncover the biology regulating parasite transmission, dissemination, and the stability of infection. Understanding bottlenecks to infection and the *in vivo* tissue niches that facilitate chronic infection and spread has the potential to reveal new aspects of parasite biology.

Highlights

Stochastic bottlenecks encountered during infection can have a disproportionately large influence on successful parasite colonization.

The relationship between parasite infection of discrete tissue niches and long-term persistence or transmission is not fully understood.

Cellular barcoding combined with spatial proteomics and transcriptomics has the potential to connect gene expression phenotypes with infection outcomes during colonization.

Understanding protozoan population dynamics

Parasitic protozoans including *Plasmodium* spp., *Leishmania* spp., *T. gondii*, and *Trypanosoma* spp. cause the debilitating diseases of human global health concern ranging from malaria, leishmaniasis, toxoplasmosis, and sleeping sickness to Chagas' disease [1]. The success of these eukaryotic pathogens can be linked to their multifaceted life cycles, which include transmission between multiple eukaryotic or insect host species (Figure 1) [2]. To facilitate transmission, protozoan parasites have evolved strategies to traffic to tissue niches distal to the site of primary infection, such as the skin or small intestine [3,4]. Establishing residency in these distal sites is often linked to changes in parasite gene expression, parasite differentiation, and chronic infection of permissive tissue niches. For example, the conversion of *T. gondii* tachyzoites into bradyzoites within neuronal tissue and skeletal muscle, or *Trypanosoma brucei* and *Plasmodium falciparum* immune evasion via surface antigen variation [5–8]. Infection and inflammation in chronic niches like the cardiac muscle (*Trypanosoma cruzi*) or the brain (*T. gondii*) can have profound negative consequences for tissue homeostasis. To protect these tissues, elaborate host barrier structures have evolved to limit the potential for pathogen colonization while facilitating nutrient and immune cell accessibility [9]. The complexity of protozoan parasite life cycles, coupled with limited 'druggability' of many chronic tissue niches has also limited the success of therapeutic tools to combat protozoan infections [10]. The need for targeted treatment options and a deeper understanding of how parasite localization influences fitness has stimulated scientific interest in understanding how the host environment shapes the population dynamics of protozoan infections [1].

During an infection, parasite population dynamics are influenced by selective and stochastic pressures (Figure 1). Host **selective pressures** (see Glossary) can include nutrition, immune polarization regulated by host genotype, and environmental inputs like drug treatment or mutagen exposure. Distinct from these selective mechanisms are **stochastic bottlenecks**. These

¹Department of Microbiology, Immunology, and Cancer Biology at the Carter Immunology Center, University of Virginia School of Medicine, Charlottesville, VA 22908, USA

²Department of Life Sciences, Imperial College London, South Kensington Campus, London SW7 2AZ, UK

³These authors made equal contributions

*Correspondence: m.child@imperial.ac.uk (M.A. Child) and se2s@virginia.edu (S.E. Ewald).



typically have a physical basis, such as transmission bottlenecks or an endothelial barrier, that constricts the parasite population in a manner that is often independent of an individual parasite's genetic fitness in the previous environment [11]. Whole-genome CRISPR knockout screens have revealed parasite genes conferring fitness *in vivo*. This approach is now being combined with host genetic tools and drug treatments to understand the parasite biology that facilitates adaptation to host selective pressures [12–16]. By contrast, stochastic bottlenecks have been less studied. Mechanisms typically used by pathogens to overcome selective pressures, include recombination or horizontal gene transfer, and occur at the level of individual genomes. In this regard, stochastic population bottlenecks can have an outsized impact on the long-term population structure of a species by dramatically impacting population size and having nonselective impact upon total population genetic diversity. Appreciating the magnitude, location, and timing of stochastic bottlenecks should clue researchers into the infection niches where gene expression and host effector mechanisms are the most critical for the outcome of infection (Figure 1). This review explores the ways in which the obligate intracellular parasite *T. gondii* and the largely extracellular parasite *T. brucei* interact with host tissue barriers, how these barriers influence colonization, and the tools available to study parasite population dynamics within these sites.

Protozoan traversal of host barriers and tissue niche colonization

T. gondii can infect a remarkably wide range of euthermic intermediate hosts, with acute infection resulting in flu-like symptoms that typically resolve. However, chronic infection is thought to be lifelong and immune suppression can trigger parasite **recrudescence** and potentially life-threatening toxoplasmosis. *T. gondii* is orally infectious, acquired through the ingestion of meat containing bradyzoite tissue cysts, or food and water contaminated with oocysts shed from the feline definitive host (Figure 2) [17]. This obligate intracellular parasite invades the **intestinal epithelial barrier (IEB)** and can be visualized replicating within enterocytes as early as 5 days postinfection [18,19]. It is not known whether *T. gondii* infects enterocytes by traversing the F-actin-rich enterocyte brush border directly, by disrupting tight junctions and invading the basolateral side of these cells, or by first infecting M cells in the villi crypts [20,21]. While this is an important early niche for the parasite, epithelial damage and recognition of commensal microbiota plays an important role in the activation of a local, protective immune response [22–24].

In the first week after infection, *T. gondii* is found in stromal cells of the lamina propria, as well as infiltrating monocytes, neutrophils, and dendritic cells [25]. Activated innate immune cells can restrict *T. gondii* growth via cell-autonomous immunity, however, naive monocytes and dendritic cells play a critical role in trafficking *T. gondii* to other tissues, including lymph nodes, spleen, adipose depots, liver, lung, skeletal muscle, cardiac muscle, and the central nervous system (CNS) [26–28]. By 4 weeks postinfection, parasites are cleared to levels that fall below PCR or histological detection in most tissues other than the skeletal muscle, cardiac muscle, and CNS; however, human organ transplant data indicate that many tissues harbor sufficient parasite load to induce infection of *Toxoplasma*-negative transplant recipients, including lung, kidney, heart, and liver [29]. To access these tissues, *T. gondii* must traverse the tissue–vascular barrier which generally consists of endothelial cells in close apposition to pericytes which also interface with tissue stromal cells or fibroblasts (Figure 2) [30].

For most tissue sites the mechanism of trans-endothelial migration has not been explored, the exception being the **blood–brain barrier (BBB)**. The brain contains the highest chronic parasite load per gram of tissue [31], making it a critical site for transmission to the feline definitive host. Neural infection is associated with parasite differentiation into bradyzoite cysts, a program that is regulated epigenetically and transcriptionally, and associated with slower parasite growth

Glossary

Arachnoid barrier: the brain meninges barrier that regulates transport between the dura and subarachnoid space composed of tight junction expressing epithelial-like cells.

Blood–brain barrier (BBB): the barrier that regulates transport between the blood and brain parenchyma composed of endothelial cells connected by tight junctions, pericytes, and astrocyte projections.

Blood–cerebrospinal fluid barrier: the epithelial cell barrier of the choroid plexus (CNS) regulating small molecule transport between the vasculature and circulating cerebrospinal fluid.

Glycosylphosphatidylinositol (GPI): a post-translational modification where a phosphoglyceride is attached to the C terminus of a protein, anchoring it to the membrane.

Intestinal epithelial barrier (IEB): the gastrointestinal tract barrier that mediates selective absorption of nutrients and restricts pathogen entry consisting of mucus and epithelial cells connected by tight junctions.

Next-generation sequencing (NGS): massively parallel sequencing technology to determine the order of deoxyribonucleic acids or ribonucleic acids in genomes or targeted regions.

Recrudescence: the process of active parasite exit from a host cell facilitating cell-to-cell spread.

Selective pressures: external or environmental circumstances that shape an organism's ability to survive. Selective pressures lead to the stabilization of genome alterations (mutations) that benefit the fitness of an organism and its progeny over evolutionary time scales.

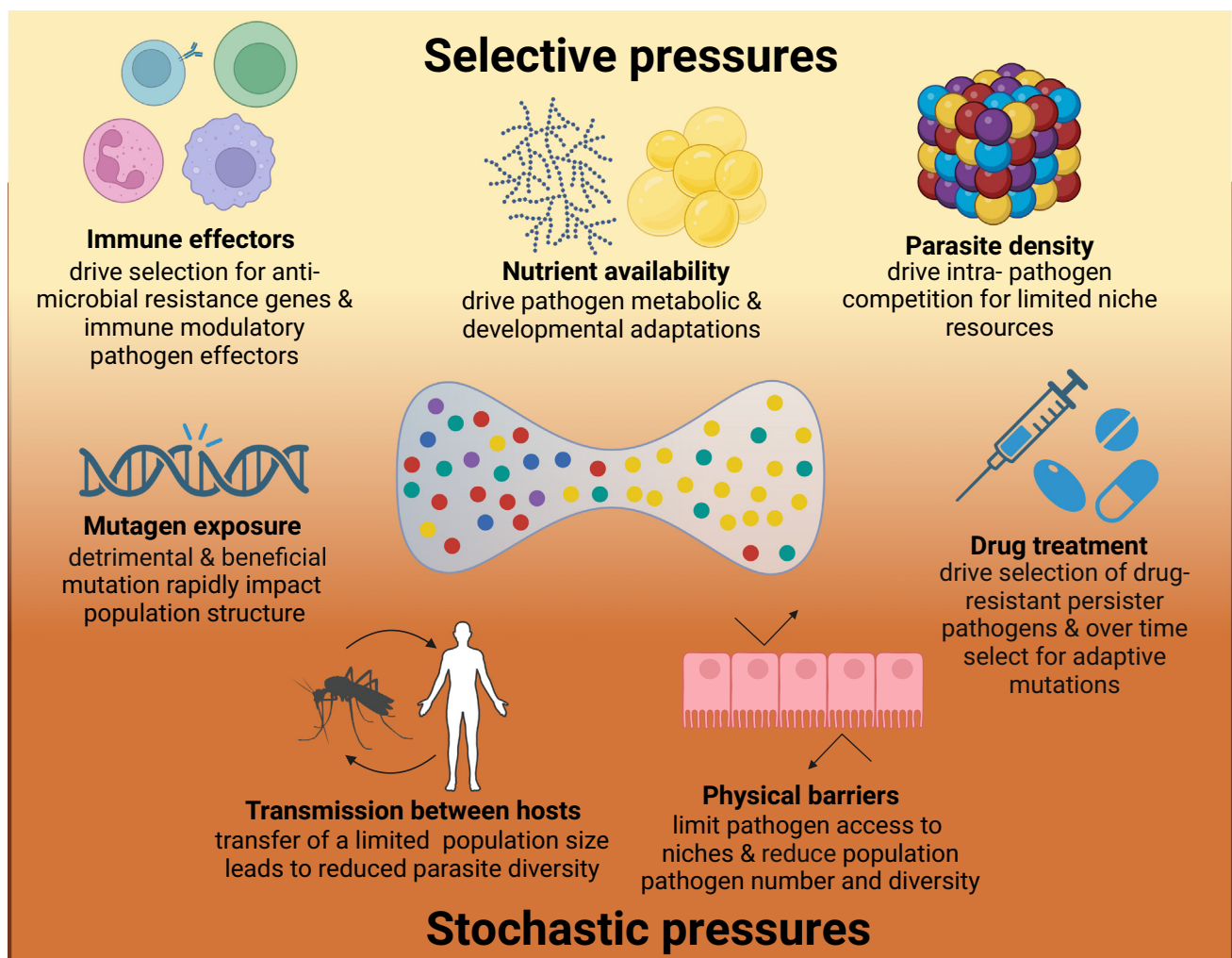
Sequence-tagged analysis of microbial populations (STAMP): a tool to assess microbial population bottlenecks and estimate founder population size based on measuring the relative frequency of neutral barcoded alleles across samples.

STAMPR: a successor to STAMP that utilizes an iterative barcode removal algorithm to consider the contribution of clonal expansion to bottleneck widths.

Stochastic bottlenecks: random events that limit the diversity of a population resulting in a limited pool of founder organisms often leading to genetic drift. Examples include physical barriers or limited transfer of organisms during pathogen transmission.

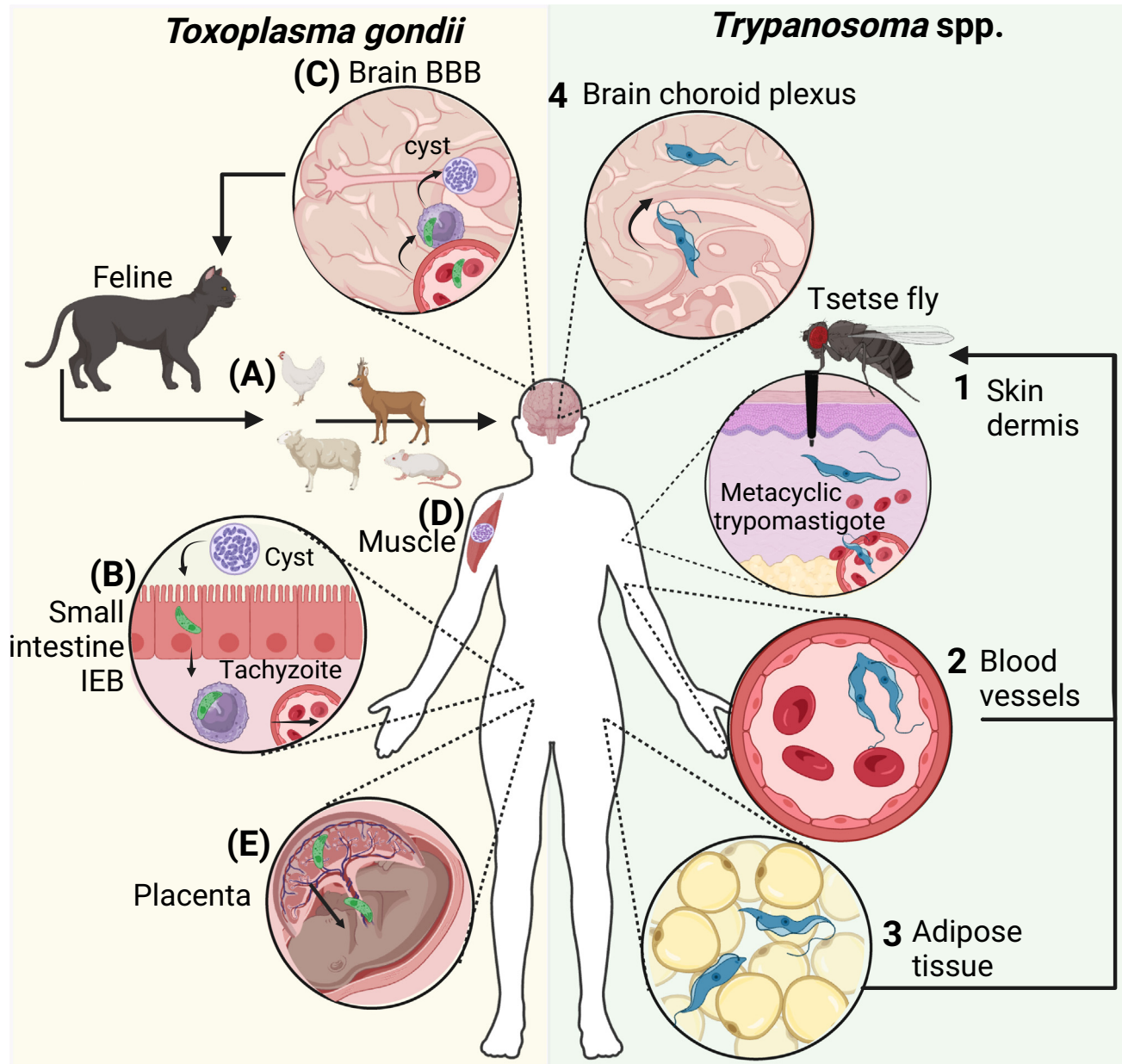
and immune evasion [32]. Immune suppression can lead to severe visual impairment, brain damage, and death [31]. At the BBB, three mechanisms of parasite entry have been proposed: transcellular infection of endothelial cells [33], paracellular trafficking of free parasites [26,34,35], or paracellular trafficking within extravasating immune cells (Figure 3) [36,37]. The precise anatomical site(s) of parasite entry into the brain and function of the **blood–cerebrospinal fluid barrier** and the meningeal **arachnoid barrier** are currently unknown [38,39]. Both CNS infection and gastrointestinal infection have been intense areas of interest, however, little is known about how effective anatomical barriers are at limiting parasite entry and stable niche colonization, or how they regulate dissemination to other tissues.

Variant surface glycoprotein (VSG): the major surface component of *T. brucei* mediating immune evasion. One of over 100 VSG-encoding genes is expressed at a time. As antibody responses are raised to one VSG, those parasites are cleared and clones arising from recombination and expression of a new VSG type emerge.



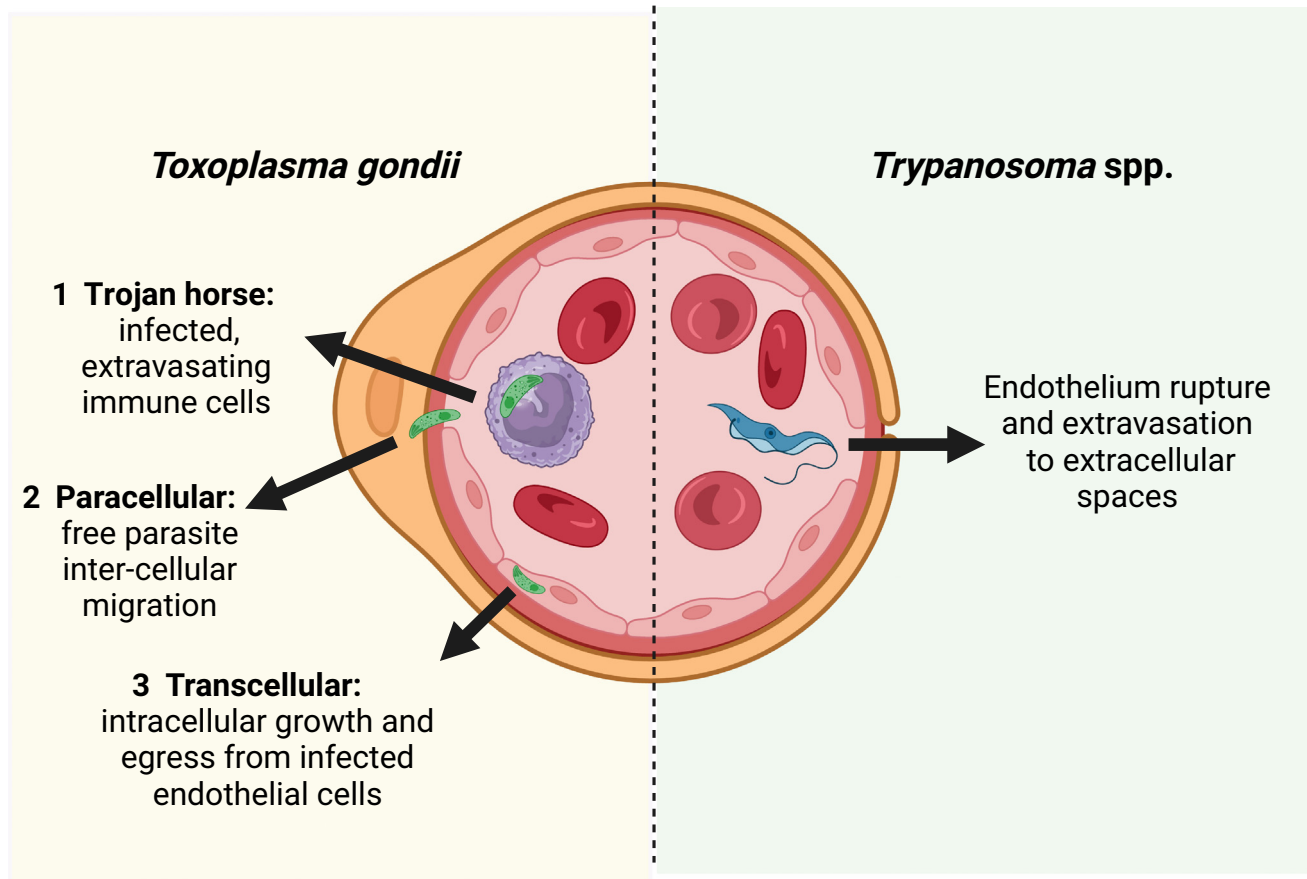
Trends in Parasitology

Figure 1. Selective pressures and stochastic bottlenecks shape the population structure of parasites during infection. Parasites frequently encounter stochastic bottlenecks upon transmission to a new host, these can include physical tissue barriers, or chance introduction of a fraction of a salivary gland population of parasites by the feeding behavior of a vector. This results in a limited founder population genetic pool that can have long-term consequences for adaptation to a new environment. Mutagen exposure or drug selection can lead to rapid loss of compromised individuals from the population or stabilization of genetic changes that confer fitness that can shape long-term population dynamics. By comparison, within-host selective pressures imparted by immune effectors, competition for nutrient availability operate on longer, evolutionary time scales. Competition between individual pathogens can drive selection of members within a population that can utilize distinct niches, energy sources, or means of transmission, for example. Figure created using BioRender.



Trends in Parasitology

Figure 2. *Toxoplasma gondii* and *Trypanosoma brucei* confront multiple dissemination barriers to establish infection in the host. (A) Felidae species support the sexual stage of *T. gondii* development, shedding oocysts into the environment. Warm-blooded intermediate hosts consume oocysts or bradyzoite tissue cysts. (B) Following ingestion, *T. gondii* invades the small intestinal epithelium. Parasites grow in endothelial cells and use immune cells as 'trojan horses' to access the vasculature and disseminate through the host. (C, D) *T. gondii* enters tissues through vascular barrier, including the blood-brain barrier (BBB), skeletal muscle, and cardiac muscle as the major sites of chronic infection and transmission. (E) *T. gondii* can be vertically transmitted to fetuses following traversal of the placental barrier. (1) *Trypanosoma* spp. are transmitted by bites of the tsetse fly. Trypomastigotes are transferred into the dermis as the fly pool feeds, disrupting blood and lymph vessels with its proboscis. (2) *Trypanosoma* migrates extracellularly to the vasculature then (3) disseminates into a range of tissues, including adipose depots, which may be sites of variant surface glycoprotein (VSG) recombination and chronic infection. Skin and blood colonization may facilitate transmission to the tsetse fly. (4) Late-stage *Trypanosoma* infection can lead to spinal fluid infection and colonization of the brain via the choroid plexus causing host-maladaptive, lethal inflammation. Abbreviation: IEB, intestinal epithelial barrier. Figure created using BioRender.



Trends in Parasitology

Figure 3. Models for transvascular dissemination of *Toxoplasma gondii* and *Trypanosoma brucei*. Three models have been proposed for *T. gondii* to traverse the vasculature. (1) Trojan horse: consists of infected immune cells directly transporting *T. gondii* into tissue niches during extravasation. (2) Paracellular: extracellular *T. gondii* migrates between endothelial cells that have been compromised or have increased permeability (vasodilation). (3) Transcellular: *T. gondii* infects and replicates within the endothelium then egresses directly into the tissue parenchyma. *T. brucei* is thought to traverse the endothelium by penetrating and subsequently rupturing the endothelial cells, then invading the extravascular space between cells. Figure created using BioRender.

T. brucei is an extracellular parasite transmitted by bites from tsetse flies, causing human African trypanosomiasis and nagana disease in livestock. Two main subspecies cause human African trypanosomiasis: *T. brucei rhodesiense* is associated with acute disease that lasts days to weeks and *T. brucei gambiense* is associated with chronic illness that lasts weeks to months. Tsetse flies are pool feeders, that use their proboscis to lacerate blood vessels and feed on blood and lymph fluid in the mammalian host. As a result, *T. brucei* is deposited in the dermal layer of the skin [40]. Parasites are observed in connective tissue associated with collagen bundles [41], which can develop into chancres as early as 2 days postinfection [42,43]. The first wave of macrophage, neutrophil, natural killer (NK) cell, and T cell infiltration, is critical to control parasite load. However, *T. brucei* can remain proliferative in the skin, eliciting a secondary neutrophil response and immunological environment that is beneficial for the parasite in the early stage of the infection [44]. During the later stages of the infection the skin contains quiescent, transmissible parasites which can differentiate into the insect procyclic form, supported by evidence where teneral tsetse flies fed on mice with low to undetectable parasitemia became infected [45]. Whether this population represents a tissue-resident pool from the initial infection or a secondary reseeding population from circulation is unclear. Together, these data indicate the dermal

environment can pose both a barrier to systemic colonization by *T. brucei* entry as well as a potential long-term reservoir for transmission.

Blood parasitemia can be detected as early as 1 day postinfection, marking the early hemolymphatic stage of trypanosomiasis. Hydrostatic pressure and chemokine sensing have been proposed to underpin parasite migration from the bite site to lymph and blood vessels although the precise mechanism is not clear [46]. This migratory event may represent an initial stochastic bottleneck that parasites must overcome to enter the circulation. As the infection progresses, *T. brucei* accesses the brain parenchyma, crossing the BBB via the choroid plexus (Figure 2) [47–49]. If untreated, parasite growth is poorly controlled causing meningoencephalitis and host death [43]. In contrast to *T. gondii* infection, the brain is not a transmission niche for *T. brucei* [47]. This indicates that the selective pressures shaping the entry of *T. gondii* and *T. brucei* into the brain are likely distinct. *T. gondii* must access the brain without triggering lethal inflammation. Hypervirulent strains that kill the host before encystation can occur but cannot be transmitted. This selective disadvantage is exemplified by the dominance of Type II strains in Europe and North America relative to hypervirulent Type I *T. gondii* [50]. In keeping with this model, CNS infection by *T. brucei* does not negatively impact the parasite's ability to be transmitted from the skin or bloodstream and frequently results in pathological inflammation [51,52]. *T. brucei* CNS infection also disrupts the host sleep–wake cycle which has been hypothesized to benefit *T. brucei* by giving the insect vector longer to feed and therefore increasing the likelihood of parasite transmission [53]. Historically, *T. brucei* was thought to traffic from the blood or lymphatic system directly to the CNS; however, recent studies have discovered that parasites cross the vascular and/or lymphatic barriers to persist within a range of tissues including adipose tissue (Figure 2) [54,55], heart [56], lung [57], ovaries [58], testes [59], and spleen [60]. Recent studies evaluating **variant surface glycoprotein (VSG)** diversity have revealed that extravascular niches may seed subsequent rounds of blood infection by clones that have recombined their VSG locus and are resistant to circulating antibodies raised against the primary infection VSG type [61]. The majority of parasites isolated from the blood are quiescent [62], suggesting that the replicative niche for VSG switching is extravascular [63]. Occupying these niches may facilitate host immune evasion and limit contact with therapeutic agents that cannot efficiently penetrate deep tissues. In addition to the skin, adipose tissue is colonized early in infection and contains a high load of proliferative parasites. Adipose parasites present a distinct morphology that is intermediate to the characteristic 'slender' and 'stumpy' forms found in blood (Figure 2), and it has been suggested that this unique morphology may facilitate *T. brucei* migration between cell–cell junctions [54,55,64]. Adipose-resident parasites also exhibit metabolic shifts that may reflect niche specific-nutrient availability and parasite persistence [55]. The levels of available nonesterified fatty acids are higher in the adipose tissue than in the blood [65]. Although *T. brucei* can synthesize lipids, it scavenges host lipids for several essential metabolic pathways including biosynthesis of **glycosylphosphatidylinositol (GPI)**, a membrane-anchoring moiety appended to procyclins and VSGs. To satisfy this major nutritional requirement, *T. brucei* has been shown to hijack the adipose T cell response to liberate fatty acids [66], a response that ultimately leads to adipose wasting and cachexia in chronically infected mice. Thus, adipose residency may benefit the parasite by providing both nutritional resources, and a tolerogenic immune environment for persistent infection. These emerging models highlight a complex interplay of selective and stochastic events, and underscore the impact that bottlenecks imparted by the infection discrete tissues have on the composition of parasites available for persistent infection transmission.

Tools to dissect stochastic bottlenecks in protozoan infections

Pooled mutant libraries have emerged as powerful tools to understand genes regulating parasite fitness. Barcoding is often applied in this context as a unique identifier for the gene conferring

loss-of-function or gain-of-function phenotypes. By design, these studies link some barcodes to fitness-conferring phenotypes which means that they cannot be used to assess stochastic bottlenecks in the population [13–16,67–69]. Furthermore, there is no assurance with these techniques that a parasite has only integrated a single barcoded cassette rather than multiple barcoded gene disruptions. Heterogeneous integration of multiple barcodes per parasite would undercut the utility of using each one as a unique identifier of a parasite lineage. Although an approach that pairs multiplexing with barcoded pooled-mutant libraries has recently been developed by the Lourido laboratory [15], this requires extremely high inoculum (lethal during acute infection) to obtain ample coverage or both multiplex identifiers and barcoded gene interruption cassettes.

For these reasons, the impact of stochastic bottlenecks have remained mostly unstudied in protozoan parasite infections. Progress within viral and prokaryotic infection disciplines has benefited from the cross-pollination of concepts and approaches typically employed by population geneticists. Lineage tracing methods, such as cellular barcoding, are notable examples, recently reviewed in relation to human disease by Sankaran and colleagues [70]. Cellular barcoding uses naturally occurring allelic variation or engineered molecular identifiers to ‘barcode’ cells of interest [71–73]. Experimentally introduced, unique selectable markers, sequences or tags can serve as ‘alleles’ to quantify changes in frequency within genetically complex populations. This has been used in lineage tracing [74,75], functional profiling [76–78], and spatiotemporal profiling of pathogen population dynamics (Table 1).

A recent advance has been the use of wild-type isogenic tagged strains (WITS). Here, a unique molecular identifier is inserted into a neutral locus therefore having no effect on cell fitness [79]. The WITS approach has been widely used in bacterial infection studies, such as dissecting stable niche colonization by *Salmonella* [80]. One limitation of the original methodologies relates to the small number of barcoded strains that can be identified, leading to coarse-grain understanding of stochastic pressures impacting within-post population dynamics. Next-generation WITS methods seek to overcome this limitation by combining a greater number of markers with **next-generation sequencing (NGS)** for quantitative analysis of highly complex barcoded strain libraries, and population genetic mathematical analyses. In principle, the higher the number of genetic markers (i.e., barcodes) used, the greater the resolution for discerning the width of the bottleneck [81]. Exemplifying this, **sequence-tagged analysis of microbial populations (STAMP)** and its successor, **STAMPR**, quantify the relative abundance of individuals in a tagged isogenic population [82,83]. The method applies mathematical equations derived from population genetic theory and NGS to accurately estimate bottleneck sizes within the host, while also providing information on the spatiotemporal dynamics. For these methods, the bottleneck width, or the founding population size (N_b), is calculated using initial sample size (i.e., number of sequences) and sample size at a given sampling time. It is derived from several population genetics approaches, including effective population size, N_e [84] and equations from Krimbas and Tsakas [85]. STAMPR improved upon the assumption that changes in allele frequency correspond to stochastic, homogenous movement of populations through a bottleneck. It incorporates considerations of complex colonization patterns such as repeated bottleneck events and heterogenous growth rates after a bottleneck event. As an alternative approach, the genetic relatedness of populations can be determined by calculating their genetic distance from one another, using the chord distance equations such those defined by Cavalli-Sforza [86]. Studying colonization using STAMP and STAMPR revealed complex dissemination patterns in a host in a temporal manner initially in bacterial species including *Vibrio cholerae* [83], *Salmonella* [87], extraintestinal *Escherichia coli* [82], and *Listeria monocytogenes* [88]. However, there are limitations to the effectiveness of this approach at low numbers of barcodes, which may be problematic when

Table 1. Studies investigating population dynamics using cellular barcoding *in vivo*^a

Organism	Barcoding technique	Integration via	Quantification method	Number of barcodes	Refs
<i>Borrelia burgdorferi</i>	WITS carrying two 20 bp DNA tag	Insertion into linear lp25 plasmid	PCR	7	[105]
<i>Escherichia coli</i>	WITS carrying 20 bp DNA tag (STAMPR)	Homologous recombination	NGS (Illumina)	~1100	[106]
	WITS carrying 40 bp DNA tag	Tn7 mediate integration	qPCR	10	[107]
<i>Listeria monocytogenes</i>	WITS carrying ~30 bp DNA tag (STAMP)	Phage integrase-mediated recombination	NGS (Illumina)	200	[88,108]
	WITS carrying 40bp DNA tag	Phage integrase-mediated recombination	qPCR	20	[109]
<i>Pseudomonas aeruginosa</i>	WITS carrying ~30 bp DNA tag (STAMP)	Integrase-mediated recombination	NGS (Illumina)	~4000	[110]
<i>Salmonella</i>	WITS carrying 40 bp DNA tag	Lambda-Red recombination	qPCR	8	[79,80, 111–113]
	WITS carrying 40 bp DNA tag	Lambda-Red recombination	rtqPCR	7	[114–117]
	WITS or MITS carrying 40 bp DNA tag	Lambda-Red recombination	NGS (Illumina)	8	[118,119]
	WITS carrying 21 bp DNA tag (STAMP)	Lambda-Red recombination	NGS (Illumina)	85	[87]
<i>Streptococcus pneumoniae</i>	Two strains carrying 'OVA' or 'AVO' peptides each	Janus-cassette mediated recombination	rtPCR	2	[120]
<i>Toxoplasma gondii</i>	WITS carrying 6 bp DNA tag	CRISPR-Cas9 mediated homologous recombination	NGS (Illumina)	96	[89]
	Wild-type RH strains carrying 8 bp UMI	CRISPR-Cas9 mediated homologous recombination	NGS (Illumina)	65 ± 29 per gRNA	[15]
<i>Trypanosoma brucei</i>	WITS carrying 6 bp DNA tag	CRISPR-Cas9 mediated homologous recombination	NGS (Illumina)	96	[89]
	Genetically homogenous cells carrying 40 bp DNA tag	Homologous recombination	NGS (Roche)	8	[90]
<i>Vibrio cholerae</i>	WITS carrying 30 bp DNA tag (STAMP)	Homologous recombination	NGS (Illumina)	~500	[83,121]
<i>Yersinia pseudotuberculosis</i>	WITS carrying 40 bp DNA tag	Transformation with plasmid containing tag	Southern blot	33	[122]

^aAbbreviations: gRNA, guide RNA; MITS, mutant isogenic tagged strains; rtqPCR, real-time quantitative PCR; UMI, unique molecular identifier.

studying host–pathogen interactions where a low inoculation dose of pathogen is required. The application of this approach to protozoan parasites has the potential to reveal dynamics of the population structure within a host.

Barcoding methods have only recently been applied to study population dynamics in protozoan infections. Wincott *et al.* developed an approach to label *T. gondii* or *T. brucei* with 96 unique DNA barcodes and monitor population dynamics in tissue culture and *in vivo* [89]. This study challenged the assumption that chronic infection of the CNS is established by a limited number of *T. gondii* entry events. Instead of identifying a clonal or near-clonal population in the brain, the majority of the barcodes identified in the inoculum were found in mouse brains one month after infection. Although this study was designed to introduce 96 barcodes at an equal ratio, natural variation in relative abundance was observed in the inoculum. Intriguingly, low abundance barcodes in the inoculum were able to become highly abundant in the brain at chronic infection, further supporting the stochastic nature of parasite access to the CNS and niche establishment [89]. However, host–parasite systems which tolerate a low inoculum pose a challenge for the application of population genetic tools designed for bacterial systems. For example, STAMPR requires approximately 500 barcodes and was designed to study pathogen systems where inocula are generally several logs greater in magnitude.

Most studies evaluating *T. brucei* colonization of extravascular spaces have relied on bioluminescent or fluorescent reporters to quantify parasite load. However, the spatiotemporal relations of the populations within each niche has been difficult to assess until recently. In one pioneering study, eight barcoded strains of *T. brucei* were used to assess the population dynamics of parasite transfer to tsetse flies, following a blood meal and transmission to the murine host [90]. This study indicated that both tsetse flies and mammalian bite recipients can be colonized by more than one founder allelic type of *T. brucei* [90]. As only eight barcoded strains *T. brucei* were used, it is likely that the increased number of cellular barcodes would reveal new aspects of host colonization. The approach recently developed by Wincott *et al.* allows for at least 96 unique barcodes to be simultaneously incorporated within a neutral locus, and should provide greater resolution of the magnitude of stochastic events impacting *T. brucei*'s within-host population dynamics during colonization [89]. These studies provide the foundations for exploring parasite diversity within discrete tissue niches, the inter-relationship of each niche over time, and host response mechanisms that regulate niche accessibility.

Concluding remarks

T. gondii and *T. brucei* survive within the host environment despite encountering immunological and physical barriers as seen in the IEB and BBB for *T. gondii*, and in the blood, skin, and BBB for *T. brucei*. Cellular barcoding approaches can facilitate quantification of the magnitude of the population bottleneck encountered at these barrier sites and how residence in protected tissue spaces regulates parasite virulence strategies are largely open questions (see [Outstanding questions](#)). Less is known about how this biology operates at sites like the placenta, which facilitates vertical transmission of *T. gondii* and *T. brucei*, and barrier sites in definitive hosts (feline species for *T. gondii* and the Tsetse fly for *T. brucei*). Such approaches can be used to understand the ecology and evolution of coinfection dynamics between parasite species or strains and the host environment [91]. The relevance of the scenario is nicely illustrated for *T. brucei*, where field samples indicate that the insect vector is typically coinfecting with different trypanosome species [92,93]. The use of independently barcoded populations of these parasites would allow for the coinfection dynamics to be ascertained with high spatiotemporal resolution, and provide insight into the impact of coinfections on parasite biology and disease pathology.

The insertion of unique sequences into the genome, a requisite for cellular barcoding, can now be readily and widely achieved with other protozoan parasites. Since 2015 *PlasmoGEM* has provided vectors for targeted introduction of barcoded alleles into the *Plasmodium berghei* genome [78]. More recently, a library of uniquely barcoded *P. falciparum* parasites was generated using the non-essential *prf3* locus and used to study the relationship between fitness and drug resistance [13]. A strategy for barcode-sequencing has also been developed for *Leishmania* spp. and applied to loss-of-function screens [77]. LeishGEdit incorporated a barcoding strategy into a functional screen for flagellar mutants, which could be adapted for within-host population genetic studies [67]. The method developed for the cellular barcoding of *T. gondii* and *T. brucei* [89], could potentially be applied to other trypanosome species such as *Trypanosoma congolense* where CRISPR-Cas9 genome editing was recently adapted [94].

The technical application of lineage tracing approaches to study host colonization by these parasites is still in its infancy. For example, lineage recorder strategies introduce Cas9-generated 'scars', either insertions or deletions, that allow the descendants of a particular cell to be determined [95]. Recorders add temporal information to population genetic studies, allowing new questions to be posed about infection. For example, when a host becomes immunocompromised during *T. gondii* infection, where does the recrudescence population originate from? Are recrudescence parasites mostly cleared or is this process important for maintaining a chronic infection? We currently lack therapeutic tools to clear parasite cysts. Understanding the biology that sustains the

Outstanding questions

What are the stochastic bottlenecks limiting population genetic diversity for parasitic protozoan infections during host colonization?

How do coinfections influence the within-host population dynamics of parasites?

Does recrudescence of *T. gondii* bradyzoites result in clonal infections?

Which host tissue niche contains the *T. brucei* population responsible for vascular reseeded and maintaining antigenic diversity?

Are parasite replication rates tuned to the specific nutrient environment of the colonized tissue niche, and is this unifying strategy underpinning persistence?

chronic population may lead to new intervention strategies for toxoplasmosis in the context of immune suppression or ocular toxoplasmosis? As generational counters, recorder-type systems would also provide insight into other basic aspects of parasite biology, such replication rate. Our current understanding of parasite replication kinetics is largely based on tissue culture infection. Many basic questions remain open including what is the *in vivo* growth rate? Is growth rate tuned to the specific nutritional state of a primary host cell type or niche environment – as would seem to be the case for adipose-residence *T. brucei* – and how does this then feed into ideas of persistence?

Technological advances in whole-tissue imaging (e.g., light sheet microscopy) [96], tools for multiparameter imaging *in situ* imaging (e.g., digital spatial profiling) [97], and spatial proteomics and transcriptomic tools (e.g., 10x/Visium, lightSeq, AutoSTOMP, DeepVisual Proteomics) [98–103] have the potential to enhance our understanding of the spatial distribution of host–pathogen interactions [104]. A recent study by Quintana *et al.* assessed localization, phenotype, and transcriptional signatures of *T. brucei* in the mouse brain relative to host cells using 10x/Visium spatial transcriptomics. They found that ‘slender’ and ‘stumpy’ morphology *T. brucei* were present in distinct brain regions and that neural infection led to the recruitment and expansion of follicular-like T cells in the brain parenchyma [7]. One limitation of tissue imaging and spatial-omic methods is that they typically assess the parasite as a homogeneous population without considering the relationships between individual parasites in space and as they disseminate over time. Merging cellular barcoding with imaging techniques has the potential to teach us about host–pathogen interactions. For example, if a niche is occupied by a single parasite clone does this make that site resistant or permissive to future colonization events? What are the parasite or local host gene expression programs that facilitate dominant infection by a clone or clearance of less successful lineages? How does abundance or localization of a parasite clone relate to successful transmission? Combining burgeoning technologies in this way has the potential to provide unique insight into potential colonization bottlenecks at host barrier sites, revealing how and where various pathogens enter tissue niches and the host and parasite gene expression programs that contribute to pathogen dissemination and/or clearance.

It is apparent that there remains much to be understood about how stochastic bottlenecks shape protozoan parasite infections, and hopefully with tools to quantify population dynamics, colonization, and tissue niche bottlenecks can finally be widened.

Acknowledgments

This work was supported by grants: National Institutes of Health (NIH) T32AI007046-46 (D.M.), R21AI156153 (S.E.E.), R35GM138381 (S.E.E.), Wellcome Trust and Royal Society 202553/Z/16/Z (to M.A.C.). F.T. is supported by a WP Studentship from the Department of Life Sciences, Imperial College, UK. For the purpose of open access, the author has applied a Creative Commons Attribution (CC BY) license to any Author Accepted Manuscript version arising.

Declaration of interests

The authors declare no conflict of interest.

References

1. Sutrave, S. and Richter, M.H. (2021) The Truman Show for protozoan parasites: a review of *in vitro* cultivation platforms. *PLoS Negl. Trop. Dis.* 15, e0009668
2. Yeager, R.G. (1996) Protozoa: structure, classification, growth, and development. In *Medical Microbiology*, Chapter 77, Available from www.ncbi.nlm.nih.gov/books/NBK8325/
3. Brescia, P. and Rescigno, M. (2021) The gut vascular barrier: a new player in the gut–liver–brain axis. *Trends Mol. Med.* 27, 844–855
4. Harris-Tryon, T.A. and Grice, E.A. (2022) Microbiota and maintenance of skin barrier function. *Science* 376, 940–945
5. Matta, S.K. *et al.* (2021) *Toxoplasma gondii* infection and its implications within the central nervous system. *Nat. Rev. Microbiol.* 19, 467–480
6. Milne, K. *et al.* (2021) Mapping immune variation and var gene switching in naive hosts infected with *Plasmodium falciparum*. *Elife* 10, e62800
7. Quintana, J.F. *et al.* (2022) Single cell and spatial transcriptomic analyses reveal microglia-plasma cell crosstalk in the brain during *Trypanosoma brucei* infection. *Nat. Commun.* 13, 5752
8. So, J. *et al.* (2022) VSGs Expressed during natural *T. b. gambiense* infection exhibit extensive sequence divergence

111. Dybowski, R. *et al.* (2015) Single passage in mouse organs enhances the survival and spread of *Salmonella enterica*. *J. R. Soc. Interface* 12, 20150702
112. Gül, E. *et al.* (2023) Intraluminal neutrophils limit epithelium damage by reducing pathogen assault on intestinal epithelial cells during *Salmonella* gut infection. *PLoS Pathog.* 19, e1011235
113. Bakkeren, E. *et al.* (2019) *Salmonella* persisters promote the spread of antibiotic resistance plasmids in the gut. *Nature* 573, 276–280
114. Maier, L. *et al.* (2014) Granulocytes impose a tight bottleneck upon the gut luminal pathogen population during *Salmonella* Typhimurium colitis. *PLoS Pathog.* 10, e1004557
115. Kaiser, P. *et al.* (2013) Lymph node colonization dynamics after oral *Salmonella* Typhimurium infection in mice. *PLoS Pathog.* 9, e1003532
116. Kaiser, P. *et al.* (2014) Cecum lymph node dendritic cells harbor slow-growing bacteria phenotypically tolerant to antibiotic treatment. *PLoS Biol.* 12, e1001793
117. Hausmann, A. *et al.* (2020) Intestinal epithelial NAIP/NLRC4 restricts systemic dissemination of the adapted pathogen *Salmonella* Typhimurium due to site-specific bacterial PAMP expression. *Mucosal Immunol.* 13, 530–544
118. Rossi, O. *et al.* (2017) Within-host spatiotemporal dynamics of systemic *Salmonella* infection during and after antimicrobial treatment. *J. Antimicrob. Chemother.* 72, 3390–3397
119. Nguyen, B.D. *et al.* (2020) Import of aspartate and malate by DcuABC Drives H₂/fumarate respiration to promote initial *Salmonella* gut-lumen colonization in mice. *Cell Host Microbe* 27, 922–936.e6
120. Li, Y. *et al.* (2013) Within-host selection is limited by an effective population of *Streptococcus pneumoniae* during nasopharyngeal colonization. *Infect. Immun.* 81, 4534–4543
121. Gillman, A.N. *et al.* (2021) The infectious dose shapes *Vibrio cholerae* within-host dynamics. *mSystems* 6, e0065921
122. Barnes, P.D. *et al.* (2006) *Yersinia pseudotuberculosis* disseminates directly from a replicating bacterial pool in the intestine. *J. Exp. Med.* 203, 1591–1601