Bottling it all up: using parasite population dynamics to understand dissemination and identify susceptibility pathways in visceral leishmaniasis

Ciara Mary Loughrey

Doctor of Philosophy (PhD) in Medical Sciences

The University of Hull and The University of York

Hull York Medical School

December 2024

Abstract

Leishmania donovani causes the systemic multi-organ disease visceral leishmaniasis (VL). However, little is understood about the mechanisms controlling parasite dissemination between host tissues. Specifically, it is not known where or when parasite populations are reduced by immunological 'obstacles'; understanding this process could enable identification of new methods to block dissemination and limit clinical disease.

Here, we combined CRISPR genome editing with Sequence Tag Analysis of Microbial Population Dynamics (STAMP) and community ecology methods to determine bottlenecks and dissemination patterns using a C57BL/6J mouse model of VL. To do this, we generated an isogenic barcoded library of 102 *L. donovani* lines.

By calculating the 'Founder Population' by STAMP in multiple tissues at multiple time-points, we observed that permissiveness to parasite colonisation varied between tissues depending on life cycle stage (promastigote or amastigote) used for infection. Analysis of genetic distance between parasite populations identified the liver as a key source of parasites colonising other tissues at later time-points following promastigote infection.

To understand the mechanisms underlying early bottleneck events, we used a C3-deficient mouse model. The results revealed distinct variations in the role of complement for regulating parasite dissemination in male and female mice.

Furthermore, we developed a re-infection model, using the barcoded library alongside an *L. donovani* line expressing luciferase, to understand the population dynamics of primary and secondary VL infections. Our results suggest that protective immunity is tissue-specific, and that host cells may transport parasites to the liver, triggered by secondary infection.

Finally, analysis of skin revealed a patchy parasite distribution, with evidence of stochastic seeding from multiple source tissues; this work further indicated that the skin is protected from secondary infection establishment.

Overall, our findings provide insights into the mechanisms regulating population bottlenecks and colonisation dynamics, which could have important implications for parasite and host evolution and disease outcomes.

Contents

Abstract p. 2

Contents p. 3

List of Figures and Tables p. 8

Acknowledgements p. 16

Declaration p. 18

Chapter 1: Introduction p. 19

1.1 Leishmaniasis p. 19

1.1.1 Epidemiology and disease pathology p. 19

1.1.2 Risk Factors p. 20

1.1.2.1 Malnutrition p. 20

1.1.2.2 Co-infections and immunosuppression p. 21

1.1.2.3 Sex differences p. 22

1.1.2.4 Genetics and pre-exisiting conditions p. 23

1.2 *Leishmania* parasitology p. 24

1.2.1 *Leishmania* species p. 24

1.2.2 Life cycle of *Leishmania* p. 24

1.2.3 *Leishmania* genetics and genomics p. 28

1.2.3.1 Genetic manipulation of *Leishmania* p. 28

1.2.3.2 *Leishmania* gene expression p. 29

1.2.3.3 Genomic plasticity and instability p. 30

1.3 Immuno-biology of leishmaniasis p. 32

1.3.1 Host cell interactions p. 32

1.3.1.1 Macrophages, dendritic cells and monocytes p. 32

1.3.1.2 Neutrophils p. 34

1.3.1.3 Other putative host cells p. 35

1.3.2 Complement p. 36

1.3.3 Type 1 interferons (IFNs) p. 36

1.3.4 T cells and the Th1 response p. 37

1.3.5 B cells p. 39

1.3.6 Immune evasion and parasite persistence p. 42

1.4 Microbial ecology p. 43

1.4.1 Concepts of microbial community ecology p. 43

1.4.2 Microbial population dynamics *in vivo* p. 46

1.5 Project Aims p. 48

Chapter 2: Materials and Methods p. 49

2.1 Parasites p. 49

2.2 Culture media p. 50

2.3 Animals p. 50

2.3.1 Mouse strains p. 50

2.3.2 Infections p. 50

2.3.3 Serum re-constitution p. 51

2.3.4 *In vivo* Imaging System (IVIS) imaging of mice p. 51

2.3.5 Ethics p. 52

2.3.6 Exclusion of animals p. 52

2.4 Primers p. 52

2.5 Barcoded parasite library generation p. 54

2.6 PCR and Sanger sequencing validation of barcode insertion p. 55

2.7 *In vitro* infection of bone marrow macrophages p. 56

2.8 Macrophage lysosomal-associated membrane protein 1 (LAMP1) staining and analysis p. 56

2.9 Barseq DNA extraction, PCR amplification and sequencing p. 57

2.10 Tissue section staining p. 60

2.11 ELISA quantification of serum concentration of Type 1 IFN p. 60

2.12 Immuno-phenotyping of liver cells by flow cytometry p. 61

2.13 *In vitro* complement-mediated parasite lysis assay p. 63

2.14 Population dynamics analysis p. 63

2.14.1 Barcode extraction p. 63

2.14.2 Barcode processing p. 67

2.14.3 Sequence Tag-based Analysis of Microbial Population Dynamics (STAMP) bottleneck and genetic distance analysis p. 67

2.14.4 Simulation of bottlenecks for analysis of STAMP accuracy p. 68

2.14.5 Diversity analysis using Shannon Index p. 68

2.15 Minimum spanning tree (MST) network analysis p. 69

2.16 Assembly process analysis p. 69

2.17 Statistical analysis p. 71

2.18 Figure generation p. 71

2.19 Code and data availability p. 71

Chapter 3: Developing a Barseq-STAMP approach to understand *in vivo* bottlenecks in *L. donovani* infection p. 73

3.1 Introduction p. 73

3.2 Results p. 76

3.2.1 Strategy for library generation p. 76

3.2.2 Determination of required library size for bottleneck assessment p. 77

3.2.3 Barcode insertion p. 81

3.2.4 Validation of barcode neutrality *in vitro* p. 86

3.2.5 Validation of barcode neutrality in *in vitro* infection of macrophages p. 91

3.2.6 PILOT experiment in vivo to assess feasibility and optimise Barseq-PCR protocol p. 99

3.2.7 Barseq-STAMP metrics in PILOT feasibility experiment p. 103

3.3 Discussion p. 106

Chapter 4: Colonisation and dissemination of *L. donovani* in an amastigote and promastigote infected C57BL/6J mouse model of VL p. 112

4.1 Introduction p. 112

4.2 Results p. 119

4.2.1 Establishing a promastigote and amastigote *L. donovani* infection model in C57BL/6J mice p. 119

4.2.2 Parasites undergo a bottleneck to colonisation and the liver is an early colonisation site in both promastigote and amastigote VL infection models p. 122

4.2.3 Parasite life stage on infection impacts the relationship between bottleneck size and subsequent population diversity p. 128

4.2.4 Tissue-seeding dynamics differ depending on parasite stage on infection p. 136

4.2.5 Liver lobes show consistent colonisation and population diversity after promastigote infection p. 141

4.2.6 Network analysis reveals differential centrality of tissues in different model systems p. 144

4.2.7 Assembly processes shaping parasite populations *in vivo* differ between promastigote and amastigote infection models p. 152

4.3 Discussion p. 160

Chapter 5: The Role of Complement in Early Bottlenecks and Dissemination Patterns in the C57BL/6J model of VL p. 174

5.1 Introduction p. 174

5.2 Results p. 178

5.2.1 Complement plays a role in imposing bottlenecks to colonisation in the early stages of *L. donovani* infection of C57BL/6J mice p. 178

5.2.2 Impact of C3-deficiency on parasite relatedness between tissues p. 182

5.2.3 Mouse serum does not kill promastigote parasites *in vitro* p. 186

5.2.4 The role of complement in early infection bottlenecks is both tissue and sex dependent p. 191

5.2.5: Early *L. donovani* dissemination dynamics are more dependent on complement in male mice p. 195

5.2.7 Neutrophil and macrophage numbers in the liver after *L. donovani* infection are not affected by C3-deficiency p. 198

5.2.8 Systemic IFNα/β in response to *L. donovani* infection is complement-dependent p. 208

5.3 Discussion p. 210

Chapter 6: Tissue-specific population dynamics in a VL re-infection model p. 223

6.1 Introduction p. 223

6.2 Results p. 224

6.2.1 Establishing a re-infection VL model p. 224

6.2.2 Parasites are not cleared from tissues after 10 to 12 weeks of infection p. 227

6.2.3 Prior amastigote infection is protective against subsequent colonisation in the bone marrow, liver and spleen p. 230

6.2.4 Prior infection has little impact on parasite dissemination between tissues in subsequent infections p. 232

6.2.5 Primary infection is not re-activated by a secondary challenge p. 234

6.2.6 Evidence that latent parasites traffic back to the liver from other tissues after a subsequent infection occurs p. 240

6.2.7 Patchy parasite colonisation with low diversity characterises skin dissemination in an immune-competent mouse p. 244

6.2.8 Differences in host cell availability at 2 weeks post re-infection do not account for liver re-colonisation p. 252

6.3 Discussion p. 261

Chapter 7: General Discussion p. 272

References p. 284

Appendix p. 351

Abbreviations p. 395

List of figures and tables

Figure 1.1: *Leishmania* life cycle.

Figure 1.2: The adaptive immune response to VL infection is tissue-specific.

Figure 1.3: Microbial population dynamics.

Figure 2.1: Barseq PCR protocol for population dynamics analysis using barcode abundance.

Figure 2.2: Sequence reads output for PRO and AMA experimental sample sequencing using Barseq-STAMP PCR and Illumina sequencing.

Figure 3.1: Schematic illustration of the concept of genetic distance (GD).

Figure 3.2: Schematic illustrating the experimental methodology for using genetic barcodes in *L. donovani* to investigate parasite dissemination via population dynamics.

Figure 3.3: Simulation strategy for assessment of required library size.

Figure 3.4: Accuracy of STAMP-calculated FP for a range of simulated barcode library sizes.

Figure 3.5: Barcoded *L. donovani* library generation strategy.

Figure 3.6: Initial barcode insertion results using single and double selection.

Figure 3.7: Generation and validation of a pool of 102 isogenic barcoded *LV9 L. donovani* lines.

Figure 3.8: Validation of isogenic barcoded library growth *in vitro*.

Figure 3.9: Comparison of Gamma distribution to experimental data.

Figure 3.10: Schematic illustration of assays for validation of isogenic barcoded library infectivity *in vitro*.

Figure 3.11: Infection burden in BMDMs for five individual lines from the barcoded LV9 T7 cas9 *L. donovani* library.

Figure 3.12: Representative images of LAMP1 staining in BMDMs infected with promastigote *L. donovani* parasites for 1, 24 and 48 hours.

Figure 3.13: LAMP1 recruitment to parasite phagosome in response to infection with five individual barcoded *L. donovani* lines.

Figure 3.14: PILOT experiment using a 10 barcode pool in five 4 week infected BALB/c mice to validate Barseq-STAMP protocol.

Figure 3.15: PCR and Illumina sequencing of barcodes in five 4 week infected BALB/c mice.

Figure 3.16: Barseq-STAMP metrics for Founder Population (bottleneck stringency) and Genetic Distance (population relatedness) for tissues obtained from five 4 week infected BALB/c mice.

Figure 4.1: Natural infection versus artificial infection murine model for VL.

Figure 4.2: Illustration of the application of ecological analysis concepts to assess within-host pathogen dynamics in the context of a murine VL model.

Figure 4.3: Establishing a promastigote (PRO) and an amastigote (AMA) infection model for *L. donovani* infection of C57BL/6J mice.

Figure 4.4: Bottlenecks to colonisation in different tissues in C57BL/6J mice infected with amastigote and promastigote *L. donovani* parasites at 2 days, 14 days and 28 days post-infection.

Figure 4.5: Bottlenecks to colonisation in different tissues in C57BL/6J mice, split by sex after infection with amastigote and promastigote *L. donovani* parasites at 2 days, 14 days and 28 days post-infection.

Figure 4.6: Population diversity quantified by Shannon Index in different tissues in C57BL/6J mice infected with amastigote or promastigote *L. donovani* parasites at 2 days, 14 days and 28 days post-infection.

Figure 4.7: Population diversity quantified by Shannon Index in different tissues in C57BL/6J mice, split by sex, after infection with amastigote or promastigote *L. donovani* parasites at 2 days, 14 days and 28 days post-infection.

Figure 4.8: The relationship between bottleneck size and population diversity after infection with amastigote or promastigote *L. donovani* parasites.

Figure 4.9: Genetic distance analysis of amastigote or promastigote *L. donovani* infection parasite populations in different tissues in C57BL/6J mice at 2, 14 and 28 days p.i.

Figure 4.10: Genetic distances between tissue parasite populations for the average amastigote or promastigote *L. donovani* infected mice at 2, 14 or 28 days p.i.

Figure 4.11: Population dynamics within individual liver lobes of promastigote *L. donovani* infected mice at 28 days p.i.

Figure 4.12: Network analysis of tissues in amastigote or promastigote *L. donovani* infected mice to model parasite population relatedness via fitting of a minimum spanning tree (MST).

Figure 4.13: Connectedness of parasite populations in different tissues after amastigote or promastigote *L. donovani* infection of male and female C57BL/6J mice at 2, 14 and 28 days p.i.

Figure 4.14: Betweenness centrality of different tissues after amastigote or promastigote *L. donovani* infection of male and female C57BL/6J mice at 2, 14 and 28 days p.i.

Figure 4.15: Computational analysis of assembly processes in microbial community structuring.

Figure 4.16: Beta diversity of parasite populations in different tissues after amastigote *L. donovani* infection of male and female C57BL/6J mice at 2, 14 and 28 days p.i.

Figure 4.17: Beta diversity of parasite populations in different tissues after promastigote *L. donovani* infection of male and female C57BL/6J mice at 2, 14 and 28 days p.i.

Figure 4.18: Proportion of interactions dominated by each assembly process, grouped by tissue, after amastigote or promastigote *L. donovani* infection of male and female C57BL/6J mice.

Figure 4.19: Initial bottlenecks and their possible long term impacts in amastigote and promastigote *L. donovani* infection.

Figure 4.20: Proposed mechanisms of *L. donovani* parasite dissemination patterns in C57BL/6J mice.

Figure 4.21: Some key questions to address on dissemination in promastigote and amastigote *L. donovani* infection.

Figure 5.1: The complement system and its manipulation by *Leishmania* parasites.

Figure 5.2: Schematic of experimental plan to investigate the role of complement in early infection population dynamics of *L. donovani* in C57BL/6J mice.

Figure 5.3: Founder populations in multiple tissues in complement-deficient and wild-type C57BL/6J mice infected with promastigote *L. donovani* parasites at 2 days post-infection.

Figure 5.4: Genetic distance analysis for comparisons between multiple tissues in complement-deficient and wild-type C57BL/6J mice infected with promastigote *L. donovani* parasites at 2 days post-infection.

Figure 5.5: Genetic distance analysis for comparisons between multiple tissues in complement-deficient and wild-type C57BL/6J mice infected with promastigote *L. donovani* parasites at 2 days post-infection.

Figure 5.6: Experimental protocol to assess *L. donovani* parasite lysis by murine serum complement *in vitro*.

Figure 5.7: Flow gating strategy for assessment of live versus dead *L. donovani* promastigotes.

Figure 5.8: Murine serum fails to kill *L. donovani* promastigotes.

Figure 5.9: Live cell percentage using flow cytometry analysis of promastigote *L. donovani* parasite death after incubation with male and female C57BL/6J (WT) or B6.129S4-C3tm1Crr/J (C3KO) serum.

Figure 5.10: Schematic of experimental plan to investigate the role of complement in early infection population dynamics of *L. donovani* in C57BL/6J mice.

Figure 5.11: Founder populations in multiple tissues in complement-deficient (with or without serum 'addback') and wild-type C57BL/6J mice infected with promastigote *L. donovani* parasites at 2 days post-infection.

Figure 5.12: Genetic distance analysis for comparisons between multiple tissues in complement-deficient mice (with or without serum 'addback') and wild-type C57BL/6J mice infected with promastigote *L. donovani* parasites at 2 days post-infection.

Figure 5.13: Exemplar images of whole liver sections stained for Ly6G and F4/80 to quantify neutrophils and macrophages.

Figure 5.14: Exemplar staining of liver sections for Ly6G and F4/80 to quantify neutrophil and macrophage proportions in male and female C57BL/6J and B6124S-C3tm1Crr/J mice infected with *L. donovani* promastigotes at 2 days p.i.

Figure 5.15: QuPath analysis protocol for classification of Ly6G and F4/80 positive cells in *L. donovani* infected mouse livers.

Figure 5.16: Total neutrophils as a percentage of total cells in livers of male and female C57BL/6J and B6124S-C3tm1Crr/J mice infected with *L. donovani* parasites at 2 days p.i.

Figure 5.17: Total neutrophils expressing high or low-medium Ly6G as a percentage of total cells in livers of male and female C57BL/6J and B6124S-C3tm1Crr/J mice infected with *L. donovani* parasites at 2 days p.i.

Figure 5.18: Total macrophages as a percentage of total cells in livers of male and female C57BL/6J and B6124S-C3tm1Crr/J mice infected with *L. donovani* parasites at 2 days p.i.

Figure 5.19: Serum interferon-alpha/beta (IFN) expression in male and female C57BL/6J and B6124S-C3tm1Crr/J mice at 2 days post-infection infection with *L. donovani* parasites.

Figure 5.20: Understanding the role of complement in tissue colonisation by *L. donovani.*

Figure 6.1: Schematic of experimental plan to investigate the population dynamics of *L. donovani* in a re-infection model using C57BL/6J mice.

Figure 6.2: Exemplar *ex vivo* imaging of tissue parasite burden of initial amastigote infection in female C57BL/6J mice re-infected with promastigote *L. donovani* parasites after a prior amastigote infection.

Figure 6.3: Tissue parasite burden, quantified by *ex vivo* imaging, of initial amastigote infection in C57BL/6J mice re-infected with promastigote *L. donovani* parasites after a prior amastigote infection.

Figure 6.4: Founder populations in secondary infection parasite populations in multiple tissues in female C57BL/6J mice re-infected with promastigote *L. donovani* parasites after a prior amastigote infection.

Figure 6.5: Genetic distance analysis for comparisons in secondary infection parasite populations in multiple tissues in female C57BL/6J mice re-infected with promastigote *L. donovani* parasites after a prior amastigote infection.

Figure 6.6: Schematic of experimental plan to investigate the population dynamics of the primary infection of *L. donovani* in a re-infection model using C57BL/6J mice.

Figure 6.7: Exemplar *ex vivo* imaging of tissue parasite burden of secondary promastigote infection in female C57BL/6J mice re-infected with promastigote *L. donovani* parasites after a prior amastigote infection.

Figure 6.8: Tissue parasite burden, quantified by *ex vivo* imaging, of secondary promastigote infection in female C57BL/6J mice re-infected with promastigote *L. donovani* parasites after a prior amastigote infection.

Figure 6.9: Founder populations in the initial infection in multiple tissues in female C57BL/6J mice re-infected with promastigote *L. donovani* parasites after a prior amastigote infection.

Figure 6.10: Genetic distance analysis for comparisons of initial parasite populations in multiple tissues in female C57BL/6J mice re-infected with promastigote *L. donovani* parasites after a prior amastigote infection.

Figure 6.11: Patchy distribution of skin parasites in female C57BL/6J mice re-infected with promastigote *L. donovani* parasites after a prior amastigote infection.

Figure 6.12: Skin colonisation in female C57BL/6J mice re-infected with promastigote *L. donovani* parasites after a prior amastigote infection.

Figure 6.13: Skin parasite population relatedness to other tissues in female C57BL/6J mice re-infected with promastigote *L. donovani* parasites after a prior amastigote infection.

Figure 6.14: Gating strategy for flow cytometry analysis of liver immune cells in female C57BL/6J mice re-infected with promastigote *L. donovani* parasites after a prior amastigote infection.

Figure 6.15: Monocytes, macrophages and dendritic cells in the livers of in female C57BL/6J mice re-infected with promastigote *L. donovani* parasites after a prior amastigote infection.

Figure 6.16: Neutrophils in the livers of in female C57BL/6J mice re-infected with promastigote *L. donovani* parasites after a prior amastigote infection.

Figure 6.17: T cells, B cells and NK cells in the livers of in female C57BL/6J mice re-infected with promastigote *L. donovani* parasites after a prior amastigote infection.

Figure 6.18: The liver in a re-infection C57BL/6J VL model.

Figure 6.19: Skin trafficking proposal in a C57BL/6J VL model.

Figure 7.1: Illustration of colonising Founder Population and Genetic Distance to other tissues for the average female mouse 28 days post-infection with either amastigote or promastigote *L. donovani* parasites.

Figure 7.2: Future work to build on the findings presented.

Appendix Figure 1

Appendix Figure 2

Table 2.1: Primers used in this project.

Table 2.2: PCR reaction conditions for Barseq PCR 1 and 2.

Table 2.3: Markers, antibodies and channels for immuno-phenotyping by flow cytometry.

Appendix Table 1

Acknowledgements

I would like to dedicate this thesis to my Nana. Thank you for being the best grandmother in the world; I could not have done this without you.

Foremost, I would like to thank my supervisors, Paul Kaye and Jeremy Mottram, without whom this work would not have been possible. Thank you for your support, guidance and patience throughout the last four years. It has been a priviledge to learn from you both and I hope I can put the skills you have given me to good use in my future career.

I would like to thank everyone in the Mottram and Kaye labs (both past and present) for their help, advice, support and stimulating scientific discussion, as well as for providing training in various aspects of this project. Foremost, I would like to acknowledge Naj Brown, without whom I could not have completed this project – whilst also managing the lab, she has been instrumental in providing training, an extra pair of hands and a supportive ear for me throughout my PhD. Thank you for all of your help! Other notable mention to: Juliana Carnielli, for her assistance with the generation of the barcode library and the PCR optimisation process, as well as for being interested every time I bothered her with some new data; Helen Ashwin for her assistance with the animal work in this project, providing me with excellent training on various techniques, as well as for her help in the image analysis and tissue handling for flow cytometry; Shoumit Dey, for his patient guidance and discussion around aspects of data analysis and code writing, as well as for his training in flow cytometry; Charlotte Hughes, for her assistance with technical protocols for parasite transfections and for her sunny approach which can brighten stressful lab days; Natalie Prow, for her help with running ELISAs and her emotional support during tough lab and thesis writing days, and for all the trips to the gym/yoga to distress! Thanks also to Ed Muscutt, Sergios Antoniou, Rachel Neish, Romina Nievas, Nicola Baker, Nidhi Dey, Vincent Geoghegan, Jodie Dixon, Kat Van Bocxlaer, Damian Perez Mazliah, Marcela Montes de Oca, Becky Wiggins, Dan Holbrook, Elmarie Myburgh, Nath Jones, Eden Ramalho de Araujo, Mathieu Cayla, Joana Correia Faria, Mohamed Osman, Ewan Parry, Giulia Bandini, Jessie Zenga and

everyone else in the York parasite community and the York immunology community for their support, guidance and for being amazing colleagues and friends!

Special thanks also must go to Jon Pitchford, for his indispensable advice on mathematical modelling, his patient and enthusiastic discussions, and for his encouragement throughout my PhD, as well as for the generous offer of help when my bike broke down at Christmas! Thank you also to my Thesis Advisory Panel members, Nathalie Signoret and Heidi Baseler, for their brilliant support and advice throughout my project.

Other members of the York community who were vital for my training and for completion of this project include the staff in the BSF, who provided excellent training, advice and support on all aspects of animal work, and the genomics team within the TF - Lesley Gilbert, Sally James and Sam Donninger – who worked exceptionally hard to sequence some rather difficult samples throughout this project. My gratitude also goes to Alastair Droop, for his encouragement and support with data analysis, and who was vital for the development of the Barseq extraction scripts. He was also incredibly supportive generally and provided some highly interesting insights in more general discussions about science and data. Thanks also to the H block lab managers, Marie Labarthe-Last and James Fox, and to all stores, infrastructure, administration and cleaning staff, without whom none of this work could have been done. The students I had the pleasure to teach and interact with in the lab also deserve thanks – they provided motivation even when things were difficult, showing interest in my work and more general discussions, and writing some truly wonderful essays in response to some slightly left-field questions! Thank you all for your enthusiasm and engagement. I am also very grateful to the wider HYMS community and HYMS itself for their funding to complete this project.

I would also like to thank the previous supervisors who have invested their time and effort in training me. Specifically, Lilach Sheiner and Andrew MacClean, along with all other members of the Sheiner lab, were instrumental in inspiring me to pursue parasitology, as well as providing me with extensive feedback and support when applying for PhDs in this field. My supervisor from my time at GSK, Malini Menon, has also been a crucial part of my development as a scientist, encouraging me to

have confidence in my work, and supporting me through a very difficult personal circumstance.

I would like to thank my family, friends and especially my godmother, Jude, for their unwavering support throughout this process. Thank you for believing in me.

Finally, I would like to express my gratitude to the mice which were involved in this study; this work would not have been possible without them and as such, they deserve to be acknowledged here.

Declaration

I confirm that this work is original and that if any passage(s) or diagram(s) have been copied from academic papers, books, the internet or any other sources these are clearly identified by the use of quotation marks and the reference(s) is fully cited. I certify that, other than where indicated, this is my own work and does not breach the regulations of HYMS, the University of Hull or the University of York regarding plagiarism or academic conduct in examinations. I have read the HYMS Code of Practice on Academic Misconduct, and state that this piece of work is my own and does not contain any unacknowledged work from any other sources.

Chapter 1: Introduction

1.1 Leishmaniasis

1.1.1 Epidemiology and disease pathology

The leishmaniases are a collection of diseases caused by protozoan parasites of the genus *Leishmania*. There are currently around 53 described species, with 20 causing human disease (Akhoundi et al., 2016). Different species cause different diseases, ranging from self-healing, but disfiguring, cutaneous leishmaniasis (CL), to mucocutaneous leishmaniasis (MCL), which damages mucosal tissues, to the severe and often fatal visceral form of the disease (VL) (Kaye and Scott, 2011; Akhoundi et al., 2016). Additional rarer forms include diffuse cutaneous leishmaniasis and disseminated cutaneous leishmaniasis (Kaye et al., 2021). Post-kala azar dermal leishmaniasis (PKDL) is a complex condition involving the emergence of skin lesions following resolution of VL, and is considered a public health concern, due to its likely role in disease persistence in some communities (Zijlstra et al., 2016).

CL, the most common form, is widely distributed, with 0.6-1 million cases globally per year, across the Mediterranean, Middle East, Americas and Central Asia (WHO Factsheet, 2023). Whilst not usually life-threatening, CL infection can cause a significant psychosocial burden, with associated stigma against infected individuals (Bennis et al., 2018; Kaye et al., 2021; Wenning et al., 2022). MCL occurs only in South America (Akhoundi et al., 2016). VL affects an estimated 50,000 people per year, and is endemic in 56 countries (WHO Factsheet, 2023; Burza et al., 2018). The vast majority of cases occur in Sudan, South Sudan, Bangladesh, India, Brazil and Ethiopia (Torres-Guerrero et al., 2017). It is classified as one of the Neglected Tropical Diseases (NTDs), predominantly impacting poor communities.

In patients, CL usually presents as an ulcer in the skin, which can leave permanent scars, or as multiple ulcers, particularly in diffuse forms. MCL involves parasite metastasis to mucous membranes, with lymphatic involvement also often present; gradual destruction of the nasal or oral cavity can occur. In VL, presentation is defined by fever, anorexia and fatigue, with hepatosplenomegaly, anaemia,

leukopenia and thrombocytopenia. VL is almost always fatal if untreated, usually due to immune dysregulation and secondary infection. PKDL, evidenced by skin nodules or papules which can contain parasites, can be mild or severe, and this appears to relate to the region in which the patient acquired the infection (Pace, 2014; Torres-Guerrero et al., 2017).

Commonly, treatment requires intravenous pentavalent antimonials. Other drugs include the only oral medication currently approved, miltefosine, along with amphotericin B and paromomycin. However, problems exist for all drug regimens, including side effects, accessibility, treatment failure and growing drug resistance. New drugs are sorely needed, with immunotherapy posited as one promising avenue (Torres-Guerrero et al., 2017; Wijnant et al., 2022).

There is currently no approved vaccine; leishmaniasis vaccines have proven difficult to design and implement, due to many factors including difficulties identifying suitable antigens, lack of funding and limited understanding of the human immune response to infection (Kaye et al., 2021). Few candidates have reached clinical testing. Vaccine designs have included a live attenuated *L. major* immunisation approach (Karmakar et al., 2021; Volpedo et al., 2022) and an adenovirus vector vaccine with two *Leishmania* antigens (Younis et al., 2021; Younis et al., 2024). Recent reports of a protocol for controlled human challenge, with no major adverse effects reported in 14 healthy volunteers after *L. major* infection (Parkash et al., 2024), provide promising avenues to accelerate vaccine development.

1.1.2 Risk factors

There are various risk factors associated with both infection itself and disease severity; some clinically-relevant ones are briefly discussed below, and a range of environmental and sociological factors are reviewed in Valero and Uriarte, 2020. However, as an NTD, many underlying mechanisms for leishmaniasis pathology are currently poorly understood, and multiple risk factors are likely not identified at all.

1.1.2.1 Malnutrition

It is known that malnutrition on both a macro and micro-nutrient level can contribute to immune dysfunction, but the underlying mechanisms and thus possible therapeutic applications are not established for VL – the picture is further

complicated by the impact VL has on appetite and digestion of nutrients (Harrison et al., 1986). Protein-energy malnutrition is common in VL patients across many regions (Cerf et al., 1987; Harrison et al., 1986; Malafaia, 2009; Nweze et al., 2020; Rey et al., 2005), and is frequently associated with co-occurring intestinal parasite infection, whilst studies in Uganda and Ethiopia show almost 100% prevalence of anaemia among VL patients; it is not established whether this is due to a causative effect of iron metabolism on infection susceptibility or occurs as a result of VL infection (Mueller et al., 2009; Mengesha et al., 2014). Reduced zinc levels compared to controls were also observed in patients with VL in Ethiopia (Mishra et al., 2010),

Some experimental evidence exists to indicate that energy malnutrition and deficiencies in specific micronutrients, such as zinc and iron, results in more severe visceral disease for mice infected with *L. donovani* parasites, and that this may be due to defects in the macrophage Nitric Oxide (NO) production pathway (Anstead et al., 2003; Anstead et al., 2001; Nweze et al., 2020). Likewise, impaired immune responses to *L. chagasi* have been observed in BALB/c mice fed similarly deficient diets (Malafaia et al., 2009; Serafim et al., 2007). There are also indications of a role for vitamin D in the development of PKDL, possibly via macrophage activation (Mukhopadhyay et al., 2014).

1.1.2.2 Co-infections and immunosuppression

Co-occurrence of human immune-deficiency virus (HIV) and VL is common in many endemic areas and is a particular concern for patients and public health bodies; immunosuppression and dysfunction caused by both HIV and VL results in high relapse rates for those with both infections, with patients acting as reservoirs of disease in these regions (Takele et al., 2022). Evidence from a longitudinal study of VL-only and HIV-VL co-infected patients in Ethiopia have indicated that drug treatment causes a large reduction in parasite populations in both groups, without imposing drug-resistance selection on the remaining parasite population, and that re-activation of latent infection is the likely mechanism for relapse in immune-compromised patients, rather than a re-infection event (Franssen et al., 2021). In Brazil, VL patient studies have likewise identified increased risk of mortality in those infected with HIV (Druzian et al., 2015), and this link is consistent across other

regions in the Americas (Belo et al., 2014). Dysregulation of CD4+ and CD8+ T-cells is present in co-infected patients, and reduced interferon gamma (IFNγ) production, along with malnutrition, may play a significant role in relapse rates (Takele et al., 2022). The interplay of these two infections is complex and not fully established.

Evidence also suggests that co-infections with bacteria (Belo et al., 2014; Costa et al., 2010; Sampaio et al, 2010), including *Mycobacterium tuberculosis* (Druzian et al., 2015), can increase the severity of VL. In mice, co-infection with *Schistosoma* parasites can worsen control of VL infection (Hassan et al., 2006) and CL infection (reviewed in Camelo et al., 2023), whilst co-infected patients have a worse prognosis overall (Akuffo et al., 2018; Camelo et al., 2023); further research is needed to establish the underlying impact both parasites exert on each other and the immune response (Camelo et al., 2023).

1.1.2.3 Sex differences

Sex differences are a relatively unexplored area in terms of leishmaniasis susceptibility, but epidemiological evidence indicates that prevalence of disease and severity are both higher in males compared to females (Lockard et al., 2019; Torres-Guerrero et al., 2017). However, women have been shown to be less likely to engage in health seeking behaviour (Cloots et al., 2020; Wenning et al., 2022), so this may not represent true case rates. The extent to which differences in infection incidence and outcome are linked to biological differences is uncertain, since sociological effects are difficult to parse out, but evidence from other infections and immunological research suggests that differences in disease outcomes on the basis of sex are likely (Cloots et al., 2020; Klein and Flannagan, 2016).

Early studies in mice suggested a minor sex difference in control of early infection (Bradley et al., 1977). *In vivo* and *in vitro* studies have indicated that testosterone may promote replication of *Leishmania* parasites, whilst upregulating the Th2 response via interleukin 10 (IL-10) and tumour necrosis factor alpha (TNFα) expression in male mice and macrophages (de Araújo Albuquerque et al., 2021; Lockard et al., 2019; Rodriguez et al., 2018).

At the sociological level, research suggests women suffer greater stigmatisation and long-term impacts as a result of CL and MCL infections, although males are more

likely to engage with harmful treatment options. The brunt of caregiving for more severe cases or those impacting children tends to fall on women, as well as the responsibility for protecting the domestic environment against infection (Wenning et al., 2022).

1.1.2.4 Genetics and pre-existing conditions

Multiple genes have been implicated in leishmaniasis susceptibility and severity in both experimental and clinical studies (Bharati, 2022; Blackwell et al., 2009; McCall et al., 2013). Many are involved in the polarisation of immune responses towards Th1 or Th2 (Bharati, 2022; Blackwell et al., 2009). Patient studies show possible roles for *CXCR1* and *CXCR2* (Mehotra et al., 2011), *IL2* (Bucheton et al., 2007), *IL4* (Mohammed et al., 2003) and *TNF* (Karplus et al., 2002) for the development and outcome of VL infection in various regions. Deletion of the *Tnf* gene in C57BL/6 mice resulted in severe visceral disease when infected with *L. donovani* or *L. major* parasites, due to impairment of the early immune response (Murray et al., 2000; Wilhelm et al., 2001).

The *Nramp1* locus has been shown to play a major role in susceptibility to visceral leishmaniasis in mice (Malo et al., 1994), although the role this gene plays in human disease is less well-established (Bucheton et al., 2003). Specific polymorphisms appear to affect susceptibility to VL in various studies, particularly in individuals co-infected with HIV (Bharati, 2022). *NRAMP1* (since re-designated as *SLC11A1*) is involved in iron transport in macrophage phagolysosomes, with dysregulation impacting the ability to kill intracellular pathogens, including *Leishmania*, via NO production (Biggs et al., 2001; Blackwell et al., 2001; Blackwell et al., 2009). It has been suggested that this difference is due to human allele variants not resulting in a completely non-functional gene (Lipoldová and Demant, 2006). The *H2* locus is also known to play a secondary role in the ability to resolve disease in susceptible mice (Blackwell et al., 1980).

Some studies posit links between VL severity and comorbid conditions such as erythematosus systemic lupus and leukaemia (Druzian et al., 2015). Immunosuppression due to organ transplantation is also a risk for leishmaniasis infection (Akuffo et al., 2018).

1.2 *Leishmania* parasitology

1.2.1 *Leishmania* species

Leishmania species are members of the Trypansomatidae family, within the genus *Leishmania*, and can be split into five subgenera: *Leishmania*, *Sauroleishmania*, *Viannia*, *Leishmania enrietti* and *Paraleishmania*. However, there is some debate, based on new molecular analysis, contesting whether this system best reflects the true evolutionary relationships. The parasites have a dixenous life cycle, residing in both sand flies and vertebrate hosts; for most species, this host is mammalian, but species within *Sauroleishmania* infect lizards. Generally, parasites exhibit narrower preferences in terms of insect host, with a range of definitive hosts able to be infected by most species. Records of *Leishmania* infections date back to 2500BC, but the parasite is believed to have initially appeared as early as 65 million years ago, with reptiles likely acting as the early hosts (Akhoundi et al., 2016).

L. donovani and *L. infantum* parasites are both known to cause VL, but only *L. donovani* is reported to lead to PKDL development (Mukhopadhyay et al., 2014). They are believed to have diverged around 1 million years ago, segregating their distributions to India and Africa in the case of *L. donovani*, and the Americas and the Mediterranean basin in the case of *L. infantum* (Akhoundi et al., 2016).

1.2.2 Life cycle of *Leishmania*

Leishmania parasites are spread to susceptible hosts via the bite of the female phlebotomine sand fly. Within the sand fly, the parasites exist as promastigotes, but after entering mammalian host cells they differentiate into their amastigote form (Figure 1.1). Research indicates that there are actually four promastigote stages within the sand fly; amastigote parasites are initially taken up in a blood meal, differentiating into the dividing procyclic form. As they migrate through the sand fly, they become nectomonad and leptomonad parasites respectively, before finally differentiating into the infectious, non-dividing, metacyclic form. These metacyclics are the stage which is then transmitted to susceptible hosts via the skin (Gossage et al., 2003). Though initially thought to be terminally differentiated, recent evidence suggests that repeated blood meals can induce conversion of metacyclics into a

replicative ‘retroleptomonad’ stage, which serves to increase the number of the infective stages (Serafim et al., 2018). Multiple studies have indicated that sand fly saliva plays a role in the establishment of infection after biting, increasing host cell influx and decreasing the expression of cytokines involved in parasite killing, but also possessing properties which can immunise against infection (reviewed in Lestinova et al., 2017).

In endemic areas, infected people, including those initially cured patients who subsequently relapse, can act as a reservoir of disease, maintaining the cycle. As a zoonotic pathogen, *Leishmania* parasites can also live in many other mammals, with dogs heavily implicated in transmission in urban areas (Dantas-Torres, 2007; Jambulingam et al., 2017).

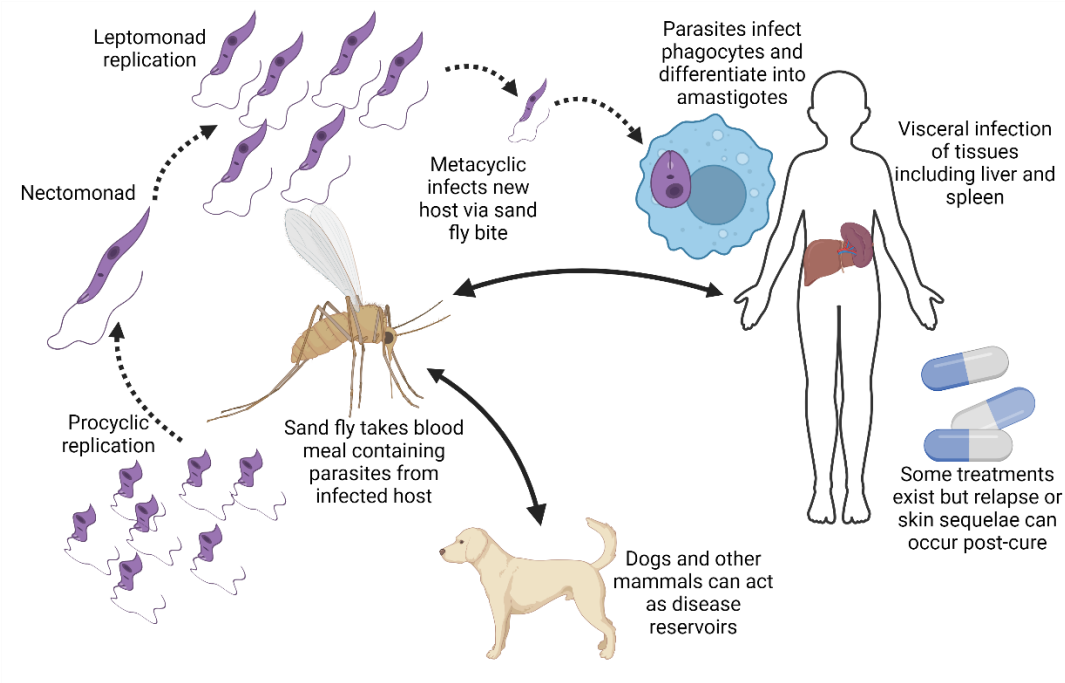


Figure 1.1: *Leishmania* life cycle. *Leishmania* parasites are spread between hosts by the sand fly vector. Within the sand fly they exist as promastigotes, but within this there are at least four distinct stages of development. When a sand fly bites a host, the sand fly metacyclic stage can infect, entering phagocytes such as macrophages and differentiating into the amastigote form. In CL, the infection remains localised to a skin lesion, whilst in VL, the parasites spread to other tissues such as the liver and spleen. Patients can be treated using various drug regimens, but relapse or post-cure skin sequelae in the form of PKDL can occur. These occurrences, along with zoonotic reservoirs, can maintain the infection cycle in endemic areas. Figure created in Biorender using information from Gossage et al., 2003.

Although metacyclic parasites are known to be more infective, and can be visually discriminated from procyclics by eye, more research is necessary to establish molecular techniques to analyse differences between these populations in detail. No specific molecular markers are currently known, and there are differences between gene expression profiles and infectivity of sand fly versus culture-derived metacyclic populations, as well as species-specific heterogeneity (Alcolea et al., 2014; Alcolea et al., 2019). Transcriptional analysis indicates that there are notable changes in gene expression profiles across the promastigote maturation process in the sand fly, indicative of pre-adaptation for amastigogenesis (Inbar et al., 2017).

The differences between amastigote and promastigote parasites are, in contrast, more established, though by no means definitively mapped. Multiple studies indicate that amastigotes are able to enter a quiescent state; dormancy is a common method microbes can use to reduce susceptibility to stressors such as drugs or limited resources (Barrett et al., 2019; Cabral et al., 2018; Ritterhaus et al., 2013). Differences in gene expression are well documented, with amastigotes generally displaying reduced expression compared to promastigote parasites. Alcolea et al (2014) compared sand fly-derived promastigote to amastigote *L. infantum* cultured *in vitro*, finding that genes related to stress responses, metabolism and translation are upregulated in promastigotes (Alcolea et al., 2014). However, all comparative analysis is caveated by the fact that differences in expression profile are likely to be species-specific, with possible other variations introduced by the use of different species of sand fly vectors or culture conditions. The comparison of gene expression in *L. major* and *L. infantum*, with their divergent tissue tropisms, exemplify this (Rochette et al., 2008); however, this study did not determine whether these differences persevere at the functional level of proteins.

Using heavy water labelling, *L. mexicana* parasites in mouse lesions have been found to reduce their metabolic rates and replication (Kloehn et al., 2015). Further work established that these quiescent parasites exist in specific spatial regions of granulomatous tissues, whilst healed lesions are less permeated by drug treatment, indicating a possible role for these quiescent parasites in persistence (Kloehn et al., 2021). Work by Dirkx et al (2024) has confirmed that quiescence also occurs in *L. donovani* and *L. infantum* infection, in the bone marrow of BALB/c mice.

Differences in a small number of surface expression markers between amastigote and promastigote *Leishmania* are well documented. A few notable examples are the amastin gene family, lipophosphoglycan (LPG) and GP63. The amastins are a gene family with homology to the related parasite *Trypanosoma cruzi*, and are expressed at higher levels on amastigote surface membranes (Wu et al., 2000). The expression and structure of LPG has been noted to change throughout differentiation, and plays roles in both interactions with host cells (Desjardins and Descoteaux, 1997; Desjardins and Descoteaux, 1998) and the sand fly vector (Mahoney et al., 1996; Sacks et al., 1995). It also varies structurally between

Leishmania species (Sacks et al., 1994). GP63, a metalloprotease, is involved in phagocytosis and survival of parasites within host cells (Matte et al., 2016; Russell and Wilhelm, 1986; Russell and Wright, 1988), and varies in structure and cellular localisation at different life stages (Hsiao et al., 2008; Medina-Acosta et al., 1989). Additionally, promastigotes can exocytose the protein, whilst amastigotes do not do this as readily, reflecting the differences in function for this molecule (Hsiao et al., 2008).

1.2.3 *Leishmania* genetics and genomics

1.2.3.1 Genetic manipulation of *Leishmania*

Many *Leishmania* species have had their genomes sequenced, and the functions of some genes are well established (Lye et al., 2010). However, understanding gene function in *Leishmania* has proven difficult, due to the specific features of the parasites' genetic regulatory processes, including a lack of RNA interference (RNAi) pathway in some species, as well as different machinery for transcriptional regulation (Duncan et al., 2017; Lye et al., 2010; Genois et al., 2014; Mannaert et al., 2012). Initial studies to investigate gene functions in *Leishmania* parasites relied on methods such as targeted gene knock-outs using homologous recombination; this is only possible for non-essential genes, or those for which culture methods can ameliorate defects in the parasite (Duncan et al., 2017). DiCre recombinase approaches have been used to generate conditional mutants, allowing study of genes which play essential roles in differentiation (Duncan et al., 2017).

CRISPR-cas9 genome editing is becoming an established technique in *Leishmania* research. Briefly, the CRISPR-cas9 genome editing system utilises CRISPR RNA to 'guide' the Cas9 enzyme to a specific target site on nucleic acids. Cas9 can then cut the DNA or RNA, producing a double stranded break which can be repaired, using homologous recombination, via insertion of a supplied 'repair cassette' sequence, allowing targeted modification of the genome, as well as other nucleic acids within the cell. In 2015, the technique was applied in *L. donovani*, allowing researchers to determine one mechanism for miltefosine resistance (Zhang and Matlashewski, 2015). The technique was further optimised in 2017, to enable application to multiple *Leishmania* species (Zhang et al., 2017). Beneke et al (2017) recently described a high-throughput CRISPR-cas9-based methodology for precise gene-

editing in *Leishmania* promastigotes. The use of CRISPR-editing is of particular value in studies involving multi-copy genes (Duncan et al., 2017).

The scalability of this method has made it invaluable for further *Leishmania* research projects, allowing rapid, high-throughput generation of genetically modified libraries for research into aspects of *L. mexicana* cell biology, including kinases (Baker et al., 2020), flagellar proteins (Beneke et al., 2019) and deubiquitinases (Damianou et al., 2020). Additionally, it has been utilised in studies, including looking at the role of RagC mutations in *L. donovani* infection pathogenesis (Lypaczewski et al., 2021), drug resistance in *L. infantum* (Queffeulou et al., 2024) and the involvement of aquaglyceroporin in antimony resistance (Potvin et al., 2021), among others. Further developments in conditional expression systems and modulation of function at the protein level will prove invaluable in the characterisation of a wider array of *Leishmania* gene functions (Duncan et al., 2017), whilst recent progress in precision base editing (unpublished work by C. Hughes and J. Mottram; Beneke and Engstler, 2023) could further revolutionise this field.

1.2.3.2 *Leishmania* gene expression

Leishmania parasites, in common with other kinetoplastids, are unable to regulate gene transcription, instead controlling expression via post-transcriptional modification of mRNA, post-translational modification of proteins, or altering gene dosage (Haile and Papadopoulou, 2007; Iantorno et al., 2017; Lahav et al., 2011). Research has indicated that these regulatory processes are involved to differing degrees at different stages of differentiation; Lahav et al (2011) showed that regulation of mRNA abundance is more prominent in promastigotes, whilst amastigotes utilise control of protein abundance more readily. Control at the translation level appears to be more important in amastigotes, with analysis of *L. donovani* indicating that ribosomes and translation factors are downregulated in the process of differentiation from promastigote to amastigote (Lahav et al., 2011).

Differences in gene expression are thought to contribute to the distinct tissue tropisms and disease pathologies between different species. One study noted large discrepancies in gene expression between *L. major* and *L. infantum*, across both promastigote and amastigote life stages (Rochette et al., 2008). Indeed, the notable difference in disease presentation of *L. donovani* infection in Sri Lanka, which can

result in cutaneous disease, has been linked to differences in the parasite genome when compared to parasites from other geographical regions (Kariyawasam et al., 2017). Comparisons between visceral and cutaneous isolates from Sri Lanka note changes in gene copy number and specific mutations in genes with varying functions; one implicated gene is the A2 gene, previously linked to visceralisation, but this gene was found not to be solely responsible for the dissemination observed in CL versus VL-causing strains of *L. donovani*. Rag C, involved in mTOR signalling, was also noted to be mutated in this line (Zhang et al., 2014). A2 has previously been implicated in virulence (Zhang et al., 1997), but further research is necessary to establish how differences in parasite genomes and gene expression are related to their disease pathologies.

1.2.3.3 Genomic plasticity and instability

Aneuploidy is generally detrimental in multicellular organisms, including humans, but in microbes it can provide an advantage in terms of adaptation to new environments or hosts, and for drug resistance development (Kwon-Chung and Chang, 2012; Lukow and Sheltzer, 2022). Alterations in gene copy number are common across multiple *Leishmania* species (Rogers et al., 2011), and are one method by which gene dosage is regulated (Iantorno et al., 2017). Work by Dumetz et al (2017) has shown that the transition between life stages is associated with rapid changes in aneuploidy in *L. donovani*. They compared parasite genomes after *in vitro* culture as promastigotes to *in vivo*-derived populations from hamster passage or sand fly infection, finding that differences in smaller scale genome variation such as single-nucleotide polymorphisms (SNPs) were minimal, but chromosome duplication was far greater in promastigote stages than in amastigotes. This could indicate a possible role in adaptation to different hosts. Differences were associated to changes in transcription of specific genes, with amastin genes expressed at higher levels in amastigotes, whilst metabolism-linked genes were more highly expressed in promastigotes (Dumetz et al., 2017).

Another study by Negreira et al (2021) investigated the degree of mosaicism within a clonal promastigote population of *L. donovani* using single cell sequencing. They found a great deal of heterogeneity in ploidy, with some chromosomes in some cells being completely lost. Their work indicates that promastigotes may undergo

rapid chromosome duplication early in transition from amastigotes, or that ploidy changes may result from hybridisation, and that specific changes are linked to selective pressures (Negreira et al., 2021). Mosaicism has also been reported in *L. braziliensis* and *L. mexicana* populations (Rogers et al., 2011). Hybridisation has previously been reported in *Leishmania in vitro* and *in vivo* (Louradour et al., 2020; Romano et al., 2014), whilst defective cell segregation is another putative model for high levels of aneuploidy (Black et al., 2023). Since *Leishmania* do not use selective transcription to regulate gene expression, differences in transcription as a result of ploidy changes can provide one method of introducing functional diversity to the population (Seco-Hidalgo et al., 2015).

Aneuploidy has been linked to drug resistance in promastigote *L. major* and *L. infantum*, as has increased expression of specific genes in circular DNA and chromosomal deletion of others (Leprohon et al., 2009; Ubeda et al., 2008). Aneuploidy in field isolates of *L. donovani* has been reported, and this was linked to antimony-resistance (Imamura et al., 2016). However, in all of these cases, the strains were analysed after lab culture, meaning these differences may not be representative of the parasite population *in vivo* in amastigote form.

Gene dosage can be regulated by copy number variation (Iantorno et al., 2017). Copy number variations are also common across species, appearing more regularly in disomic chromosomes (Rogers et al., 2011), and are associated with adaptive responses to stresses, such as drug exposure (Ubeda et al., 2014). These duplications appear to be reversible, exemplifying the plasticity of *Leishmania* genomes (Ubeda et al., 2014). *Leishmania* genomes contain many repetitive DNA regions, which have been implicated in increased genome plasticity (Ubeda et al., 2014); they are thought to lead to tandem repeats and amplification of specific regions via hairpins, single strand and double strand breaks, which trigger DNA repair mechanisms and cause duplications (Black et al., 2023).

1.3 Immuno-biology of leishmaniasis

1.3.1 Host cell interactions

1.3.1.1 Macrophages, Dendritic Cells (DCs) and monocytes

Leishmania parasites are predominantly parasites of myeloid cells, such as macrophages. Upon infection via sand fly bite, they are phagocytosed by these cells and transform from promastigotes to amastigotes. In macrophages, amastigotes subsequently multiply in phagolysosomes (Kaye et al., 2020; Kaye and Scott, 2011). Phagocytes can exert leishmanicidal activity when exposed to IFNγ (Murray et al., 1983), but *Leishmania* parasites have evolved methods to subvert this killing capacity (Kima and Soong, 2013; Ray et al., 2000). Multiple studies have characterised the ability of *Leishmania* parasites to modify the phagolysosome in a manner that enables their own survival. *L. donovani* promastigotes can block superoxide generation in an LPG-dependent manner (Lodge et al., 2006), reprogramming macrophages to downregulate their response to IFNγ (Nandan and Reiner, 1995; Nandan et al., 1999). Others have found that LPG and GP63 are involved in the ability of *L. donovani* promastigotes to block lysosome maturation and fusion with endosomes (Desjardins and Descoteaux, 1997; Scianimanico et al., 1999; Verma et al., 2017).

The inflammasome can also be activated by multiple *Leishmania* species (de Carvalho et al., 2019; Lefèvre et al., 2013; Lima-Junior et al., 2013), with contradictory effects, in some cases exacerbating and in others ameliorating disease progression (Zamboni and Sacks, 2019). *L. donovani*, *L. major* and *L. amazonensis* are able to inhibit the inflammasome, with GP63 implicated as one mechanism for this (Gupta et al., 2017; Shio et al., 2015).

Monocytes are also known to act as a host cell for *Leishmania* parasites. They are recruited early after infection, but their influx can be triggered by sand fly biting in the absence of parasites (Romano et al., 2017). Trafficking appears to be mediated by C-C chemokine receptor 2 (CCR2) (Romano et al., 2017; Terrezas et al., 2017). In both human and murine infection, *L. major* parasites have been shown to resist reactive oxygen species (ROS)-mediated killing inside monocytes in the acute phase. However, on secondary infection, these cells are recruited via IFNγ signalling and

can phagocytose parasites more rapidly. These parasites are susceptible to killing and monocytes also show a more mature phenotype when previously challenged (Romano et al., 2017). Other studies have supported this role in re-infection for monocytes, with Glennie et al (2017) reporting that skin-resident memory T cells recruit inflammatory monocytes on secondary infection with *L. major*. These monocytes rapidly kill parasites via ROS and inducible nitric oxide synthase (iNOS), and differ in their expression of various markers of activation when compared to those in a primary infection model (Glennie et al., 2017).

Inflammatory monocytes have also been implicated in VL severity; in a murine *L. donovani* infection model, these monocytes are recruited to the spleen and liver, where they can host parasites, whilst blocking their CCR2-mediated influx can reduce parasite burdens in these tissues. The phenotype of these monocytes differs in a tissue-specific manner, such that splenic monocytes are more permissive to parasite survival (Terrezas et al., 2017); this effect was found using an amastigote infection model, so the role of monocytes in VL infection may differ in the more natural promastigote injection or sand fly bite models. In the bone marrow, CD4+ T cells are able to activate inflammatory monocytes via IFNγ when exposed to *L. donovani*, and this can help to control parasite burdens in this tissue (Romano et al., 2021). Monocytes have also been implicated in antigen presentation and the induction of later adaptive immune responses (Romano et al., 2017), whilst depletion has been noted to alter the secretory actions of CD4+ T cells (Terrezas et al., 2017).

DCs are also known to harbour *Leishmania* parasites and are involved in antigen presentation to induce an adaptive immune response to the parasites. Promastigote *Leishmania* infection can interfere with this antigen-presenting role via downregulation of surface markers including major histocompatibility complex (MHC), and can alter the cytokine expression profile of these host cells (Contreras et al., 2014; Neves et al., 2010). Additionally, *L. donovani* infection has been shown to prevent DC trafficking to the lymph nodes via induction of Transforming Growth Factor beta (TGFβ) (Yadav et al., 2023). A similar effect was noted for *L. amazonensis*, but no mechanism was determined (Hermida et al, 2014). However, *in vitro* experiments have indicated that bone marrow DCs (BMDCs) show increased

migration after *L. infantum* infection; this migratory capacity seems to be dependent on *Leishmania* species (Luz et al., 2023).

1.3.1.2 Neutrophils

Another key early event in VL, and other leishmaniases, is neutrophil recruitment; neutrophils have been established to strongly contribute to disease severity and are a key early responder after sand fly bite (Carlsen et al., 2015; Ribeiro-Gomes and Sacks, 2012; Romano et al., 2017), but may play differing roles dependent on the tissue and timing of their action (Ribeiro-Gomes and Sacks, 2012). Complement activation can trigger neutrophil recruitment upon initial infection (Jacobs et al., 2005), as can localised disruption of the extracellular matrix (ECM) (de Menezes et al., 2016). There is evidence of neutrophil infection by VL and CL-causative species (Carlsen et al., 2015; Chang, 1981; Lauts et al., 2002; Oualha et al., 2019; Pearson and Steigbigel, 1981; Ribeiro-Gomes and Sacks, 2012; Romano et al., 2017; Terrezas et al., 2017), and that neutrophils may act as a 'sanctuary' for parasite survival (Carlsen et al., 2015; Ribeiro-Gomes and Sacks, 2012; Romano et al, 2017). Romano et al (2017) found that this protective role can occur in both primary and secondary infections with *L. major*. Parasites appear to preferentially infect neutrophils in the immediate aftermath of inoculation, but then transition to nearby monocytes over the proceeding hours (Romano et al., 2017). However, other studies show neutrophils are capable of killing phagocytosed parasites (Chang, 1981; Pearson and Steigbigel, 1981). Parasite species appears to be important and can differentially alter the phenotype of infected neutrophils (Oualha et al., 2019), potentially altering their parasite killing behaviour.

Neutrophil Extracellular Traps (NETs) are one method neutrophils use to kill pathogenic microbes; their production can be induced by *Leishmania* (Guimarães-Costa et al., 2009; Rochael et al., 2015). NETs have been shown to act in a manner which can increase uptake of promastigote parasites at infection site, and *L. donovani* specifically can resist killing via NET-osis (de Menezes et al., 2016; Gabriel et al., 2010). LPG appears to contribute to this evasive capacity in *L. donovani* promastigotes (Gabriel et al., 2010), whilst nucleases have been implicated in the case of *L. infantum* (Guimãraes-Costa et al., 2014). Conversely, exposure of various cutaneous species to histones has been shown to result in promastigote death and

reduced infectivity, whilst *L. amazonensis* and *L. mexicana* amastigotes can resist histone-mediated killing *in vitro* (Guimarães-Costa et al., 2009; Wang et al., 2011), raising the question of whether this difference applies *in vivo* and whether life stage differences are also present for visceral species of *Leishmania*.

1.3.1.3 Other putative host cells

Evidence supports the ability of *Leishmania* parasites to infect host cells other than the traditionally recognised phagocytes. Karagiannis et al (2023) used a dual-single cell RNA-seq study of both spleen and bone marrow of BALB/c mice infected with *L. donovani* amastigotes to investigate how parasitized cells respond to infection. Findings indicated that bone marrow Hematopoietic Stem Cells (HSCs) were highly infected, alongside a smaller number of eosinophils. In the spleen, NK cells, basophils and megakaryocytes also showed evidence of infection. HSC infection was subsequently confirmed *in vitro*, within 24 hours of parasite exposure. Differences in the expression of key receptors were noted between infected and uninfected cells. Their approach theoretically allowed analysis of the parasite transcriptional expression profile in infection, but in practice this was not possible. A major caveat is that the data is entirely derived from a single infected female mouse, and as such may not be representative generally, and cannot account for any differences between infection profiles in other mouse strains, ages or males. Further studies (Dirkx et al., 2022; Dirkx et al., 2024) have since confirmed HSC infection occurs *in vivo*, and they have found that these cells can alter their gene expression profiles to provide a protective environment for parasites, which may be linked to relapse of infection. Bone marrow parasites in a BALB/c VL model were noted to be able to resist drug treatments and to recolonise other tissues after paromomycin clearance, with *ex vivo* confirmation of their ability to tolerate high stress; similar gene signatures were found in HIV-VL co-infected patient HSCs (Dirkx et al, 2022). HSC-residing parasites appear to have an enhanced propensity to enter quiescence, in contrast to macrophage-derived amastigotes, and are apparently more infective (Dirkx et al., 2024). Reticular fibroblasts have also been implicated as cells able to harbour persistent *L. major* parasites, which may contribute to relapse (Mandell and Beverly, 2017).

1.3.2 Complement

Both visceral and cutaneous models show a complex interplay with the complement system. *Leishmania* parasites can fix complement via different pathways (Ambrossio et al., 2021; Mosser et al., 1986; Mosser and Edelson, 1984; Tirado et al., 2021), which can both enhance MAC-mediated lysis of parasites (Mosser and Edelson, 1984; Hoover et al., 1984; Tirado et al., 2021) and can also be exploited for increased recruitment of, entry into, and parasite survival within, phagocytes (Brittingham et al., 1995; Jacobs et al., 2005; Matte et al., 2021; Mosser and Edelson., 1984). Importantly, the promastigote non-metacyclic life stage has been shown to have a higher susceptibility to opsonisation than the amastigote or metacyclic parasite stage (Mosser and Edelson, 1984; Tirado et al., 2021).

1.3.3 Type 1 interferons (IFNs)

Whilst IFNγ is critical for parasite killing by phagocytic host cells and induction of the Th1 response (reviewed by Kima and Soong, 2013), the role of Type 1 IFNs in leishmaniasis is more complex and contradictory (Kaye and Scott, 2011). Neutrophils have been shown to promote visceral disease in an IFNβ-dependent manner (Dias et al., 2019). Whilst Type 1 IFNs have been more widely researched in terms of viral infection (McNab et al., 2015), IFNβ has been found to increase parasite burdens and block macrophage-mediated killing via ROS in CL (Khouri et al., 2009; Mattner et al., 2000). However, it can lessen disease progression at lower doses after *L. major* infection (Mattner et al., 2004). Dias et al (2019) described a role for Neutrophil Elastase expression in macrophages leading to induction of IFNβ after *L. donovani* infection, which increased parasite survival, whilst Xin et al (2010) showed that Type 1 IFN was important for clearance of *L. amazonensis in vivo*. Mice lacking IFNα/β receptor (IFNAR) had reduced parasite burdens, with lower levels of monocytes but increased neutrophils recruited to the infection site compared to controls; neutrophils lacking IFNAR were more able to kill parasites when injected into a wild-type mouse line, whilst the reverse experiment led to reduced parasite death (Xin et al., 2010). Other intracellular bacterial infections involving macrophage host cells have indicated that Type 1 IFNs may be important, with *Mycobacterium tuberculosis* and *Listeria monocytogenes* both showing increased survival when IFNα/β signalling is induced (McNab et al., 2014; McNab et al., 2015;

Kotov et al., 2023). IFNα/β is involved in the trafficking of host cells to the lung (Antonelli et al., 2010; McNab et al., 2015), monocyte differentiation (Antonelli et al., 2010; Mariotti et al., 2004) and in neutrophil mediated plasmacytoid dendritic cell (pDC) activation (Kotov et al., 2023) in tuberculosis; Type 1 IFN signalling may be involved in reducing the response of macrophages to IFNγ triggering of ROS-mediated *Leishmania* killing capacity (McNab et al., 2014; McNab et al., 2015; Rayamajhi et al., 2010).

1.3.4 T cells and the Th1 response

Chronic VL infection is markedly tissue-specific, with the liver capable of regulating *Leishmania* parasite burdens, whilst the spleen becomes the site of chronic infection and inflammation (Bankoti and Stäger, 2011). It has been established that a Th1-biased response is important in disease mitigation and parasite clearance, whilst Th2 responses are detrimental in CL, but not so in VL infection (Das et al., 1999; Lehmann et al., 2000; Kaye et al., 1991; Scott et al., 1988; Scott et al., 1990) (Figure 1.2A). CD4+ T cells are essential for parasite control, via activation of phagocytic killing, but the role of CD8+ T cells is less well characterised, and appears to differ in different contexts (Kaye and Scott, 2011; Stäger and Rafati, 2012), whilst *L. donovani* parasites have been noted to induce CD8+ T cell exhaustion in chronic infection models (Joshi et al., 2009). In CL, there appear to be subsets of CD8+ T cells which secrete different cytokines and differ in function in primary infections. For both VL and CL secondary infections, CD8+ cells are necessary for control of the new infection (Stäger and Rafati, 2012). CD4 and CD8 T cell secretion of cytokines such as TNFα, IL2 and IFNγ is associated with VL resolution both in patient studies (Rodrigues et al., 2021) and in animal models (Das et al., 1999; Kaye et al., 2004; Lehmann et al., 2000). In the bone marrow, CD4+ IFNγ secretion can activate inflammatory monocytes in the presence of *L. donovani* parasites (Romano et al., 2021).

A key aspect of liver control is presumed to be the formation of granulomas, which are noted to be a response to many microbial infections. A detailed review of the role of these structures is found in Murray, 2008 and Moore et al., 2013. Briefly, the initial events are instigated by infection of Kupffer cells, which results in recruitment of monocytes and T cells and the expression of Th1 cytokines. Recruitment of CD8+

T cells to liver granulomas via Kupffer cell interactions is important for control of liver parasite load (Beattie et al., 2010; Stäger and Rafati, 2012). TNFα plays a critical role in granuloma development, and the involvement of B cells has also been noted (Kaye et al., 2004; Moore et al., 2012; Moore et al., 2013; Murray, 2008).

IL-12 is also a key player in liver resolution; a study in BALB/c mice found that IL-12 knock-out resulted in reduced granuloma formation and IFNγ induction, and a delayed resolution profile, whilst in the spleen, IL-12 was essential for controlling parasite load and structural dysregulation in later infection stages (Engwerda et al., 1998). IL-4, meanwhile, has been implicated in worsening CL infection with both *L. mexicana* and *L. major* via polarisation towards a Th2 response (Satoskar et al., 1995; Mohrs et al., 1999). Conversely, the role of this cytokine in *L. donovani* infection exacerbation appears to be minimal, and IL-4 may be beneficial in the development of liver granulomas and parasite control (Stäger et al., 2003; Satoskar et al., 1995). This protective effect was illustrated in a study by McFarlane et al (2019), in which IL-4 knock-out mice had increased *L. donovani* parasite burdens compared to wild-type mice; mice whose T cells could not respond to IL-4, in contrast, were not more susceptible to infection, indicating that T cells were not the key responder to this cytokine. They hypothesised that IL-4 could thus be protective via induction of IL-12 secretion by DCs (McFarlane et al., 2019).

Early work identified IL-10 as important in *L. donovani* infection, with mouse knock-out models showing resistance to the parasite. Whilst IL-10 is not involved in either phagocyte uptake of parasites or early establishment in the liver and spleen, clearance of both tissues occurs in mice lacking IL-10 by day 14 of infection; this ability is at least partially dependent on the production of Th1 cytokines IL-12 and IFNγ (Murphy et al., 2001). There is significant evidence of a role for IL-10 in parasite persistence, with chronic VL infection in the spleen linked to TNFα mediated induction of IL-10, along with extensive splenic remodelling and loss of marginal zone macrophages early in infection (Engwerda et al., 2002; Kaye and Scott, 2011; Montes de Oca et al., 2020; Terrezas et al., 2017).

Central memory T cells appear to play little role in ongoing *Leishmania* immunity (Peters et al., 2014), whereas tissue-resident CD4+ memory T cells in the skin have been shown to contribute to lasting protection in CL (Glennie et al., 2015; Scott,

2020). This occurs via recruitment of circulating effector T cells (Glennie et al., 2015), and via inflammatory monocyte activation in *L. major* infections (Glennie et al., 2017). Parasite persistence and an equilibrium between the Th1 effector T cells and Tregs/inducible Tregs (iTregs) appear to be important for maintenance of long-term immunity (Peters et al., 2014; Peters and Sacks, 2006; Scott et al, 2004). Furthermore, γδ T cells have previously been shown to be increased by *L. donovani* infection – these cells predominate in the gut mucosa, triggering immune suppression, and could be linked to parasite burden heterogeneity between individuals in this tissue (Jawed et al., 2019).

1.3.5 B cells

In contrast to T cells, relatively little is known about the role of B cells in the adaptive response to leishmaniasis, despite the importance of memory in vaccine design. One characteristic VL symptom is hypergammaglobulinemia, a result of innate B cell activation (Silva-Barrios and Stäger, 2019; Silva-Barrios et al., 2016). Early in infection, B cell activation can exacerbate visceral disease (Silva-Barrios et al., 2016); interaction with *L. donovani* parasites has been shown to activate B cells, increasing MHCII and CD86 expression and the production of Type 1 IFNs (Stogerer et al., 2023; Silva-Barrios et al., 2016). Mature B cells are recruited to granulomas in the liver after *L. donovani* infection, where they interact with T cells; these B cells were not found to be regulatory or antigen specific (Moore et al., 2012). In the spleen, B cell deficiency has been shown to improve T cell responses to *L. donovani* (Bankoti et al., 2012). Activated splenic B cells induce a Th1-polarised response (Silva-Barrios et al., 2016) and in the marginal zone (MZ) are capable of antigen presentation (Bankoti et al., 2012).

A recent paper describes a new role for B cells in the context of VL. *L. infantum* promastigotes appear to be able to attach to the surface of B cells, and can then differentiate into amastigotes, both *in vitro* and *in vivo*. These events can trigger activation and an increase in B cell numbers in multiple tissues. It was additionally reported that the amastigotes in question are capable of infecting macrophages (Dirkx et al., 2024). Other studies have found a similar phenomenon in *L. donovani*, and have also noted that the infective capacity of these parasites is not only

retained, but that specialised B cell protrusions can facilitate the transfer of parasites between B cells and host cells (Bankoti et al., 2012; Stogerer et al., 2023).

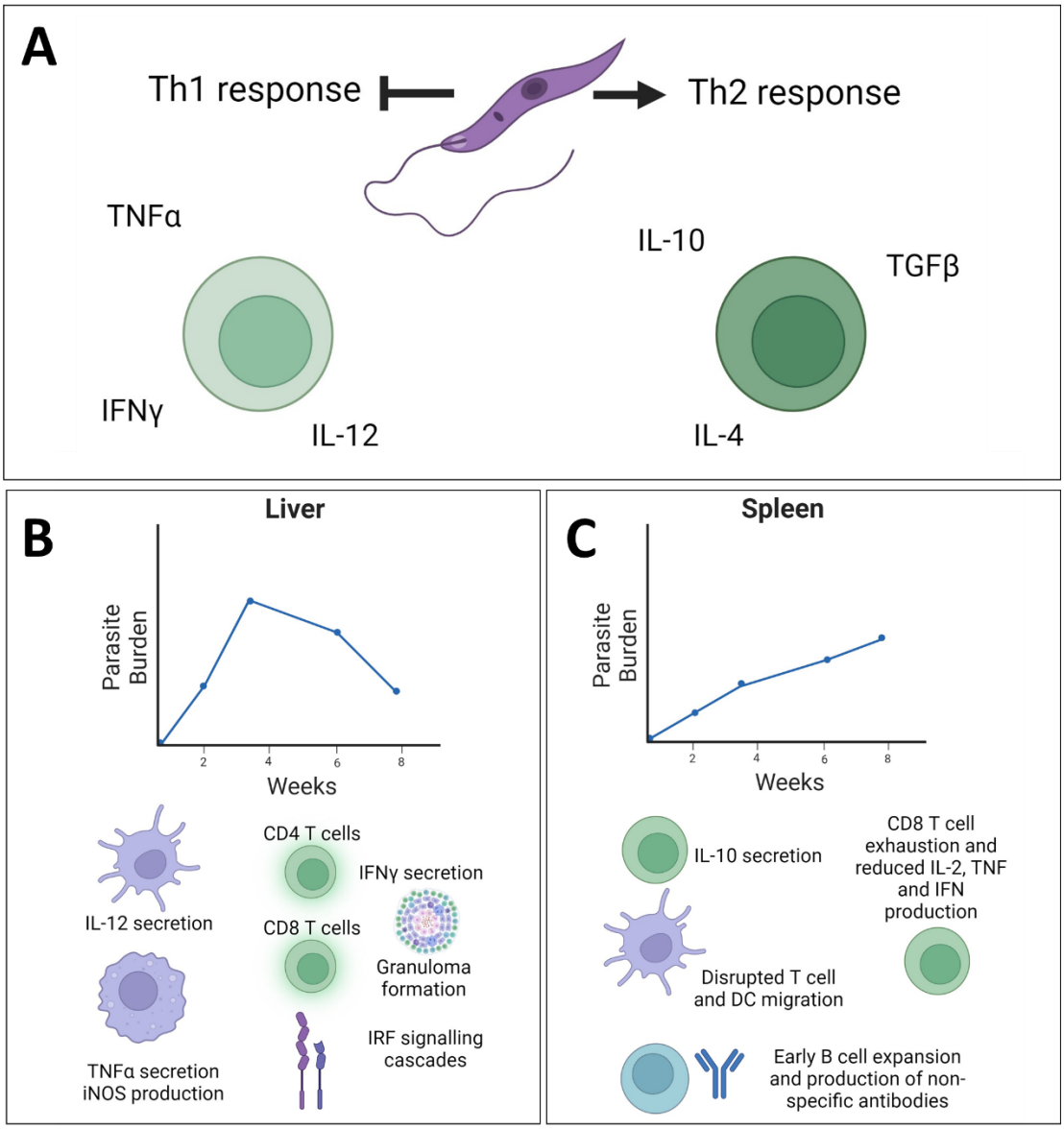


Figure 1.2: The adaptive immune response to VL infection is tissue-specific. A). *Leishmania* parasites are favoured by a failed Th1/regulatory response, and have multiple mechanisms by which they can bias a host towards this and away from the more detrimental Th1 response (see text for details). B). The liver is characteristically able to control parasite burdens in long term VL infection, with a reduction in parasites around 6-8 weeks in mouse models. The liver uses a Th1-dominated response to achieve this, along with granulomas which contain parasites and ameliorate disease. C). The spleen, in contrast, is not able to regulate parasite burdens, which continue to increase over a chronic infection. Spleen immune responses are characterised by CD8 T cell exhaustion, with increased IL-10 and B cell expansion, DC dysregulation and splenic disfigurement. Figure created in Biorender using information taken from Bankoti and Stäger, 2011.

1.3.6 Immune evasion and parasite persistence

Various *Leishmania* species are known to disrupt the adaptive immune response to improve their own survival (reviewed in Gupta et al., 2013 and Peters and Sacks, 2006) (Figure 1.2A). *Leishmania* parasites can dysregulate antigen-presentation by macrophages and dendritic cells (DC) by interfering with MHCII expression and stability (Gupta et al., 2013; Pradhan et al., 2021). The parasites can also induce IL-10 and TGFβ secretion by Tregs (Peters and Sacks, 2006), which can prevent DC trafficking to lymph nodes (Yadav et al., 2023). DC migration within the spleen is disrupted in *L. donovani* infection via a TNFα IL-10 CCR7 axis (Ato et al. 2002), and *L. donovani* infection can also alter the balance of DC1 vs DC2 to favour parasite survival (Mishra et al., 2023). *L. major* antigens can induce a Th2 response in mouse lymphocytes *in vitro*, increasing IL-4 and decreasing IFNγ expression (Tabatabee et al., 2011). Multiple *Leishmania* species are reported to block IL-12 secretion, which is key for the induction of Th1 differentiation and IFNγ-mediated killing (Belkaid and Sacks, 1998; Carrera et al., 1996; Gupta et al., 2013; Peters and Sacks, 2006; Weinheber et al., 1998). Studies indicate that cytokine secretion by B cells stimulated with *L. major* can increase susceptibility in BALB/c mice by influencing this axis towards Th2 (Ronet et al., 2009), whilst *L. donovani* has been shown to activate B cells, inducing IL-10 and Type 1 IFN production which can dampen protective T cell responses and increase levels of low-affinity antibodies (Silva-Barrios et al., 2018).

Resolution of disease does not necessarily result in sterile clearance of parasites from tissues (Mendez et al., 2004; Peters and Sacks, 2006). In murine VL models, parasites are known to persist in the spleen and bone marrow, even with apparent resolution in the liver (Bankoti and Stäger, 2011) (Figure 1.2B and C). In patients, the potential for relapse is well established for visceral and cutaneous leishmaniasis infections, including after apparent cure via chemotherapy (Osman et al., 1998; Mukhopadhyay et al., 2014; Santos et al., 2023). Latent parasites are known to re-emerge in the skin in cases of PKDL but the underlying mechanism is as yet unclear; T cell phenotypes differ between Sudanese and Asian patients, but systemic dysregulation of T cell cytokine secretion appears to be involved (Mukhopadhyay et al., 2014). Possible roles for drug regimens, UV exposure, re-infection and

immunosuppression, including after HIV infection, are reported, with loss of immune regulation of latent parasites believed to be a contributing factor (Mukhopadhyay et al., 2014).

1.4 Microbial ecology

The advent and subsequent advances of molecular biology and computational techniques have enabled the application of macro-ecology approaches to microbiological systems. Some adaptations to specific concepts are required when considering these generally more complex systems of large numbers of microbes, particularly where interactions between host and commensal or pathogen are concerned (Nemergut et al., 2013).

1.4.1 Concepts of microbial community ecology

Community ecology relates to the groups of different organisms which interact with each other, defining how these interactions occur and how they shape the formation of populations within the community. Commonly, the community studied does not actually comprise the entire set of organisms present, but only a subset.

The forces shaping population assembly can be referred to as assembly processes (Figure 1.3A). Broadly, assembly processes can be divided into deterministic and stochastic (for a detailed review, see Nemergut et al., 2013), where a deterministic process is evidence for selective forces which are niche-specific, whilst stochastic processes are those occurring due to dispersal-related factors or random drift (Chase, 2010; Chase and Myers, 2011; Hanson et al., 2012; Nemergut et al., 2013; Zhou and Ning, 2017). Selective forces can be further split into homogeneous and heterogeneous, where homogeneous indicates two environments favour similar species traits and heterogeneous indicates the opposite (Stegen et al., 2023; Stegen et al., 2015). Finally, speciation or diversification can also lead to stochastic variance over time; in microbial systems the concept of a species is perhaps harder to define than in traditional macro-organisms, so this concept is usually reconceptualised as mutation in these cases (Hanson et al., 2012; Nemergut et al., 2013; Vellend, 2010). Smaller communities can generally be well accounted for using stochastic processes

and drift dynamics; drift especially is more important when considering rare species in a population (Hanson et al., 2012; Sloan et al., 2006).

Priority effects and historical contingency can also be deterministic, in that early arriving individuals can alter the niche in ways which favour (niche facilitative modification) or are detrimental to (niche pre-emption and niche inhibitory modification) later arrivals (Debray et al., 2022). Niche pre-emption occurs when early arrivals use resources in such a way that few are available to late-comers. It has been shown that this process is more important where individuals are closely related, as competition for the same resources increases, whilst in less hospitable environments, this effect lessens, as early individuals are not able to grow as rapidly in these hostile conditions (Peay et al., 2012; Tan et al., 2012; Debray et al., 2021; Tucker et al., 2014). Niche facilitative modification occurs when early arrivals modify the niche such that newcomers are more able to survive if others are already present (Debray et al., 2021). This could be via suppression of host immunity in the case of pathogenic infections (Potnis et al., 2014; Halliday et al., 2020). Conversely, niche inhibitory modification can alter the environment to reduce colonisation ability of those entering later (Debray et al., 2021). One example could be tissue destruction by pathogenic organisms making later establishment more difficult (Foster et al., 1994; O'Keefe et al., 2021).

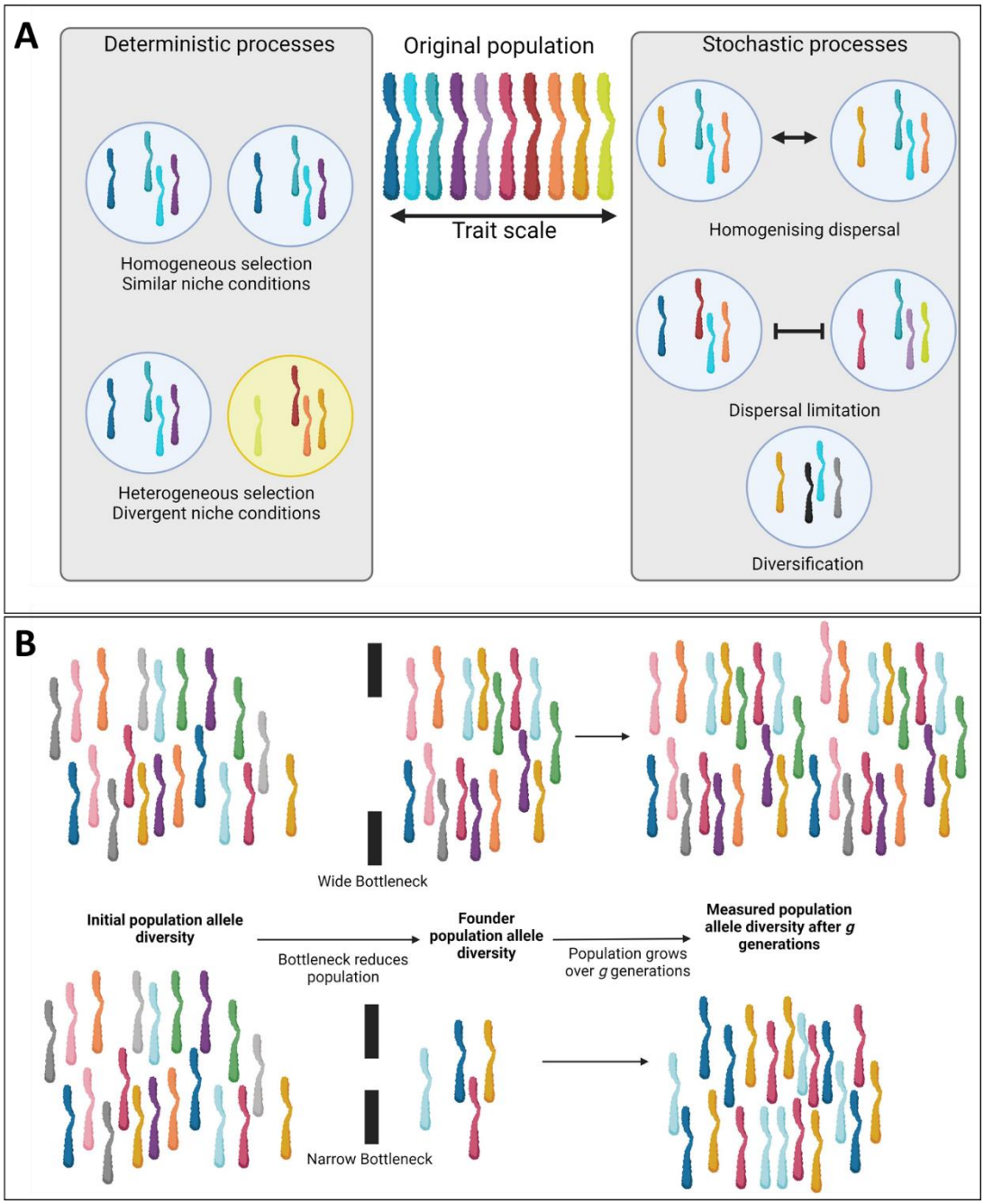


Figure 1.3: Microbial population dynamics. A). A schematic representation of the assembly processes which can shape microbial communities. These are grouped into deterministic and stochastic. Deterministic processes result in trait-based selection dependent on the environmental conditions in each niche. Conditions select for similar traits in homogeneous selection, whilst in heterogeneous selection, the environmental conditions differ (represented by the circle colour difference), meaning divergent traits are selected for. Stochastic processes include homogenising dispersal, resulting in similar populations. If dispersal is limited, then populations will differ. Diversification is the appearance of new species or alleles within the population. See Nemergut et al., 2013 for a detailed review. B). Schematic representation of bottlenecks in microbial colonisation. When a bottleneck is

'wide' (i.e. permissive), a larger number of individuals from a population can pass through it, resulting in a larger 'founder population' than when a bottleneck is 'narrow' (i.e. restrictive). A narrow bottleneck curtails the population more, giving a small founder population, which means that the diversity of alleles present is also reduced. These founder populations grow to become larger populations, and measuring the diversity of alleles present at a later time-point can allow determination, via mathematical modelling, of the original founder population size. This is a proxy measure for the bottleneck width itself. Based on a figure in Abel et al., 2015b. Figure created in Biorender.

The concept of bottlenecks (Figure 1.3B) is more commonly associated with ecology, describing an obstacle which is applied to a population and thus reduces its size and diversity (Abel et al., 2015b). This results in changes to the population structure, which can alter evolutionary processes over time. For example, a tight bottleneck can result in a 'founder effect', where a specific trait becomes amplified despite conferring no competitive advantage, and even in some cases a significant disadvantage to the individual (Yuste et al., 1999). Likewise, genetic drift can result in random differences between populations, due to no selective pressure, but simply as a result of a reduction in population size and diversity (Abel et al., 2015b). In a microbial system, such bottlenecks can influence the development of drug resistance, adaptability to additional hosts, growth patterns and changes in virulence (Ebert, 1998; Grenfell et al., 2004; Levin et al., 2000). These are all important factors to understand when studying pathogens, as they can influence how they spread within and between hosts and cause infection, and how effective treatment options are. Bottlenecks can be caused by and operate in different ways; for example, a toxic bottleneck, such as drug treatment, may kill a proportion of the population, whilst a resource limit, such as host cell availability, may impose a ceiling on population size. These different forces can then shape populations in different ways, although in reality most bottlenecks will be a more complex mixture of multiple factors (Abel et al., 2015b).

1.4.2 Microbial population dynamics *in vivo*

Much research on *in vivo* population dynamics has concentrated on the microbiome (Debray et al., 2021). However, within-host pathogen population dynamics is a field

of growing interest and importance in the understanding of infectious diseases. Recent studies have analysed population bottlenecks in many bacterial infection models (Abel et al., 2015; Bachta et al., 2020; Campbell et al., 2023; Hotinger et al., 2024; Hullahalli and Waldor, 2021; Louie et al., 2019; Zhang et al., 2017). Recently, Wincott et al (2022) applied these methods in mouse models of infection with the eukaryotic parasite *Toxoplasma gondii* and found that multiple colonisation events contribute to the establishment of the parasite in the brain; even low abundance clones can become dominant in the population within a tissue in some cases. This is likely due to differences in clonal expansion combined with variation in the host immune response, but the authors did not explore the mechanisms underlying this effect.

Further work has illustrated how this type of analysis can be used to decipher immunological interactions between host and pathogens; in a murine model of *Escherichia coli* infection, infectious dose was found to be linked to toll-like receptor 4 (TLR4) activation, which induced a tissue-specific bottleneck in the spleen. This dose effect could be induced using dead bacteria mixed with the live population, supporting a mechanism reliant on surface lipopolysaccharide (LPS) immune activation (Hullahalli et al., 2023).

Other recent studies have sought to combine bottleneck analysis with determination of assembly processes. Jiang et al (2024) assessed how assembly processes contributed to pathogen populations in plans under different stressors and via different infection methods, and how this related to colonisation dynamics. They quantified bottleneck differences, and attempted to measure the relative effects of selective and stochastic forces; however, their analysis did not account for traits or strain relatedness in any way when assessing the role played by selection, and as such their conclusions with regard to differences in selective pressures appear unsupported.

Another study, by Chevée et al (2024), combined the analysis of bottlenecks and priority effects. In their work, they investigated the how *Listeria monocytogenes* colonises the brain and used sequential infections to determine whether priority effects were the cause of high levels of clonal expansion observed. Their results indicated that early arrivals had an advantage in the brain, which they hypothesised

was simply a result of greater replication time. However, they did not investigate whether other mechanisms may have altered bacterial fitness prior to reaching the brain. For example, the bacteria in the later infection may disseminate to the brain via a different route, which could alter their phenotype; alternatively, the re-infection may mean a greater bottleneck occurs at the injection site, due to a localised inflammatory response. They further compared the colonisation patterns when using different infection protocols, finding that the more natural oral infection route displayed different dynamics to the commonly used intravenous injection method. This finding shows that researchers should be cognizant of the limitations of experimental models when investigating infection dynamics.

1.5 Project Aims

The aims of this project were as follows:

- To develop methods for identifying tissue-specific parasite population bottlenecks in a murine VL model, using artificial barcoding of *L. donovani* parasites.
- To use this system to analyse how population dynamics of parasites within infected tissues and dissemination patterns between these tissues change in C57BL/6J mice over the course of infection
- To assess sex differences in the dynamics of infection in this model
- To investigate some of the immunological mechanisms underlying these population dynamics and dissemination patterns
- To develop a re-infection model of VL to understand how prior and subsequent infection affect each other at a tissue level

Chapter 2: Materials and Methods

2.1 Parasites

All parasite lines used were derived from the WHO designated MHOM/ET/67/L82 *Leishmania donovani* strain, referred to as LV9 (also known as HU3). *L. donovani* LV9 T7 cas9 was generated prior to this project by Juliana Carnielli, following the method described in Carnielli et al., 2022 (for *L. infantum* species). The transfection protocol for generation of the barcoded parasite library using this line is described in 2.5. *L. donovani*-RE9H LV9 parasites, for IVIS imaging, were generated prior to this project by insertion of a red-shifted firefly luciferase gene (Ppy-RE9H) into the ribosomal RNA locus of LV9 promastigotes (Branchini, 2010; Soysa et al., 2015).

Parasites were passaged through *B6.Rag2*$^{-/-}$ CD45.1 (RAG) or BALB/c mice for amastigote generation. Spleen was placed in un-supplemented RPMI 1640 (Gibco: 31870-025) and homogenised to a single-cell suspension using a glass homogeniser, then centrifuged for 5 minutes at 800 rpm. Supernatant was transferred to a saponin (BDH: 440914Y) coated tube and incubated for 5 minutes at room temperature, then centrifuged at 3100 rpm for 10 minutes. The pellet was then washed three times in RPMI, with centrifugation at 3100 rpm for 10 minutes for each wash, and re-suspended in RPMI, before passing through a 26G needle and counting by eye. Parasites were diluted to the appropriate infection density; most commonly this was 5x10^7 per 100 µl.

Promastigotes were cultured *in vitro* in complete 10% fetal calf serum (FCS) SDM-79 at 25 °C. Drug selection for LV9 T7 cas9 parasites was with 100 µg/ml hygromycin. Selection for barcoded lines was 75 µg/ml puromycin for initial puromycin resistant lines and 20-150 µg/ml blasticidin for blasticidin resistant lines which made up the complete library, with higher concentrations applied for initial selection and lower concentrations for subsequent maintenance *in vitro*. Maintenance of RE9H parasites was with 25 µg/ml G418.

2.2 Culture Media

For complete SDM-79, SDM-79 powder (Gibco: 074-90916N) was supplemented with either 10% or 20% (for recovery from transfection process) heat-inactivated FCS (Gibco: A5256701), 2.5mg/ml hemin (Sigma-Aldrich: 51280-5G) in 50 mM NaOH, 10 mM 6-biopterin (Sigma-Aldrich: B2517-25MG), 10 U/ml penicillin and 10 mg/ml streptomycin (Gibco: 15140-122), prepared in milli-Q (MQ) water. FCS was heat-inactivated at 56 °C for 30 minutes.

For complete DMEM, DMEM (Gibco: 21969-035) was supplemented with 20% heat-inactivated FCS (Gibco: A5256701), 2 mM L-glutamine (Gibco: 25030-024), 100 U/ml penicillin and 100 µg/ml streptomycin (Gibco: 15140-122). For macrophage differentiation media, this was further supplemented with 30% L929 cell supernatant.

2.3 Animals

2.3.1 Mouse strains

C57BL/6J mice (obtained from Envigo or bred in-house) (JAX stock #000664) were infected at 9-14 weeks of age, except for experiments requiring repeat infections, where subsequent infections were at the specified time-points. B6.CD45.1 (B6) mice (bred in-house) were used for production of bone marrow macrophages. BALB/c mice (obtained from Envigo or Charles River) (JAX stock #000651) were infected at 19 weeks of age for the PILOT experiment. B6.129S4-C3tm1Crr/J (C3tm1) mice (Wessels et al., 1995) (bred in-house, from breeders gifted by Delphine Autheman and Gavin Wright) (JAX stock #029661) were infected at 10-14 weeks of age. *B6.Rag2*$^{-/-}$ CD45.1 (RAG) mice (bred in-house) for parasite passage were infected at a minimum of 10 weeks of age, following adverse reactions in some younger mice after promastigote infection.

2.3.2 Infections

All infections were performed via tail intravenous injection of 100 µl of *L. donovani* diluted in RPMI 1640 (un-supplemented) (Gibco: 31870-025). Parasite load and life stage varied depending on experiment; amastigote infections were done using

parasites isolated as described in 2.1, with a total infection burden of 5x10^7 parasites per mouse (see Chapter 4), or with 3x10^7 parasites in the re-infection study (see Chapter 6). Promastigote infections were done with stationary phase parasites at either 1x10^8 or 5x10^7 per mouse (detailed in relevant chapters), with rationale for this dosing mentioned in 4.2. Mice were culled by CO_2 exposure for 5 minutes, or exsanguination if cardiac bleeds were taken, followed by cervical dislocation.

Parasite burdens in infected mouse spleen and liver were assessed via tissue smears fixed in methanol for ten minutes, then stained with 10% Giemsa stain (diluted in water) for 30 minutes at room temperature, before rinsing and leaving to dry. Parasites per 1000 host cells were counted by eye and multiplied by organ weight (in grams) to obtain Leishman-Donovan Unit counts (LDU) for each organ. For any mice for which organ weights were not available due to technical error, the organ weight for LDU calculations was estimated as the mean of the remaining mice in the same experimental group.

2.3.3 Serum re-constitution

For serum addback to B6.129S4-C3tm1Crr/J mice, serum was initially obtained from C57BL/6J mice via cardiac bleeding of mice, with blood samples collected in SST Blood Collection Tubes (BD: 365968). Blood was left at 4 °C for a minimum of 6 hours before centrifugation at 5000 g for 5 minutes to isolate sera. Sera was stored at -80 °C. For re-constitution, sex-matched sera was thawed and pooled. The sera was given to sex-matched B6.129S4-C3tm1Crr/J mice via a 200 µl intraperitoneal injection, between 30 minutes and 2 hours prior to infection with *L. donovani* parasites.

2.3.4 *In Vivo* Imaging System (IVIS) imaging of mice

For live mouse imaging, mice received a subcutaneous injection of 100 µl of 30 mg/ml Firefly D-luciferin (Syd Labs: MB000102-R70170) in 1X PBS, left for 7 minutes, then anaesthetised with 4.0% isofluorane/1.5LO2 per minute for 5 minutes. Mice were imaged using an IVIS Spectrum bioluminescence imaging system (PerkinElmer) to detect light emission using an open emission filter, for

exposures up to five minutes, captured with a charge-coupled device (CCD) camera, for equal times when positioned laterally and ventrally.

Ex vivo imaging of tissues was performed under the same conditions, with luciferin added to the surface of the organ and incubated for 5 minutes in dark conditions prior to exposure for 30 seconds, followed by removal of liver and spleen, and subsequent exposure of other tissues for 5 further minutes.

Analysis was performed in Living Image software for initial image processing to calculate total flux for regions of interest and to generate image figures, followed by quantitative analysis in R.

2.3.5 Ethics

All animal care and experiments were done in accordance with relevant Home Office Legislation and approval by the Animal Welfare Ethics Review Board and Named Veterinary Surgeon at University of York. Most animal experiments were performed with assistance from Naj Brown and Helen Ashwin.

2.3.6 Exclusion of animals

Animals were excluded from further analysis if they showed no sign of infection establishment. For mice which experienced infections with REH-*L. donovani*, this was based on an absence of any bioluminescence signal when organs were imaged *ex vivo*, alongside the following criteria. For mice infected for longer than 2 weeks, absence of infection was defined by absence of splenomegaly and by viewing of liver and spleen smears. For mice which were infected for less than this, only tissue smears were used. These criteria resulted in two mice being excluded from the re-infection study.

2.4 Primers

Primers were designed using CLC software, with sequences for guides and templates created using LeishGEdit (Beneke et al., 2017; Martel et al., 2017) and TryTrypDB (Alvarez-Jarreta et al., 2024). All primers used in this project are detailed in Table 2.1 or Appendix Table 1. All primers were purchased from Merck.

Primer ID	Sequence	**Protocol**
OL6137	aaaagcaccgactcggtgccacttttcaagttgataacggacta gccttattttaacttgctatttctagctctaaaac	**sgRNA PCR**
OL12794	gaaattaatacgactcactataggCTTTGCCCTCGGACGA GTGCgttttagagctagaaatagc	**sgRNA PCR [3' guide]**
OL12793	gaaattaatacgactcactataggACAGACGTCGCGGTG AGTTCgttttagagctagaaatagc	**SgRNA PCR [5' guide]**
OL12831	AAGAATCTCGTGCTTTCAGCTTCGATGTAGgtataat gcagacctgctgc	**Barcode insertion (forward primer)**
OL12833	taatacgactcactataaaactggaag	**Barseq-PCR1 (forward primer)**
OL12845	CTCGGACGAGTGCTGGGGCGTCGGTTTCCAAATT AACCCTCACTAAAG GGAGGAGAAGCTGTTCCAATCCACGcttccagttttat agtgagtcgtattaCCA ATTTGAGAGACCTGTGC	**Barcode 1 insertion (reverse primer)**
OL12881	CTCGGACGAGTGCTGGGGCGTCGGTTTCCAAATT AACCCTCACTAAAGG GAGGtctacgaggactcggtagaacttccagttttatagtgagt cgtattaCCAATTTGAG AGACCTGTGC	**Barcode 2 insertion (reverse primer)**
OL12882	GCGACGTCTGTCGAGAAGTT	**Sanger sequencing/PCR validation of barcode insertion**
OL12884	TGCGTCCGCCTTATACTTCC	**Sanger sequencing/PCR validation of barcode insertion**
OL12888, OL12889, OL12892-OL12897	See Appendix Table 1	**Barcodes 3-10 insertion (reverse primers)**
OL13279-OL13298	See Appendix Table 1	**Barcodes 11-30 insertion (reverse primers)**
Ol13411-OL13480	See Appendix Table 1	**Barcodes 31-100 insertion (reverse primers)**
OL14218-OL14223	See Appendix Table 1	**Barcodes 101-106 insertion (reverse primers)**
OL13833	ACTATTCCTTTGCCCTCGGAC	**Barseq-PCR1 (reverse primer)**

OL14195	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGA taatacgactcactataaaactggaag	**Barseq-PCR2 (forward primer)**
OL14196	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAG TACTATTCCTTTGCCCTCGGAC	**Barseq-PCR2 (reverse primer)**

Table 2.1: Primers used in this project.

2.5 Barcoded parasite library generation

Barcode sequences were designed by Katherine Newling and are listed in Appendix Table 1. 20 bp barcode sequences were selected, with a GC content between 0.4 and 0.6, no BLAST matches to mouse or *L. donovani* genomes, a low likelihood for hairpin structures and a hamming distance of at least 16 between all barcodes.

For the final library preparation, the repair templates were created using pGL2662, which contains a blasticidin-resistance gene (BSD). Forward primer OL12831 and a range of reverse primers containing each barcode (see Table 2.1 and Appendix) were used to amplify the repair template from pGL2662, as individual PCR reactions. The 40 µl reaction was comprised of 40 ng/µl pGL2662 (40 ng), 0.2 mM dNTPs, 0.2 µM forward primer and reverse primer, 1 unit Q5 polymerase (NEB: M0491L), 1X Q5 reaction buffer (NEB: M0491L) and 28 µl MQ water. This was run for initial denaturation for 5 minutes at 98 °C, then cycled 45 times at 98 °C for 30 seconds, 65 °C for 30 seconds and 72 °C for 90 seconds, then run for a final extension at 72 °C for 10 minutes. For optimising the protocol, the same method was used for double transfections but with the template pGL2667, containing a puromycin resistance gene (PAC).

For sgRNA, the 20 µl amplification PCR was made up of 2 µM OL6137, 0.2 mM dNTPs, 0.2 µM OL12793 or OL12794, 1 unit Q5 enzyme, 1X Q5 buffer and 14.6 µl MQ water. The PCR was run at 98 °C for 30 seconds, then cycled for 35 cycles at 98 °C for 10 seconds, 60 °C for 30 seconds, 72 °C for 15 seconds, then run for a final extension at 72 °C for 10 minutes.

Successful amplification was assessed by running on a 1% agarose gel (templates) or a 2% agarose gel (sgRNA), with Sybr Safe DNA Gel Stain (Invitrogen: S33102), at 100 V, to check for appropriate band sizes (approximately 2 kbp for templates and 200

bp for sgRNA). PCR products were then purified using a QIAquick PCR Purification kit (Qiagen: 28104).

Promastigote parasites were prepared for transfection, using the Amaxa P3 Primary Cell 4D-Nuceofector X Kit L (Lonza: V4XP-3024) by spinning 2x10^6 parasites per transfection at 1000 g for 10 minutes at room temperature, washing in a large volume of PBS, then spinning again. Cells were re-suspended in 90 µl electroporation buffer with 15 µl purified guide and repair cassette DNA (or sterile water for controls), before transfer to a cuvette. Cuvettes were placed in the P3 Primary Cell 4D-Nucleofector and electroporated with the Human T cells HE unstimulated programme, then immediately transferred to pre-warmed 20% FCS SDM-79 and incubated at 25 °C for a minimum of 8 hours before addition of drugs for selection. Cells usually recovered in 1-3 weeks if successful, and were passaged at this point to a fresh flask. Selection with blasticidin was done using 150 µg/ml initially, then reduced to 75 µg/ml, then 20 µg/ml for subsequent passages. Puromycin selection was done using 75 µg/ml initially for two passages, then reduced to 20 µg/ml for subsequent passages. For double selection (during optimisation), both drugs were combined at the concentrations stated.

2.6 PCR and Sanger sequencing validation of barcode insertion

To validate barcode insertion via PCR, approximately 1x10^8 promastigote parasites were centrifuged at 3100 rpm for ten minutes. Supernatant was removed and the pellet was frozen at -20 °C for 24 hours. The cell pellet was then used for DNA extraction using a Qiagen DNeasy Blood and Tissue kit (Qiagen: 69506), following the cultured cells protocol. The PCR master mix of 1X Taq RedMix (PCR Biosystems: PB10.23-02), 0.2 µM forward (OL12882) and reverse (OL12884) primers, 2 µl genomic DNA and 7.2 µl MQ water was made up in the amounts in Table 2.4 for each 20 µl PCR reaction. The PCR was run for initial denaturation for 3 minutes at 94 °C, then cycled 45 times at 94 °C for 30 seconds, 60 °C for 30 seconds and 72 °C for 60 seconds, then run for a final extension at 72 °C for 10 minutes. The PCR product was run on a 1% agarose gel at 100 V for 40 minutes to check product size (approximately 2 kB for successful insertion).

Barcodes 1-10 were sequenced by Sanger sequencing (Eurofins). To amplify the PCR product, the same primers were used, but the PCR was performed with Q5 (New England Biolabs: M0491L) instead. The 40 µl PCR reaction comprised 1 unit Q5, 1X Q5 reaction buffer (NEB: M0491L), 2 µl genomic DNA, 2 mM dNTPs, 0.2 µM forward primer OL12882 and reverse primer OL12884 and 26.8 µl MQ water. The PCR was run for 2 minutes at 98 °C, then cycled 40 times for 30 seconds at 98 °C, 30 seconds at 60 °C and 60 seconds at 72 °C, then extended for 10 minutes at 72 °C.

2.7 *In vitro* infection of bone marrow macrophages

To generate bone marrow macrophages for *in vitro* assays, the bone marrow was flushed from the femur and tibia of B6 mice and pooled, before centrifugation at 1200 rpm for 5 minutes. Cells were counted using Trypan blue (ThermoFisher: 15250061) and a haemocytometer, and re-suspended in DMEM (macrophage differentiation media) at 5x10^5 per ml, before plating and incubating at 37 °C. Media was changed at day 3 and swapped for non-differentiation complete DMEM at day 6, before resting for a further 1-2 days. Cells were then frozen for future use or used immediately for *in vitro* experiments.

To assess infection burdens *in vitro*, murine bone marrow macrophages were plated at 1x10^4 cells per well of a chamber slide and rested overnight at 37 °C, before infecting with stationary promastigote parasites at 5X multiplicity of infection (MOI). Slides were incubated for either 1 hour or 3 hours at 37 °C. Slides were fixed in methanol for 10 minutes at room temperature, then stained with 10% Giemsa stain (VWR Chemicals: 352603R), diluted in water, for 3 minutes, rinsed and dried, before counting parasites per cell for 100 cells per well.

2.8 Macrophage lysomomal-associated membrane protein 1 (LAMP1) staining and analysis

To stain for LAMP1, murine bone marrow macrophages were plated at a density of 5x10^4 per well of a chamber slide in DMEM and incubated for 24 hours at 37 °C, then infected with promastigote parasites at late-log stage at 5X MOI. Slides were

incubated at 37 °C for the specified infection time, before washing with 1X PBS. Wells were fixed and permeabilised with acetone for 10 minutes at room temperature, then blocked for 1 hour at room temperature in blocking buffer, consisting of 5% rat serum, 0.05% BSA (Sigma-Aldrich: A7906-100G) in 1X PBS. LAMP1 AF647 (Ebioscience: 51-1071-82 [discontinued]) at 2.5 µg/ml, diluted in blocking buffer, was added and incubated for one hour at room temperature. IgG2a AF647 rat was used as a control. Slides were then washed with 1X PBS 3 times and incubated with YOYO-1 (Fisher Scientific: Y3601) diluted 1:500 in 1X PBS for 30 minutes, before washing 3 times with PBS and mounting with Prolong Gold (Invitrogen: P36930). Cover-slipped slides were stored at 4 °C in the dark prior to imaging on Zeiss Axio Observer 7 Fluorescent Microscope at 63X.

Images were analysed in TissueGnostics Strataquest software to detect parasites overlaying LAMP1 (AF647) staining. Host cell nuclei were defined via YOYO-1 staining, with anything above 80 μm^2 defined as definite positive, below 20 μm^2 as definite negative and anything between the two values with strong staining also defined as positive. YOYO-1 mean intensity threshold was set at 26. Parasites were defined as YOYO-1 staining with radius between 0.5 and 5 μm^2. LAMP1+ parasites were defined as those overlaying AF647 staining above a threshold of 30. The analysis pipeline was developed with assistance from Helen Ashwin. The quantified counts of LAMP1+ parasite vacuoles per cell were used for further analysis in R.

2.9 Barseq DNA extraction, PCR amplification and sequencing

Tissue samples for liver, spleen, lymph node, gut, lung and brain were flash frozen using dry ice immediately after dissection and were stored at -80 °C until use. Femurs and tibias were cleaned to remove skin, muscle and fat tissue, before flushing out bone marrow using un-supplemented RPMI and a 26G needle. Bone marrow was then centrifuged at 3100 rpm for 10 minutes at room temperature. Supernatant was removed and the cell pellet was frozen at -20 °C until DNA extraction.

DNA extraction was performed on approximately 25 mg samples of larger organs, or 10 mg for spleen, or on whole individual lymph nodes or entire bone marrow cell

pellet, using Qiagen DNeasy Blood and Tissue Kit (Qiagen: 69506) following supplier instructions for tissue DNA for all tissues except bone marrow and cultured parasites; these samples followed cell culture DNA extraction instructions. For all samples the final DNA elution was in 50 µl of buffer AE.

Barseq amplification was done using a two-step PCR protocol consisting of Barseq-PCR1 and Barseq-PCR2 (Figure 2.1). Clean up of PCR products between the two stages and after Barseq-PCR2 was done using either Qiagen QIAquick PCR Purification Kit (for smaller sample batches) (Qiagen: 28104) or with 0.9X Agencourt AMPure XP beads (Beckman Coulter: A63882) (for high-throughput sample processing, with protocol as described below). PCRs were performed using 1 unit VeriFi Mix (PCR Biosystems: PB10.43-05), 0.4 nM 10 nm gold nanoparticles (AuNP) in citrate suspension (Sigma-Aldrich: 741957), 0.025 mM MgCl (Sigma-Aldrich: M1028), 0.4 µM primers OL12833 and OL13833 (Barseq PCR1) or OL14195 and OL14196 (Barseq-PCR2) and PCR grade water. For Barseq-PCR1, 5 µl DNA was added, whilst for Barseq-PCR2 10 µl DNA was used, with water volume adjusted accordingly. PCR cycle conditions are detailed in Table 2.2 for both PCRs.

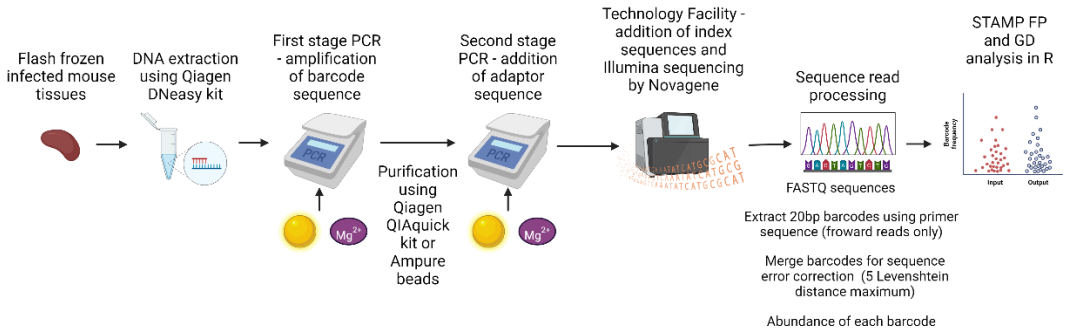


Figure 2.1: Barseq PCR protocol for population dynamics analysis using barcode abundance.

Barseq-PCR1			Barseq-PCR2		
Temperature (°C)	Time	Cycles	Temperature (°C)	Time	Cycles
98		1	98		1
98	30 sec	30	98	30 sec	6
60	30 sec		65	30 sec	
72	40 sec		72	40 sec	
72		1	72		1
12		Hold	12		Hold

Table 2.2: PCR reaction conditions for Barseq PCR 1 and 2.

PCR amplification was verified by running on a 2% agarose gel for 23 minutes at 100 V to check product size. The final PCR product (after both PCRs were run) was then purified by adding Agencourt AMPure XP beads (Beckman Coulter: A63882) at a ratio of 0.9X to PCR volume. The mix was incubated for 5 minutes at room temperature, before placing on a magnetic rack. Once the sample was clear, supernatant was removed and the remaining beads were washed twice with 75% ethanol. The beads were then dried for around 3 minutes before eluting in 32 µl Molecular Biology Grade Water (ThermoFisher: AM9932) for 5 minutes at room temperature. The samples were placed back on the magnet and allowed to clear before pipetting the DNA (supernatant) into a plate for storage at -20 °C prior to Illumina sequencing.

The samples were then used to generate a library for sequencing by Sally James, Lesley Gilbert and Sam Donninger in the University of York Technology Facility Genomics Team. They performed a further PCR with Nextera XT Indexing primers, with 9 PCR amplification cycles, using Q5 Polymerase 2X Master Mix (New England Biolabs: M0492L) according to manufacturer instructions. The samples were purified with Agencourt AMPure XP beads (Beckman Coulter: A63882) at a 0.9X ratio and eluted into low TE buffer. Pooled samples were sent for 150 base paired end sequencing on an Illumina Sequencer at Azenta Life Sciences.

2.10 Tissue section staining

Tissue sections were stored in OCT embedding matrix (Fisher Scientific: 15212776) at -80 °C after mouse harvests. Sections of 10 µm thickness were cut using a Cryostat and allowed to dry before fixing in ice cold acetone for 10 minutes. Slides were then washed twice for 5 minutes in wash buffer, comprising 1X PBS and 0.05% BSA (Sigma-Aldrich: A7906-100G) (w/v). Slides were blocked for 30 minutes at room temperature in dilution buffer, comprising wash buffer plus 5% rat serum (v/v) and kept in darkness thereafter. Primary antibodies were incubated for 1 hour at room temperature. For neutrophils, Ly6G [1A8] AF647 (5 µg/ml [stock diluted 1:100 in dilution buffer]) (Biolegend: 127610) was used, and for macrophages, F4/80 [BM8] AF594 (5 µg/ml [stock diluted 1:100 in dilution buffer]) (Biolegend: 123140) was used. Isotypes were both rat IgG2a on the relevant fluorophore. Slides were washed thrice in wash buffer for 5 minutes each, then incubated with YOYO-1 (Fisher Scientific: Y3601), diluted 1:1000 in 1X PBS, for 45 minutes at room temperature, then washed again three times for 5 minutes each, in wash buffer. Slides were further washed in 1X PBS twice, then mounted using Prolong Gold (Invitrogen: P36930), before cover-slipping and storing at 4 °C prior to imaging on the Zeiss AxioScan Z1 Slide Scanner at a resolution of 20X.

Image analysis was performed in QuPath, to count numbers of F4/80 (AF594) and Ly6G (AF647) positive cells and total cells (YOYO-1 staining). Thresholds were applied to assess fluorophore intensity, to classify cells as macrophages (AF594), neutrophils (AF647) or negative (other YOYO-1-positive nuclei). The thresholds were

defined as low-medium (1000) and high (2000) expression, quantified as cytoplasmic mean expression of AF647 and cytoplasmic maximum expression for AF594. Subsequent analysis was performed in R, to determine the proportion of total cells expressing each marker.

2.11 ELISA quantification of serum concentration of Type 1 IFN

Blood serum samples were obtained via cardiac bleeding of mice, with blood samples collected in SST Blood Collection Tubes (BD: 365968). Blood was left at 4 °C for a minimum of 6 hours before centrifugation at 5000 g for 5 minutes to isolate sera. Sera was stored at -20 °C or -80 °C prior to further analysis. ELISA assays were performed by Natalie Prow, using the Mouse IFN-alpha/beta R2 ELISA kit (Invitrogen: EM39RB) as per manufacturer instructions. Plates were imaged using the SpectraMAX 340 (Molecular Devices) and IFN concentration was calibrated using a standard curve. Subsequent analysis was performed in R.

2.12 Immuno-phenotyping of liver cells by flow cytometry

Whole mouse livers (minus segments removed for other tissue processing) were processed to single cells as follows. Livers were digested for 20 minutes at 37 °C in a buffer comprising 0.01% collagenase (Fisher Scientific: 10780004), 0.0001% DNase (VWR: A3778.0050), 0.02% BSA (Sigma-Aldrich: A7906-100G) and 10 ml HBSS (Fisher Scientific: 12549069) per sample, before adding 10% FCS RPMI to end the reaction. Livers were then mashed, centrifuged for 6 minutes at 1300 rpm and the pellet was re-suspended in 33% Percoll (VWR International: 17-0891-01) in PBS. This was centrifuged at 1700 rpm for 15 minutes with no brakes, then the pellet was re-suspended in ACK lysis buffer (Fisher Scientific: 11509876) and incubated for 5 minutes before washing in 2% FCS in PBS and centrifuging at 1300 rpm for 6 minutes.

Cells were stained using the markers and antibodies in Table 2.3. Approximately 1-1.5x10^6 cells were used per liver. Cells were centrifuged at 1700 rpm for 1 minute at 4 °C (for all subsequent steps), then re-suspended in PBS and centrifuged again.

Cells were stained with live/dead marker for 10 minutes before re-suspending in MACS buffer and centrifuging. Rat IgG (5 µl) was added to the edge of each well. Cells were then incubated with the antibody cocktail or with FMOs or single stained for controls, for 20 minutes at room temperature, then re-suspended in MACS buffer and centrifuged again. Cells were re-suspended in MACS buffer, centrifuged, then re-suspended a final time. Cells were run on the CytoFLEX LX (Beckman Coulter) and analysed using FlowJo software, with gating strategy illustrated in (Figure 6.14). Subsequent analysis was performed in R.

Marker	Antibody Supplier and Cat Number	Clone	Fluorophore	Dilution	Channel
F4/80	Biolegend: 123108	BM8	FITC	1:200	B530/30
CD20	Biolegend: 150422	SA275A11	PerCPCy5.5	1:50	B710/50
Ly6G	BD Biosciences: 551461	1A8	PE, DsRed	1:100	YG586/15
NK1.1	Biolegend: 108716	PK136	PECy5	1:20	YG670/30
Ly6C	Biolegend: 128018	HK1.4	PECy7	1:300	YG780/60
CD11b	Biolegend: 101212	M1/70	APC	1:100	R670/14
CD11c	Biolegend: 117324	N418	APCCy7	1:20	R780/60
CCR2/CD192	Biolegend: 406410	SA203G11	BV421	1:50	V450/50
Live/Dead	Biolegend: 423102	NA	LDAqua	1:1000	V525/50
CD3	Biolegend: 100232	17A2	BV785	1:50	V780/60

Table 2.3: Markers, antibodies and channels for immuno-phenotyping by flow cytometry.

2.13 *In vitro* complement-mediated parasite lysis assay

Promastigote parasites were plated at 1x10^6 parasites per well incubated in a serial dilution of the relevant pooled serum (taken via cardiac bleed) from male or female C57BL/6J or B6.129S4-C3tm1Crr/J mice, with or without prior heat inactivation, for 1.5 hours at 37 °C. Serum was serially diluted in complete SDM-79, with a range of 50% serum to 1% serum, alongside a no serum added negative control. A positive control using a serial dilution of human serum (Sigma Aldrich: H4522-100ML; Source SLCO2789) was performed alongside. The assay was performed in triplicate.

Parasites were then scored by a blind marker alongside visual scoring by the author, using a light microscope and the following scale: 1= all dead; 2= a small number of live cells, mostly dead; 3= approximately half live, half dead; 4= most alive but some dead or unhealthy appearance; 5= almost all live and healthy appearance.

For further live/dead assessment via flow cytometry, triplicate parasites for the highest serum concentrations (alongside controls) were pooled, washed in PBS at 1700 rpm for 5 minutes at 4 °C, then stained with LDAqua (Biolegend: 423102) for 10 minutes at room temperature. Parasites were washed in MACS buffer (consisting of 1X PBS with 10% 0.5M EDTA and 0.1% heat-inactivated FCS) at 1700 rpm for 3 minutes, then fixed using a cell fixation kit (BD Biosciences: 554714) for 20 minutes at room temperature. Parasites were then washed a further two times in MACS buffer as before, and left overnight at 4 °C before analysis using the CytoFLEX LX (Beckman Coulter). All steps for after the staining were performed whilst avoiding light exposure. Initial analysis to gate and quantify marker expression was performed using FlowJo software, with subsequent analysis performed in R.

2.14 Population Dynamics Analysis

2.14.1 Barcode extraction

A Rust script was used to extract barcodes from Illumina sequence Fastq files. Barcodes were extracted as the 20 bp following the sequence ATAATACGACTCACTATAAAACTGGAAG in the forward reads for each sample, and

the number of reads with each barcode sequence were counted. In order to account for sequencing errors, which may otherwise have resulted in slight alterations to the captured barcode sequences and thus erroneous total counts for the abundance of each canonical barcode, any barcode sequences which had 500 or fewer reads were merged with barcodes above this threshold if they had a Levenshtein edit distance of 5 or fewer. If any barcode was close enough in sequence to multiple barcodes above the threshold (which should not have occurred in true barcode sequences, based on the barcode designs), the lowest edit distance choice was merged. Unmatched sequence reads (those which detected no barcode) and those which did not merge were quantified for each sample for quality checking. This merging process was performed individually for every sample analysed. The canonical barcode list was then used to select the true barcodes and counts from this list for each sample, and this was used for further analysis. Barcodes which failed on insertion were included in the list, and consistently showed 0 or negligible reads in INPUT samples. The script for the analysis was developed with assistance from Alastair Droop, in the University of York Technology Facility Data Science Team, and is available in Appendix Script 1 and 2.

To check if the PCR and Illumina sequencing protocols showed replicable results, the PCR amplification of the INPUT sample was repeated twice for each experiment. In the two exemplar experiments here (PRO and AMA), this PCR was then split into two (PRO 14 and 28 days p.i INPUT) or 3 (PRO 2 days p.i INPUT and AMA INPUT). The subsequent read counts for each canonical barcode for each sequenced sample are shown in Figure 2.2A and indicate high consistency across both PCR and sequencing, with slightly increased spread of read counts in AMA compared to PRO datasets; overall this was consistently the case across datasets.

The total canonical barcode reads per *in vivo* sample were broadly consistent within each experiment, with a small number of samples showing lower reads; these predominantly consisted of liver tissue samples (Figure 2.2B). In these exemplar experiments the *in vivo* samples showed consistency in both unmatched sequence reads and in canonical barcode proportion. For PRO datasets, the median unmatched read percentage was 5.32%, with a median of 96.11% of all barcodes extracted matching canonical sequences. Only a single sample had a value above

10% for unmatched reads. For the AMA dataset, the median unmatched read percentage was 3.45% and the median for canonical barcodes across samples was 96.42%. Similar values were found when assessing other datasets.

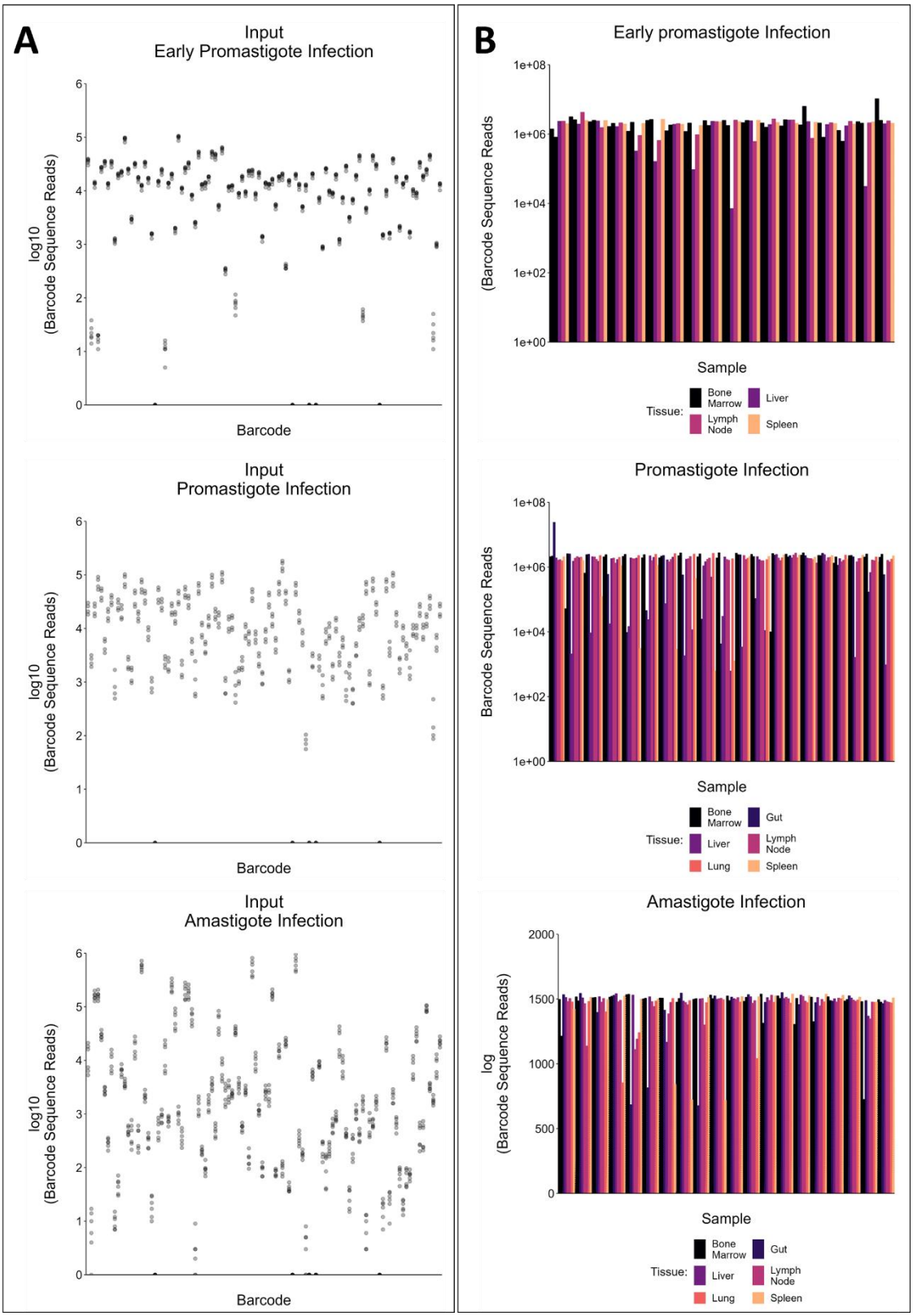


Figure 2.2: Sequence reads output for PRO and AMA experimental sample sequencing using Barseq-STAMP PCR and Illumina sequencing. A). Distribution of barcode read counts in INPUT samples for Early Promastigote (2 day p.i PRO), Promastigote (14 and 28 days p.i PRO) and AMA experiments. Each INPUT pool of parasites was amplified using the Barseq-PCR protocol as two separate reactions, which were then split into either 3 samples (Early

Promastigote and Amastigote) or 2 samples (Promastigote) each for Illumina sequencing. INPUT n = 6, 4 and 6 respectively. Grey or black indicates independent PCRs and data points indicate independent sequence runs. B). Total canonical barcode sequence reads per *in vivo* tissue sample, coloured by tissue type, for PRO and AMA experiments. Each bar represents a single tissue from each mouse.

2.14.2 Barcode processing

All barcodes with below 500 reads in the input sample for each experiment were removed prior to subsequent analysis. Barcode 59 (sequence: TGGATCTTAGGACGCAACAT) was also removed (from all experimental datasets included in this thesis) prior to subsequent analysis, for reasons discussed in 3.2.4.

2.14.3 Sequence-Tag based Analysis of Microbial Population Dynamics (STAMP) bottleneck and genetic distance analysis

STAMP analysis of Founder Population (FP) and Genetic Distance (GD) were performed as described in Abel et al., 2015, based on the method defined by Krimbas and Tsakas (1971). The functions used in this thesis were custom made in R; scripts are available in Appendix Script 3. FP was quantified as follows:

$$\hat{F} = \frac{1}{k}\sum_{i=1}^{k} \left(f_{i,s} - f_{i,0}\right)^2 / f_{i,0}(1 - f_{i,0})$$

$$FP = \frac{1}{\hat{F} - \frac{1}{S_0} - \frac{1}{S_s}}$$

Where f_i is the frequency of barcode *i,* at input ($f_{i,0}$) or in the tissue sample ($f_{i,s}$), S_0 and S_S are the sequence reads for each sample, and *k* is the total number of barcodes.

GD (D_{ch}) was quantified using the Cavalli-Sforza chord distance (Cavalli-Sforza, 1967) as follows:

$$D_{ch} = \frac{2\sqrt{2}}{\pi}\sqrt{1 - \cos\theta}$$

$$\cos\theta = \sum_{i=1}^{k} \sqrt{f_{P1,i}\, f_{P2,i}}$$

Where f_i is the frequency of barcode *i* in populations P1 and P2 being compared and k is total number of barcodes.

2.14.4 Simulation of bottlenecks for analysis of STAMP accuracy

MATLAB was used for simulated bottleneck data generation as described in 3.2.1, with subsequent analysis of the output data performed in R. Scripts are available in Appendix Script 4. The input barcode distribution was simulated as an even matrix of barcoded individuals. These individuals were then passed through a bottleneck, simulated using a random binomial distribution, based on a probability of passing the bottleneck. This was set as a range of values from 0.0000001 to 0.5. Each individual passing the bottleneck was counted in an output matrix summing the individuals for each barcode. The experiment was repeated for 100 iterations. To test the impact of variance in growth rate, the output matrix count values were varied randomly by a percentage multiplier between 95% and 105% of their initial value. To test the impact of clonal expansion, 1 in 10 barcodes were randomly selected to be amplified by a factor of 10. The count values were used to calculate the FP value as described in 2.13.2, and this was compared to the true count as described in 3.2.1.

2.14.5 Diversity analysis using Shannon Index (SI)

Diversity and population evenness were assessed using the Shannon Index, performed using the Vegan package in R. The Shannon Index was quantified as:

$$-\sum_{i=1}^{S} p_i \ \log_b \ p_i$$

Where p_i is proportion of "species" *i* (in this case barcodes) and *S* is total number of "species" (barcode total) and *b* is the default option (natural logarithm *e*), as defined by Hill (1973).

When assessing the relationship between FP and SI, a linear model with the equation:

$$y_{PRO} = \log x \times e^{-x}$$

$$y_{AMA} = \log x \times e^{-1/x}$$

where *x* is FP and *y* is SI, was fitted. The model was interpreted as the log*(x)* term providing scaling for the upper limit on the maximal diversity possible, which was directly related to the amount of diversity initially colonising the tissue (FP). The e^{-x} or $e^{-1/x}$ term provided the rate at which this relationship decayed as more parasites were added to the system, since each additional individual would impact overall diversity more when few individuals were present, as the new individual proportionally makes up a larger amount of the population.

2.15 Minimum spanning tree (MST) network analysis

To generate the MST for each mouse, a graph adjacency object was generated using GD values for pairwise distances between tissues. This was then resolved into an MST using Prim's algorithm. Clusters were assessed using the cluster_Louvain function. All steps were performed using the igraph package in R.

Harmonic centrality (Freeman, 1979; Marchiori and Latora, 2000) was defined as the mean inverse distance of a vertex to all other vertices.

Betweenness centrality (Brandes et al., 2000) was defined as the number of shortest paths passing through a vertex and was quantified as:

$$\sum_{i \neq j, i \neq v, j \neq v} g_{ivj} / g_{ij}$$

Where g_{ij} is the total number of shortest paths between the vertices *i* and *j* and g_{ivj} is the number of these paths which traverse vertex *v*.

Both calculations were performed using the igraph package. The script was developed with help from Shoumit Dey and is available in Appendix Script 5.

2.16 Assembly process analysis

The distance matrix for phylogenetic relationship between barcodes was based on abundance of each in the INPUT pool for each experiment, as these barcoded lines are isogenic and are all generated from a single clonal line. This pairwise distance matrix was then used to create a pseudo-phylogeny for all barcodes, clustering the

closest taxa. The taxa clusters were used to calculate the beta Mean Nearest Taxon Distance (bMNTD) for pairwise comparisons of the parasite population within each tissue.

This was compared to the values obtained via 1000 random simulated populations. The beta Nearest Taxon Index (bNTI) was defined as the standard effect size (SES) from the null model. If the bMNTD was more than two SES away from the null model this indicates that the difference is not due to random chance and the paired population structures were deemed to show dominance of deterministic processes. If SES > 2, this was taken as evidence of heterogeneous selection, where the two tissues have low similarity in their taxa present. If SES < -2, this indicated homogenous selection, where the two tissues possess very similar taxa. SES between 2 and -2 was taken to indicate stochastic processes were dominant; these pairs were then analysed using the Bray-Curtis Modified Raup-Crick metric (RC_{BRAY}).

The paired populations were compared to determine the likelihood of species occurring in samples from both tissues, with adjustment based on the abundance of the species in question. This was again compared to 1000 random simulations to determine if the value was different to that which would occur due to random chance. If $RC_{BRAY} > 0.95$, this indicated that dispersal limitation was the dominant process, meaning very little overlap in the species present in each sample. If $RC_{BRAY} < -0.95$, this indicated that homogenising dispersal was the dominant force, where populations were intermixing so strongly that species compositions were highly similar. If RC_{BRAY} fell between these two values, it indicated that the populations were no different than could be explained by stochastic drift as the dominant process. In a natural population this could also be the influence of diversification on the population via random mutations, but this was not a factor in the described work due to the known barcode sequences.

For calculation of the pseudo-phylogeny, the Phangorn package in R was used, applying a Canberra algorithm to calculate pairwise distances of each barcode to all other barcodes. For calculation of bNTI and RC_{BRAY} indices, the iCAMP package in R (Ning et al., 2020) was used on the Viking computer cluster for parallel computing of simulations. The script for all analysis is available in Appendix Script 6.

2.17 Statistical analysis

Statistical analysis was performed using R with the rstatix and ggpubr packages. Pairwise Mann-Whitney (MW) tests were used to compare GD, SI and FP, tissue staining cell quantification and for IFN level comparisons, with a Bonferroni correction applied to pairwise comparisons of 3 or more categories per group (e.g. timeline comparisons or those involving three treatment groups), to reduce the risk of a type 1 error occurring due to the multiple comparisons performed. Pairwise Mann-Whitney tests were also used to compare flow cytometry data between infection groups. For comparisons to uninfected control, a Bonferroni correction was applied. For the comparison of re-infected and single infection, no correction was applied as only two groups were compared. Reported p values for tests throughout the following chapters are adjusted if a Bonferroni correction was performed. Comparisons involving two categories per group were compared using a Mann-Whitney test without an adjustment applied. A p-value of less than 0.05 (after adjustment if applied) was taken as significant unless otherwise stated. Correlations were assessed using Kendall correlation, using the cor.test function. Distribution fitting was performed using R, using the fitdistrplus package. Akaike information criterion (AIC) and Bayesian information criterion (BIC) were used to select the best fitting distribution.

2.18 Figure generation

Schematic diagrams were created in Biorender, as indicated in figure legends. Other figures were generated using R, with ggplot2 (within the tidyverse package) except for MST networks, which were plotted using base R.

2.19 Code and data availability

Data files and analysis scripts for Barseq work, network analysis and community ecology analysis are available in the Barseq_thesis repository of github.com/CiaraLoughrey-scientist. Data files and analysis scripts for other analysis

in this thesis which does not fall into those categories is in the repository named Other_thesis in the same Github account.

Chapter 3: Developing a Barseq-STAMP approach to understand *in vivo* bottlenecks in *L. donovani* infection

3.1 Introduction:

L. donovani is one of the two causative species of VL. Whilst it is known that the parasite is able to colonise multiple tissues – most notably the liver and spleen – it is not yet understood how the parasites disseminate to and between these tissues or which factors may influence these tissue colonisation events in patients and in animal models of disease. Gaining an understanding of how the parasite and host interact, and how this shapes the characteristics of both, is crucial for the development of effective treatments and preventative methods. Bottleneck events are one example of an interactive process which can alter parasite population dynamics and influence future microbial evolution and disease progression.

Different methods are available to quantify the restrictiveness of bottlenecks and the degree of inter-relatedness between populations. Krimbas and Tsakas described equations (Krimbas and Tsakas, 1971) to understand insect population dynamics and these were adapted by Abel et al (Abel et al, 2015) for use in microbial systems *in vivo*, in order to understand the bottlenecks and dissemination dynamics of *Vibrio cholerae* in rabbits (Abel et al., 2015). The method, STAMP, involves utilising high-throughput sequencing to measure changes in proportions of large numbers of fitness-neutral barcodes in a pathogen population; by measuring the difference between the input and the tissue resident population at a later time-point, the effective 'founder population' (FP) size can be calculated, giving a relative measure of bottleneck restrictiveness under different conditions. The method has so far been used in many bacterial infection systems, including *Pseudomonas aeruginosa* (Bachta et al., 2020), *Citrobacter rodentium* (Campbell et al., 2023), *Listeria monocytogenes* (Louie et al., 2019; Zhang et al., 2017; Chevée et al., 2024) and *E. coli* (Hullahalli and Waldor, 2021), as well as in a study on *Toxoplasma gondii*, to date the only use of this technique in a eukaryotic pathogen (Wincott et al., 2022). It has also recently been utilised to study the tissue-specific bottlenecks affecting invasion dynamics of the plant pathogen *Ralstonia solanacearum* (Jiang et al.,

2024). Other methods of calculating bottleneck size also exist, such as modelling or probability-based methods, and different population dynamics equations (Abel et al., 2015b; Krimbas and Tsakas, 1971; Nei and Tajima, 1981; Pollak, 1983; Wright, 1950).

Genetic distance (GD) is a related but distinct concept, which compares the overall similarity of two populations (Figure 3.1), providing a relative metric of potential intermixing between the two, which enables insight into dissemination patterns and the methods by which colonisation events occur. It likewise has different methods of calculation available, with Cavalli-Sforza chord distance (Cavalli-Sforza and Edwards, 1967) a popular choice; this was also utilised for STAMP assessment of *in vivo* pathogen populations (Abel et al., 2015).

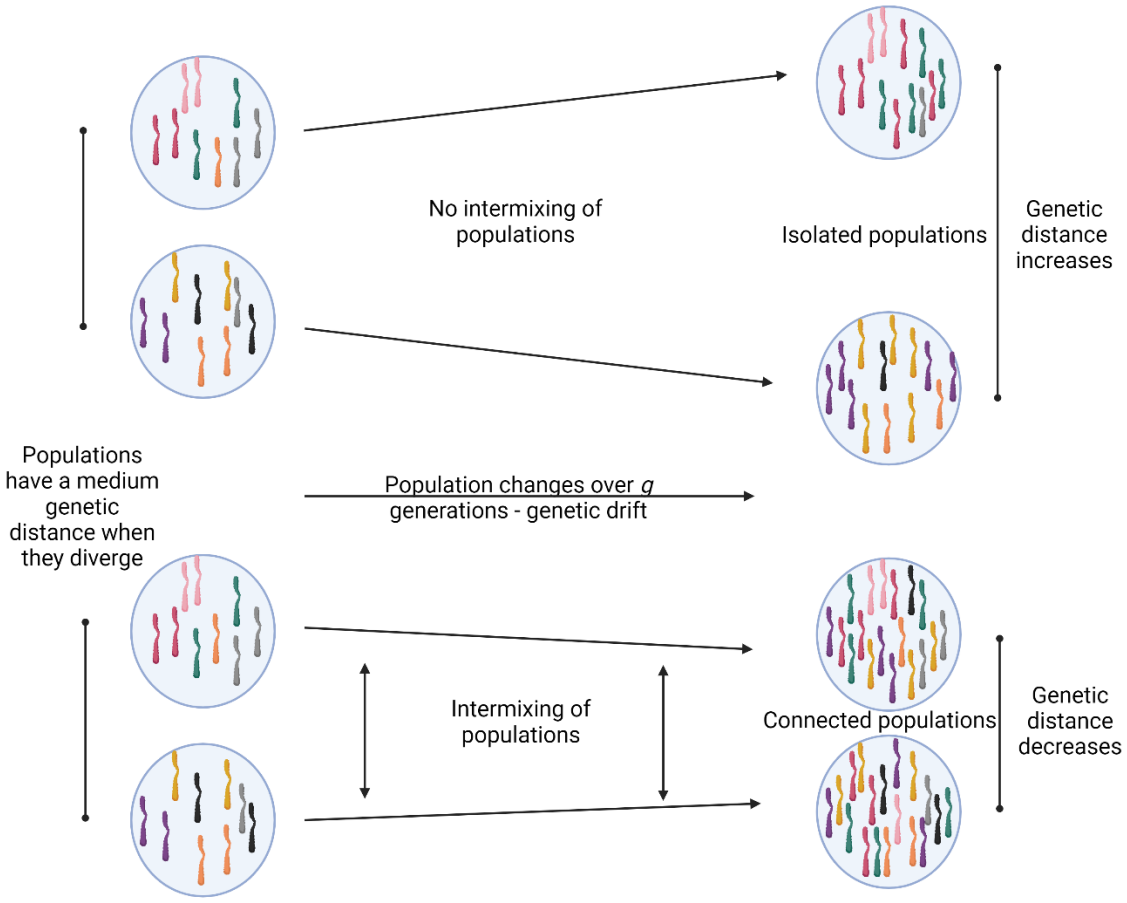


Figure 3.1: Schematic illustration of the concept of genetic distance (GD). GD can change over time as a result of inter-mixing. If populations are isolated, they will become more distant (GD will increase), as genetic drift results in changes in allele abundances. If

populations are intermixing, this will result in them becoming more closely related (GD will decrease). Figure created in Biorender.

Genetic barcoding additionally enables other functional studies of pathogen-host interactions and microbiological phenomena. Barseq is one approach which utilises barcoding, originally developed in yeast (Han et al., 2010; Smith et al., 2009); barcodes are often attached to other insertion sequences, for the deletion or mutation of selected genes on a high-throughput scale. The pool of barcoded mutants are then used for functional studies to assess the relative fitness of different mutants using quantitative barcode analysis in a similar manner to STAMP. This can lead to identification of important pathogen genes for further research. Barseq approaches have been utilised for genetics research in a diverse array of pathogens, including *Plasmodium*, *Toxoplasma* and trypanosome parasites (Bushell et al., 2017; Gomes et al., 2015; Sidik et al., 2016).

Recent advances have led to the development of CRISPR-cas9 genome editing in *Leishmania* parasites, allowing precise gene editing (Beneke and Gluenz, 2020; Zhang et al., 2017; Zhang et al., 2020; Zhang and Matlashweski, 2015). This has enabled a number of studies into *Leishmania* genetics (Carnielli et al., 2022; Jones et al., 2022; Lypaczewski et al., 2021; Potvin et al., 2021), including the utilisation of Barseq approaches in this parasite to investigate the importance of ubiquitination, flagellar proteins and protein kinases in *L. mexicana* (Baker et al., 2021; Beneke et al., 2019; Burge etal., 2020; Damianou et al., 2020). All of these previous studies used barcodes as a method to identify individual mutants within a population over the course of the experiment.

Barcoding has additionally proven useful in a variety of other ways, with rapid further developments particularly focussed on spatial transcriptomic profiling of mammalian progenitor cells and viral vector targets (Ratz et al., 2022; Jang et al., 2023). ‘Zombie’ (an acronym for “Zombie is optical measurement of barcodes by *in situ* expression”) is one such approach, utilising the inducible transcription *in vitro* of barcodes and Hybridisation Chain Reaction (HCR)-based imaging in a variety of systems to spatially track individual barcodes (Askary et al., 2020). Other recent uses for barcoding include imaging and lineage tracking of cancer cells (Quinn et al.,

2021; Rovira-Clavé et al., 2022) and experimental evolution of viruses and bacteria (Cira et al., 2018; Jasinka et al., 2020; Levy et al., 2015; Vasquez et al., 2021).

To date, an investigation into the bottlenecks in *Leishmania* dissemination within host organisms has not been published. The aim of this work was therefore to generate an isogenic barcoded library of *L. donovani* parasites to use for further investigation into bottlenecks and dissemination dynamics within a murine VL model, using a combination of the Barseq and STAMP methodologies.

3.2 Results:

3.2.1 Strategy for library generation

The approach taken to address these aims is illustrated in Figure 3.2. The intention was to use CRISPR genome editing to generate a library of *L. donovani* parasites, each with a unique but otherwise phenotypically neutral barcode. These parasites can then be used for *in vivo* studies to assess the barcode proportions in different tissues and determine the spatial and temporal bottlenecks and dissemination patterns of the parasites in mouse models.

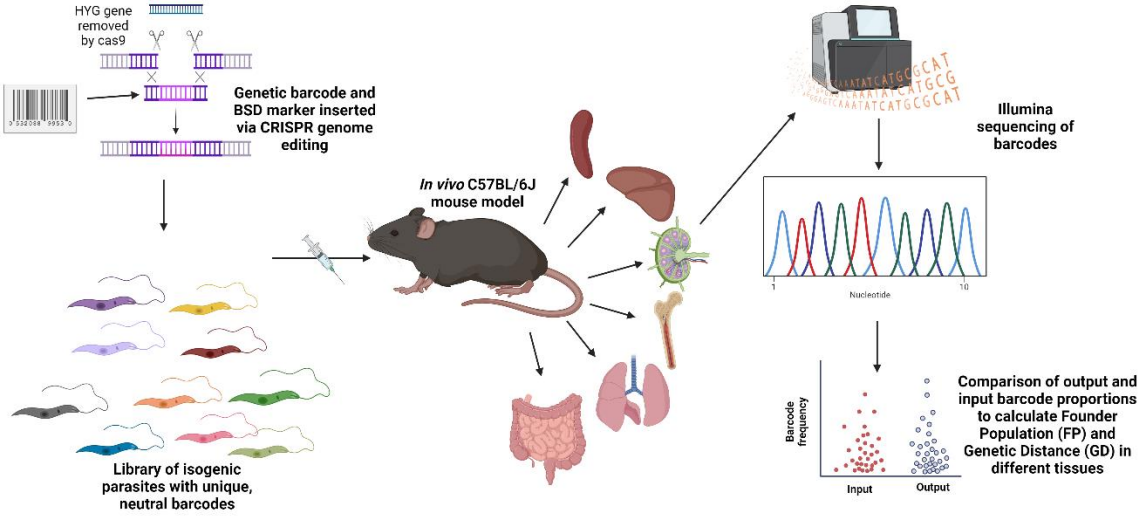


Figure 3.2: Schematic illustrating the experimental methodology for using genetic barcodes in *L. donovani* to investigate parasite dissemination via population dynamics. This figure shows how barcodes can be inserted into parasite genomes using CRISPR-cas9 genome editing, to produce a library of parasites, each with a unique barcode, but otherwise identical. This library can then be used for experimental work in *in vivo* models. For example, population bottlenecks can be assessed in different tissues in a murine model of infection by extracting DNA from these tissues and measuring barcode abundance via Illumina sequencing. The proportions of barcodes in the input and output parasite populations can be quantified and these data can be used to determine founder populations and thus bottlenecks in dissemination. Created with Biorender.

3.2.2 Determination of required library size for bottleneck assessment

To determine the number of barcoded lines necessary for accurate calculation of FP using the STAMP methodology, a simulation was developed using MATLAB (Appendix Script 4). Figure 3.3 shows the conceptual methodology for this simulation.

Briefly, a negative binomial distribution was used to calculate the probability for each individual simulated "parasite" passing through a range of 'bottlenecks', encoded by various probabilities from 10^{-7} to 0.1. Parasites were allocated a "barcode" prior to passing or not passing the bottleneck, and the number of individuals with each barcode was counted after the bottleneck event. These counts were used to calculate an FP estimate, using the STAMP method, for each

experiment. This was repeated 100 times per bottleneck probability, per barcode total. Barcode totals of 70, 105, 140, 210 and 350 were tested. The Kendall correlation between STAMP estimate and actual FP (defined as the number of counted individuals) was calculated for each barcode total and showed that all barcode numbers assessed were highly similar in their ability to accurately calculate FP (Figure 3.4A and B).

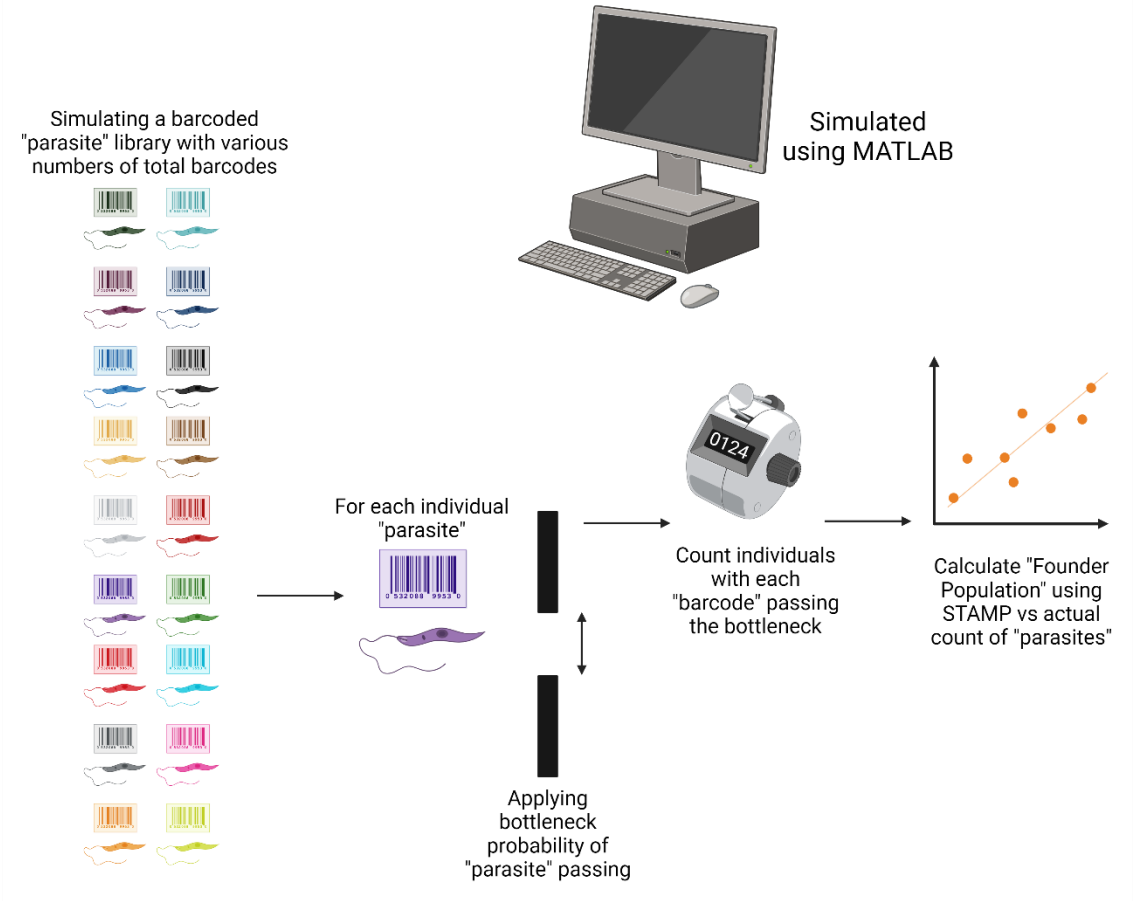
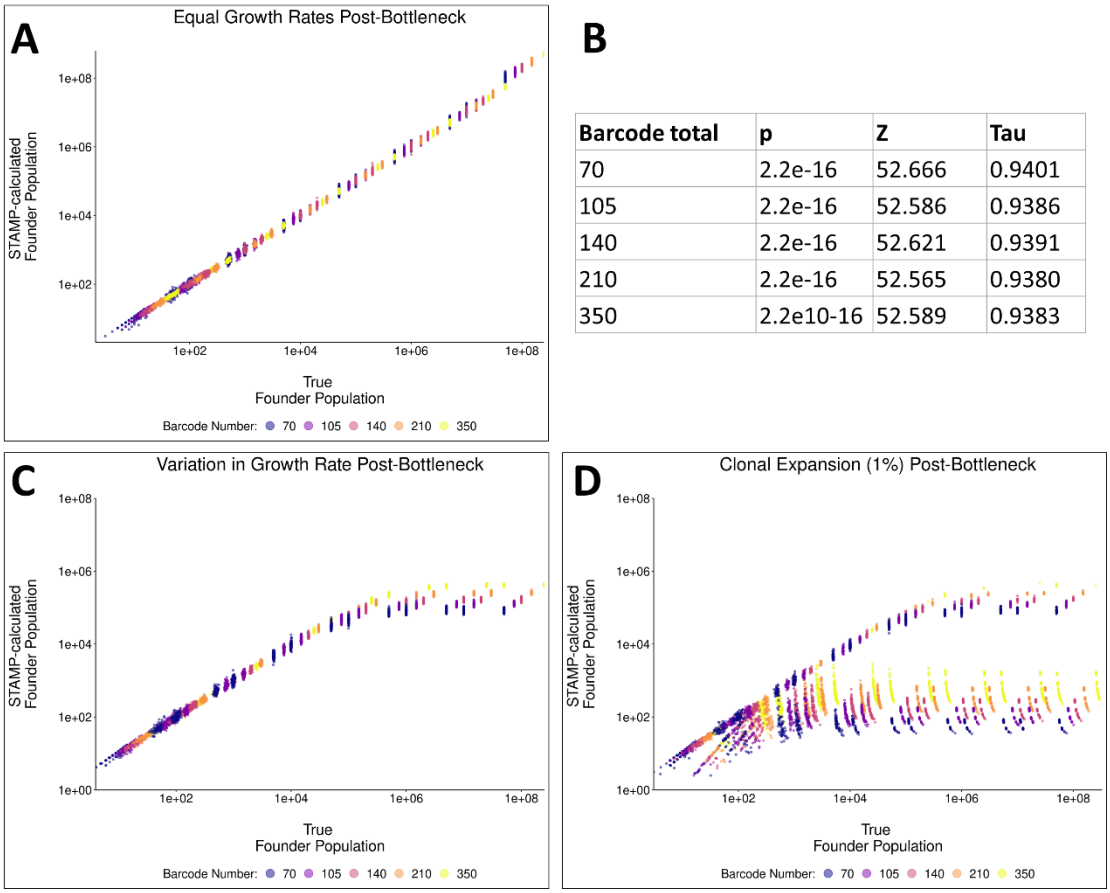


Figure 3.3: Simulation strategy for assessment of required library size. Schematic illustrating the simulation approach taken. "Parasites" are allocated a barcode prior to passing through a "bottleneck", denoted by a probability ranging from 1 in 10^{-7} and 1 in 2 for each individual, with a negative binomial distribution. After passing or not passing the bottleneck, "parasites" of each barcode are counted and these values are used to calculate FP estimates using the STAMP method. Figure partially created in Biorender.

In order to determine how differences in growth affected the accuracy of FP calculation, we varied the simulated growth rate by applying a random increase or decrease in the final count of up to 5% (Figure 4.4C). The results of this simulation indicate that FP can still be accurately calculated up to populations of above 10000 when individual growth rate varies by a small amount. However, for more permissive bottlenecks, the ability to accurately quantify FP is diminished. Similar findings were noted when growth variation was combined with random clonal expansion of 1 in 100 individuals (Figure 4.4C). Again, the number of barcodes used (across the range tested here) did not seem to alter this computational power.



Barcode total	p	Z	Tau
70	2.2e-16	52.666	0.9401
105	2.2e-16	52.586	0.9386
140	2.2e-16	52.621	0.9391
210	2.2e-16	52.565	0.9380
350	2.2e10-16	52.589	0.9383

Figure 3.4: Accuracy of STAMP-calculated FP for a range of simulated barcode library sizes. A). Correlation between true FP and STAMP estimated FP under even growth conditions, for a range of barcode totals between 70 and 350, under the full range of "bottleneck" probabilities tested. B). Kendall correlation parameters for even growth conditions. C). Correlation between true FP and STAMP estimated FP under variable growth rates post-bottleneck. D). Correlation between true FP and STAMP estimated FP when 1 in 100 individuals clonally expand at 10 times the average rate of growth. All plots show a range of barcode totals between 70 and 350, indicated by colour as specified in the key.

This analysis therefore led us to conclude that a library of approximately 100 barcodes was likely to be sufficient for estimation of FP in our VL model. This is similar to the barcode numbers used in some previously published work using the STAMP technique (Chevée et al., 2024; Jiang et al., 2024; Louie et al., 2019; Wincott et al., 2022; Zhang et al., 2017).

3.2.3 Barcode insertion

The protocol for generating our barcoded *Leishmania* parasites has been designed for use in the LV9 T7 cas9 *L. donovani* line, already generated in the Mottram lab prior to this project (method described in Carnielli et al., 2022 for *L. infantum* species, including details of plasmid constructs utilised here). The integration of the cas9 enzyme into this line allows the CRISPR-cas9 genome editing system to be used to insert the barcodes at the desired location. A barcode list was designed (by K. Newling), each of which was 20 bp, had a GC content between 0.4 and 0.6, did not BLAST match to any mouse or *L. donovani* genome sequences, had minimal likelihood of hairpins occurring and was a Hamming distance of 16 or more from all other barcodes to allow effective discrimination using Barseq.

Figure 3.5 shows the integration strategy, which inserts the barcodes into the hygromycin resistance locus (HYG) (Figure 3.5A). This was the selection marker used for the T7 cas9 integration. The insertion was initially done via transfection of parasites with both pGL2662 (encoding blasticidin resistance [BSD]) and pGL2667 (encoding puromycin resistance [PAC]), to attempt to integrate two copies of the unique barcode into each parasite line (Figure 3.5B). HYG is not an endogenous gene in *L. donovani*, and therefore loss of this gene was predicted to have no effect on parasite phenotype, meaning the barcoded line should have the same fitness as the parental line. A T3 promotor was included in the design of the integration cassette, to allow the potential for *in vitro* transcription for barcode amplification via the Zombie technique (Askary et al., 2020); this method was not subsequently used in this thesis but is a possible future use of the barcoded lines for further experimental work.

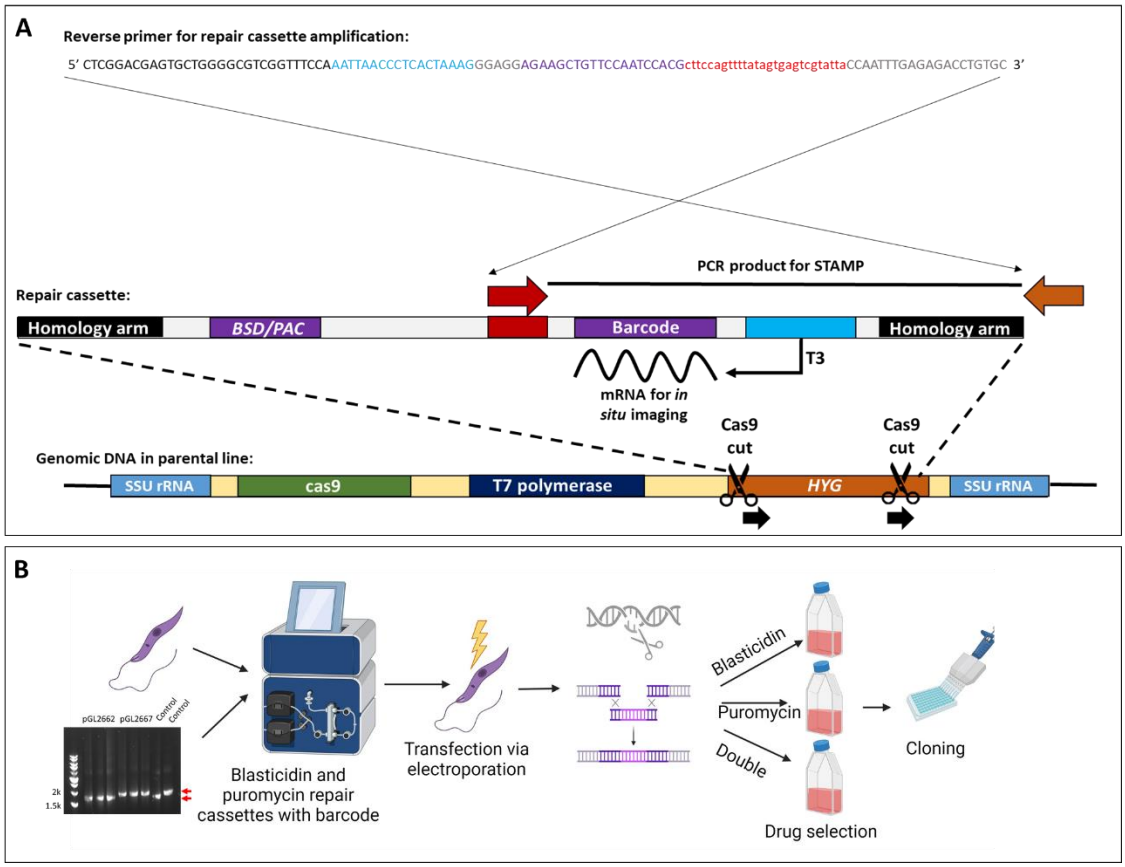


Figure 3.5: Barcoded *L. donovani* library generation strategy. A). Integration strategy. The requirement for a neutral insertion locus to minimise the possibility of barcode insertion impacting on parasite fitness meant that the hygromycin B resistance locus (HYG) (orange) in the previously integrated cas9 cassette, located in the SSU rRNA locus of the parasite genome, was chosen. Hygromycin B was used for selection of the parasite line expressing cas9 following transfection. Production of cas9 is enabled by addition of T7 promoters in the sgRNA, resulting in T7 polymerase (dark blue) transcription. The repair cassette for barcode insertion utilises either blasticidin (BSD) or puromycin (PAC) resistance (purple) for selection, and contains a unique 20 bp barcode (grey) and a T3 promoter in reverse orientation (turquoise), for possible *in situ* imaging using Zombie (see Askary et al., 2020). Additionally, a region designed to bind to a forward primer is included (red), with the reverse primer designed to bind to the remaining section of HYG after integration – these primers amplify the barcode region for Barseq sequencing (Abel et al., 2015a). The homology regions for the barcode insertion are shown in black; these are located within the HYG locus. Cas9 cut sites are also illustrated. The barcoded DNA fragment is then used as the repair template for CRISPR cas9 genome editing of the hygromycin resistance locus in the parasites. B). Schematic of double-selection transfection strategy using CRISPR genome editing to insert barcodes. Panel B created in Biorender.

A PCR was designed to check for successful integration of the repair cassette containing the barcode sequence (Figure 3.6A). Following analysis of the first two barcode insertions, which utilised double selection, it was noted that the resulting parasites did not show the expected integration product on PCR analysis and instead appeared to have either a larger integration product (in the case of Barcode 1 [B1; see Appendix Table 1 for details]) or no product detected by the PCR (in the case of Barcode 2 [B2; see Appendix Table 1 for details]) (Figure 3.6B). Subsequent genome sequence data indicated that only a single copy of HYG was present in the parental line, meaning integration of two barcodes via double selection would not be possible.

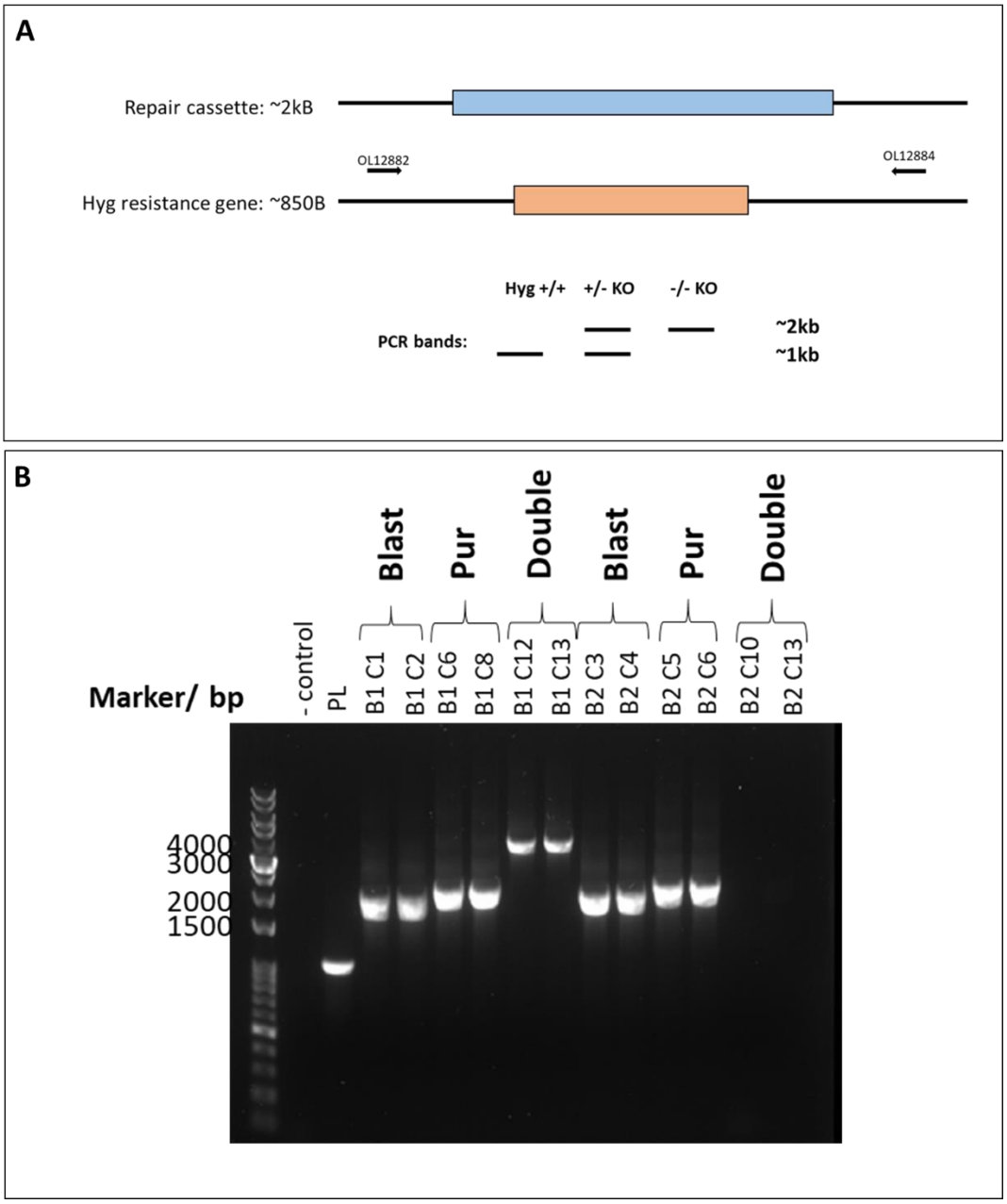


Figure 3.6: Initial barcode insertion results using single and double selection. A). Strategy for validation of barcode insertion by PCR. The primers are designed to bind outside of the integration region (i.e. around the *HYG* locus) and will amplify either the repair cassette insertion or the parental line hygromycin resistance gene. Since the repair cassette is approximately twice as long, the size differential can be used to determine whether the barcode has been inserted or not. B). Gel images (1% agarose gel) of PCR results in single and double selected lines. B1 and B2 refer to Barcode 1 and Barcode 2, as detailed in Appendix Table 1. – control refers to negative (water) control. PL refers to parental line. C refers to the clone ID. Blast indicates selection using blasticidin; pur indicates selection using puromycin; double indicates selection using both.

Attempts to use a pooled transfection approach were also unsuccessful. Therefore, the strategy used for generating the full library was to perform individual transfections for each barcode insertion, with single selection using blasticidin (Figure 3.7A). Transfection and cloning of 8 further barcoded lines was performed. These lines were validated by PCR (Figure 3.7B) and Sanger sequencing, showing successful barcode integration, alongside the T3 promotor sequence (Figure 3.7C). This strategy was then utilised for generating an additional 92 barcoded lines to complete the barcoded *L. donovani* library and these lines were validated by PCR. Four barcodes failed to generate parasite lines. The barcoded lines are referred to by the designations B1-106, as detailed in Appendix Table 1.

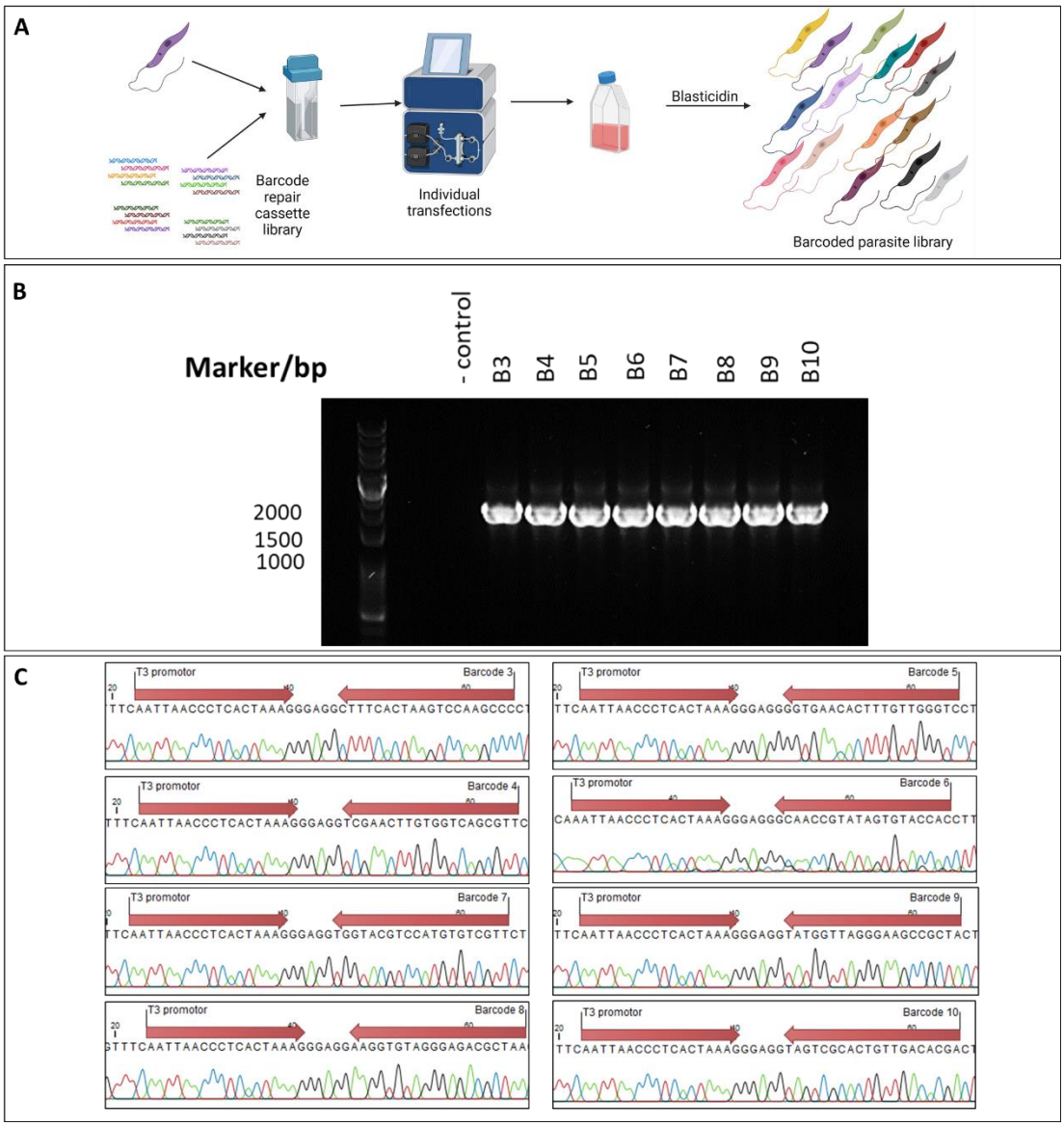


Figure 3.7: Generation and validation of a pool of 102 isogenic barcoded *LV9 L. donovani* lines. A). Strategy for genome editing using CRISPR-cas9 and electroporation, using single selection with blasticidin. B). Gel (1% agarose) images of PCR validation of eight further barcoded lines (B3-B10; see Appendix Table 1 for details). – control refers to negative (water) control. C). Sanger sequencing results showing barcode insertion and the T3 promotor sequence (pink arrows) for all eight barcodes. Analysis was performed using CLC. Panel A created in Biorender.

3.2.4 Validation of barcode neutrality *in vitro*

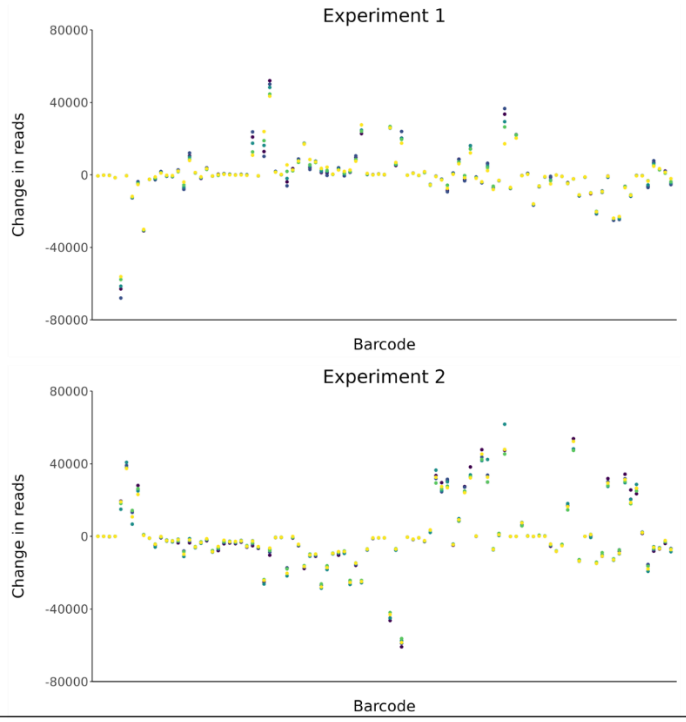
To test whether the barcodes were indeed neutral in terms of parasite fitness, the library of 102 barcoded lines was pooled and grown *in vitro* for seven days (n=5 replicates per experiment). The input DNA was purified, then the barcode was

amplified by PCR, and was sequenced via Illumina. The same protocol was followed for the five replicates, and the change in barcode abundance was assessed by calculating the difference in output versus input reads (normalised to total reads per sample) for each barcode. The data from two independent experiments is shown in Figure 3.8A and these were combined for analysis (Figure 3.8B).

Across both experiments, most barcodes did not exhibit a trend of increasing or decreasing in abundance by the same amount (Figure 3.8B). To test whether line to line variation was indeed stochastic, the barcode counts were fitted to a distribution. The best fit was found, in both cases, to be a Gamma distribution, via comparison of Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC), after a transformation which made data values absolute and divided by 100 for scaling (Figure 3.8C). Parameters fitted were: Experiment 1: shape = 0.534374227, rate = 0.004754405; Experiment 2: shape = 0.561977795, rate = 0.008490018.

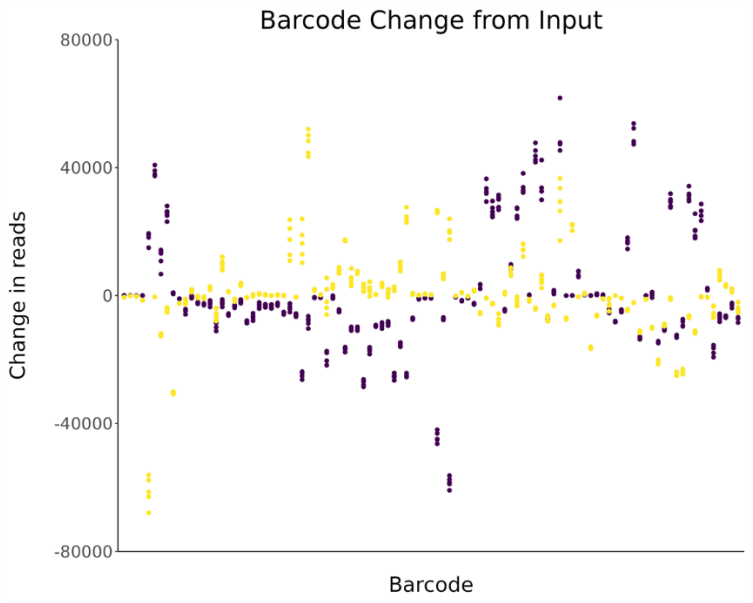
A

In vitro barcode change (Output – Input)



B

In vitro barcode change (Output – Input)



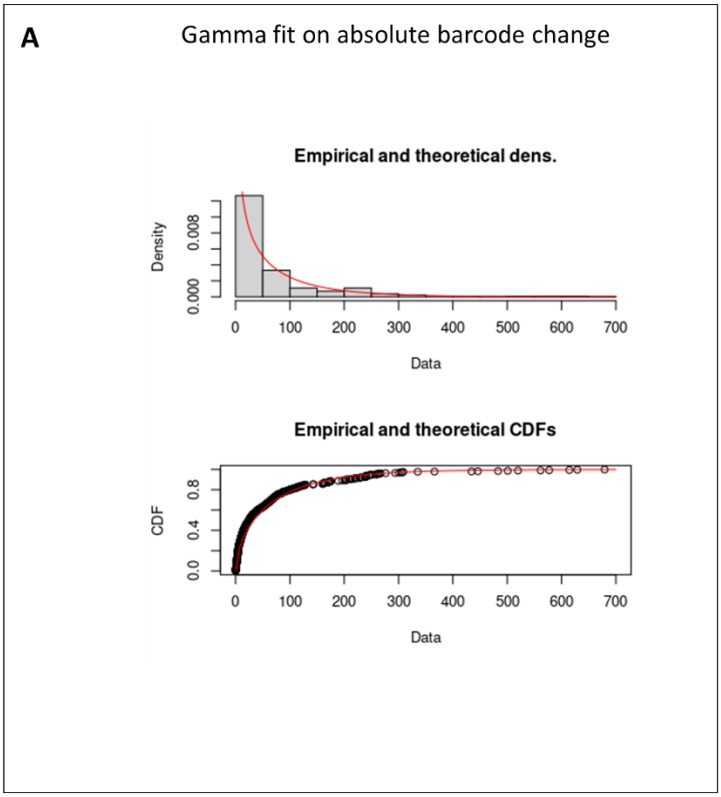
C

Best fit distributions for absolute barcode change (Output – Input)

	Experiment 1		Experiment 2	
	AIC	**BIC**	**AIC**	**BIC**
Negative binomial	9757.166	9765.615	10259.284	10267.733
Normal	10928.358	10936.807	11293.485	11301.934
Log Normal	9736.580	9745.029	2766.389	2774.838
Gamma	5106.252	5114.701	5610.200	5618.650
Logistic	10761.621	10770.070	11226.313	11234.762

Figure 3.8: Validation of isogenic barcoded library growth *in vitro*. A). Change in barcode abundance (output reads minus input reads, normalised to total sample reads) after *in vitro* growth of a pool of 102 barcoded LV9 T7 cas9 *L. donovani* parasites for seven days. Each experiment included 5 replicates (indicated by colour of points). B). Combined change in barcode abundance for both experiments, with individual experiments (all 5 replicates from each) coloured yellow (experiment 1) or purple (experiment 2). C). Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) values for different distribution fits to the experimental data.

When the theoretical and actual values were compared for the experimental data versus the distribution fit, it was found that they showed little difference (Figure 3.9A), indicating that the data fitted a random Gamma distribution in both experiments. When compared directly to a simulated distribution using the fitted parameters (Figure 3.9B), it was again found that there was no obvious difference between the experimental and simulated datasets. Thus the barcode frequency changes *in vitro* appear to be random and not the result of differences in fitness between barcodes.



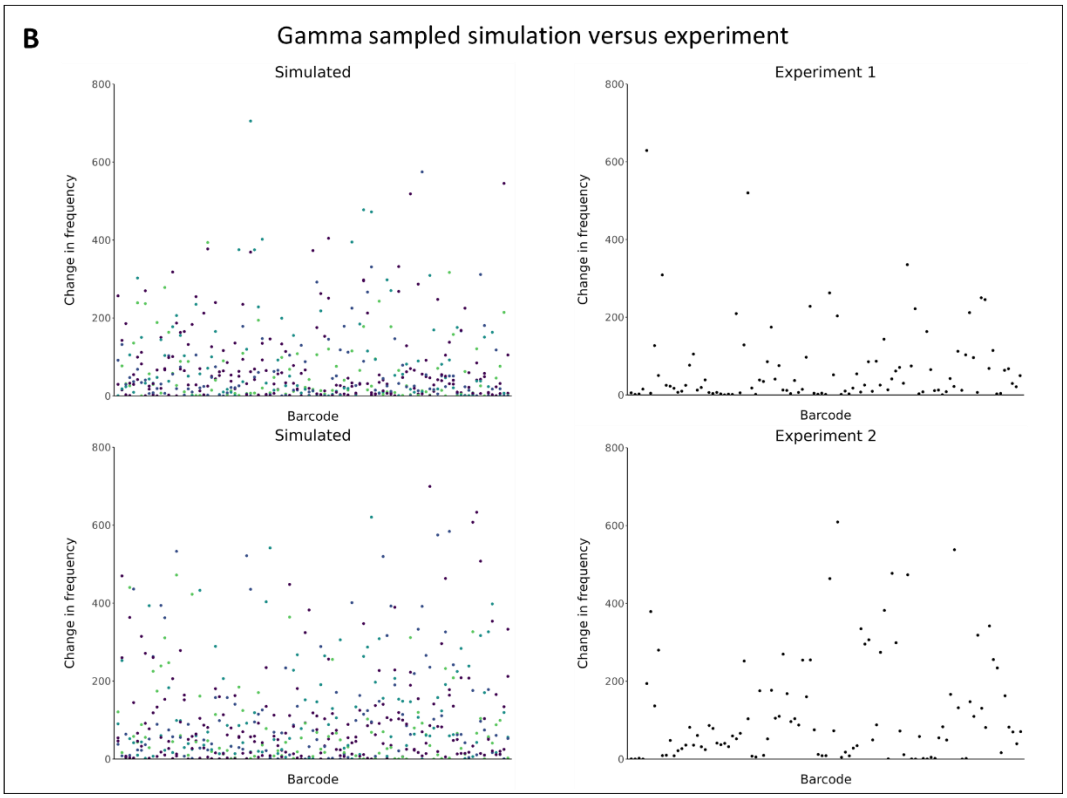


Figure 3.9: Comparison of Gamma distribution to experimental data. A). Empirical and theoretical values for a Gamma fit on the change in barcode counts after transformation (absolute value taken, divided by 100). B). Gamma sampled simulation using shape and rate values for the Gamma distribution fit on each experiment individually (left), and the corresponding experimental data (right) after transformation was applied.

These data provide evidence that the *in vitro* growth of the barcoded line varies in a manner similar to that expected if the lines were of equivalent fitness, meaning the differences observed are likely to be stochastic.

3.2.5 Validation of barcode neutrality in *in vitro* infection of macrophages

To further validate that the barcodes are neutral in terms of parasite fitness, a set of assays was developed to compare the lines to each other in terms of infectivity in macrophages (Figure 3.10).

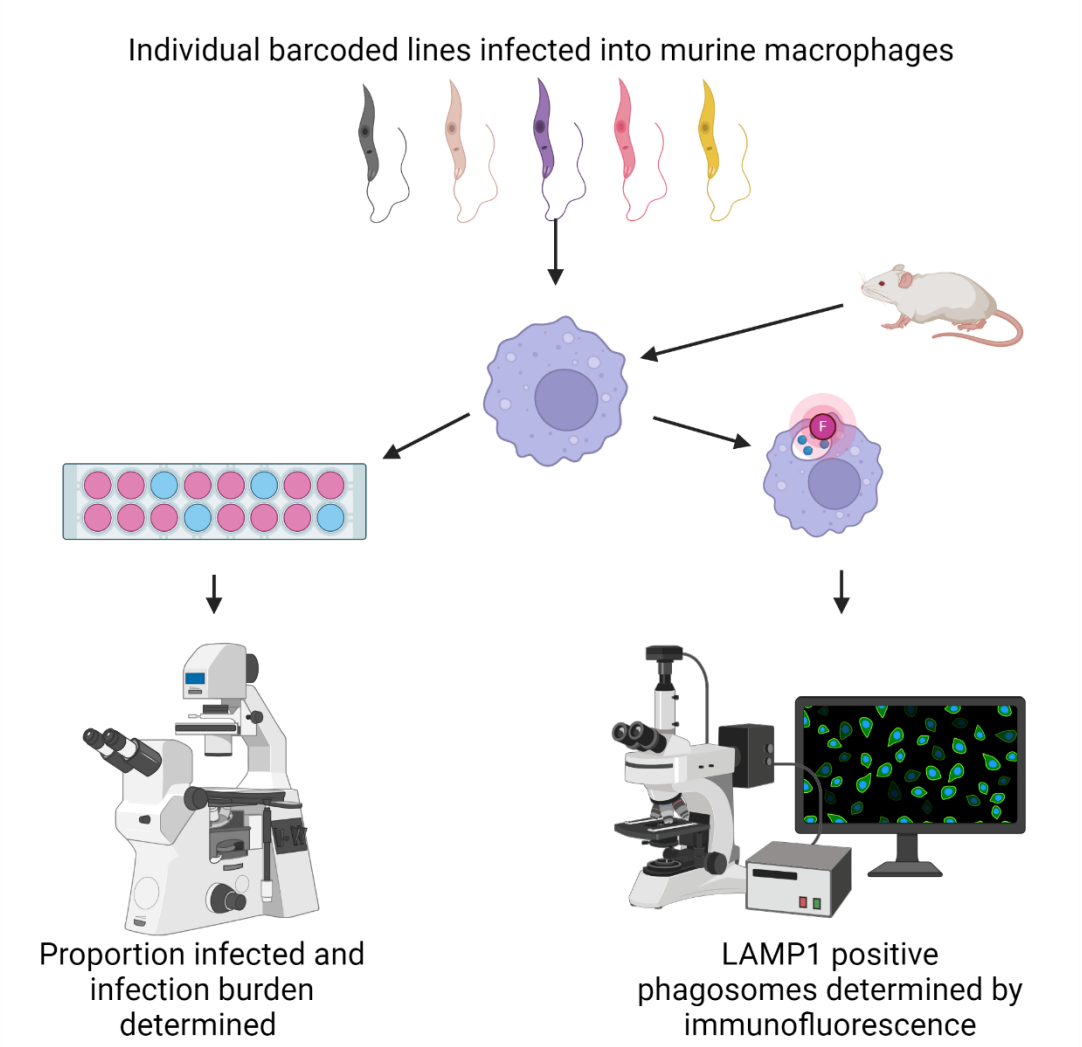


Figure 3.10: Schematic illustration of assays for validation of isogenic barcoded library infectivity *in vitro*. Five barcoded lines were individually infected into bone marrow derived macrophages (BMDMs). Parasite burden per cell was assessed by eye for each line at 1 hour and 3 hours post-infection. The number of parasites residing in lysosomal-associated membrane protein 1 positive (LAMP1+) phagosomes per cell was determined using fluorescence microscopy at 1 hour, 24 hours and 48 hours post-infection. Figure created in Biorender.

A protocol was developed to assess parasite infectivity via measuring parasite burden in murine bone marrow macrophages (BMDMs) using Giemsa staining to visualise the cells and parasites. The lines selected were B1-5 (see Appendix Table 1 for details) and were assessed individually for their ability to infect host cells. Figure 3.11A shows the resulting distributions of parasites per macrophage for the five

lines at 1 hour and 3 hours post-infection. Since pairwise comparisons between lines would be likely to generate type 1 error, we decided to instead compare each line to the average distribution of the five; this would tell us whether each distribution could be stochastic variation within that pooled distribution, and thus whether a line was behaving in a way which was unlikely to be due to random chance. A Poisson distribution was fitted to the distributions for each line, generating lambda values for the distribution of parasites of each barcoded line at each time-point. Figure 3.11B shows lambda values fitted for both time-points. All lambda values at all time-points were within 2 standard deviations of the mean lambda, indicating differences between lines were likely to be stochastic.

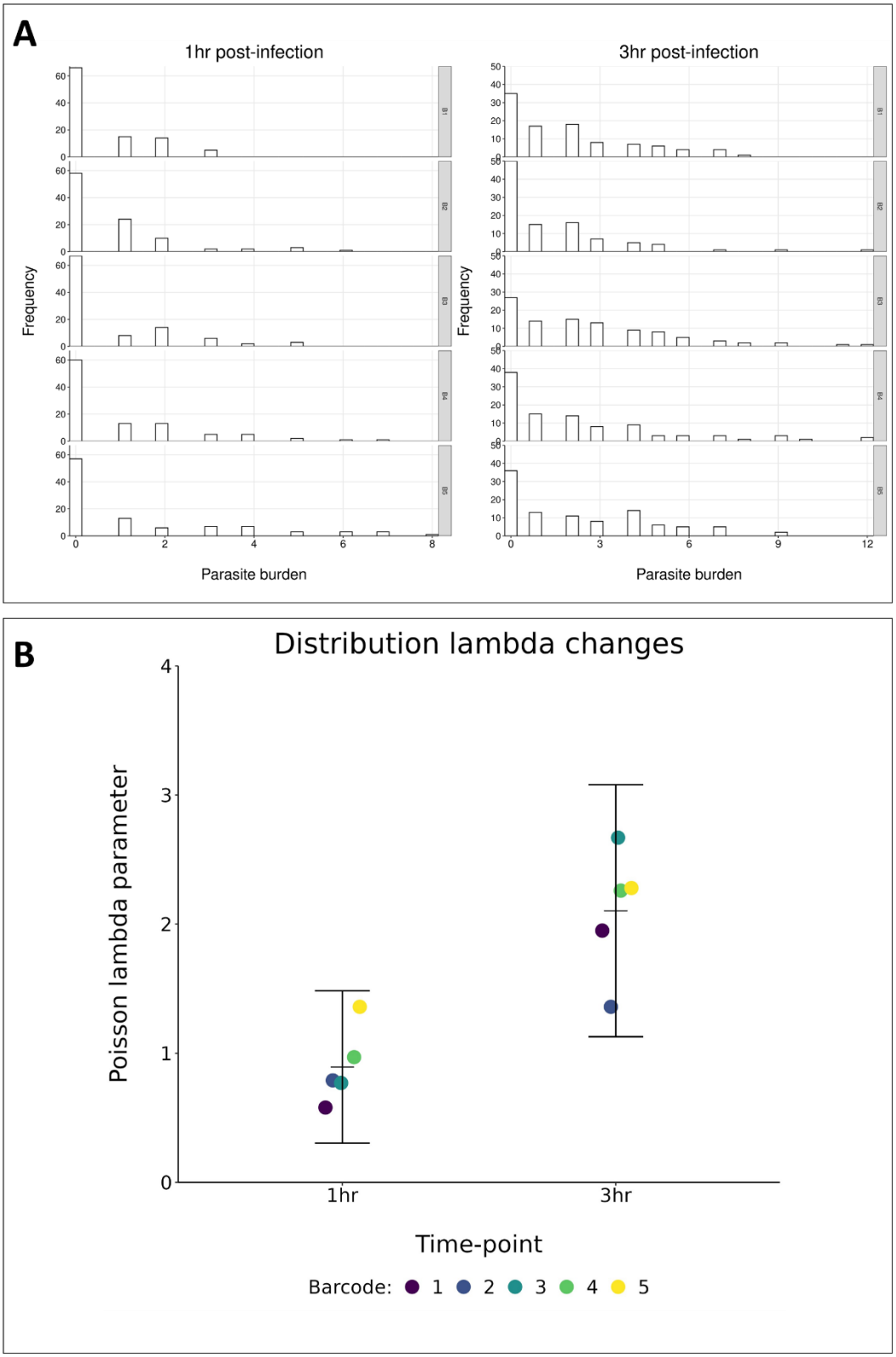


Figure 3.11: Infection burden in BMDMs for five individual lines from the barcoded LV9 T7 cas9 *L. donovani* library. A). Infection burden in B6.CD45.1 murine macrophages after 1 and 3 hours incubation with one of five barcoded parasite lines. Macrophages were plated at 1x10^4 per well of a chamber slide, and infected after 24 hr with 5X MOI promastigote single-barcode LV9 T7 cas9 *L. donovani*. Slides were incubated at 37 °C, then washed, fixed with methanol for 5 minutes and stained with 10% Giemsa for 3 minutes before counting

by eye. Analysis was performed in R Studio. B). Lambda parameter for the Poisson distribution fitted to the parasite burden at each time-point post-infection for each parasite line. Colour indicates parasite line (B1-5; see Appendix Table 1 for details). Error bars represent 2x standard deviation from the mean (black horizontal line).

A second assay measured macrophage response to infection via immunofluorescence staining for LAMP1 (Figure 3.10). LAMP1 is used as a proxy for lysosome maturation (Banerjee et al., 2016; Verma et al., 2017). BMDMs were infected with promastigote barcoded parasites; each line was infected individually as shown in Figure 3.10. Figure 3.12A shows representative images from 1 hour, 24 hours and 48 hours post-infection. Figure 3.12B shows the classification of parasites as present in a LAMP1+ vacuole (see Methods for further details on this strategy).

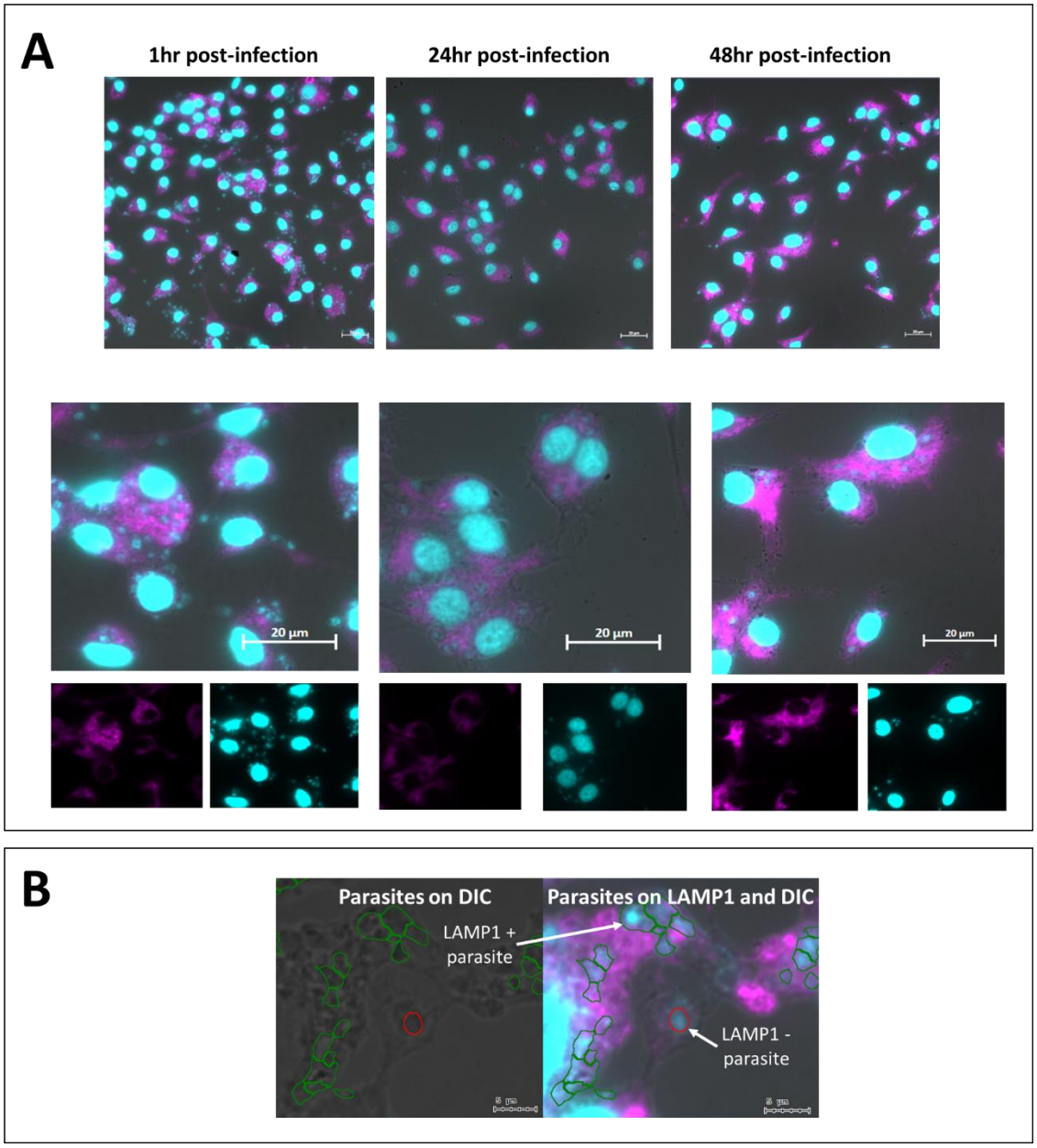


Figure 3.12: Representative images of LAMP1 staining in BMDMs infected with promastigote *L. donovani* parasites for 1, 24 and 48 hours. Macrophages were plated at 5x10^4 per well of a chamber slide, and infected after 24 hr with 5X MOI *L. donovani* with one of 5 barcode sequences (lines B1-B5) and incubated at 37 °C for 1 hr, 24 hr or 48 hr. Two replicates were performed per condition. Slides were washed, fixed with acetone for 10 minutes and stained with AF647 LAMP1 and Yoyo-1 before imaging by fluorescence and DIC microscopy. Ten images from each well were obtained, for a total of 20 images per condition. Images obtained were analysed using Strataquest and R Studio. A). Representative images from each time-point post-infection. Parasite and host cell nuclei are cyan and LAMP1 staining is magenta. B). Classification strategy of positive (green) and negative (red) parasites for co-localised LAMP1 staining using Strataquest.

Figure 3.13A shows the mean LAMP1+ parasites per macrophage for each barcode, with individual data points representing individual images (20 per condition). The data appears to show that between 1 and 24 hours, parasites per cell decrease, before increasing again by 48 hours, which could be due to initial killing of promastigotes, followed by replication after differentiation into the amastigote stage. The same strategy, with the same reasoning, for the analysis to compare the lines was taken as described above, but in this case a Gamma distribution was the most suitable fit to the data. All null values (which occurred only in the 24 hour time-point data) were removed prior to fitting. The shape parameter for each barcode at each time-point is shown in Figure 3.13B. Again, shape for all distributions falls within 2 standard deviations of the mean, indicating that differences between lines are random.

Further, when combining this with the data in Figure 3.11B, no specific line is shown to be consistently higher or lower than the others across all time-points, which reinforces that differences are likely to be stochastic and not the result of differences in fitness.

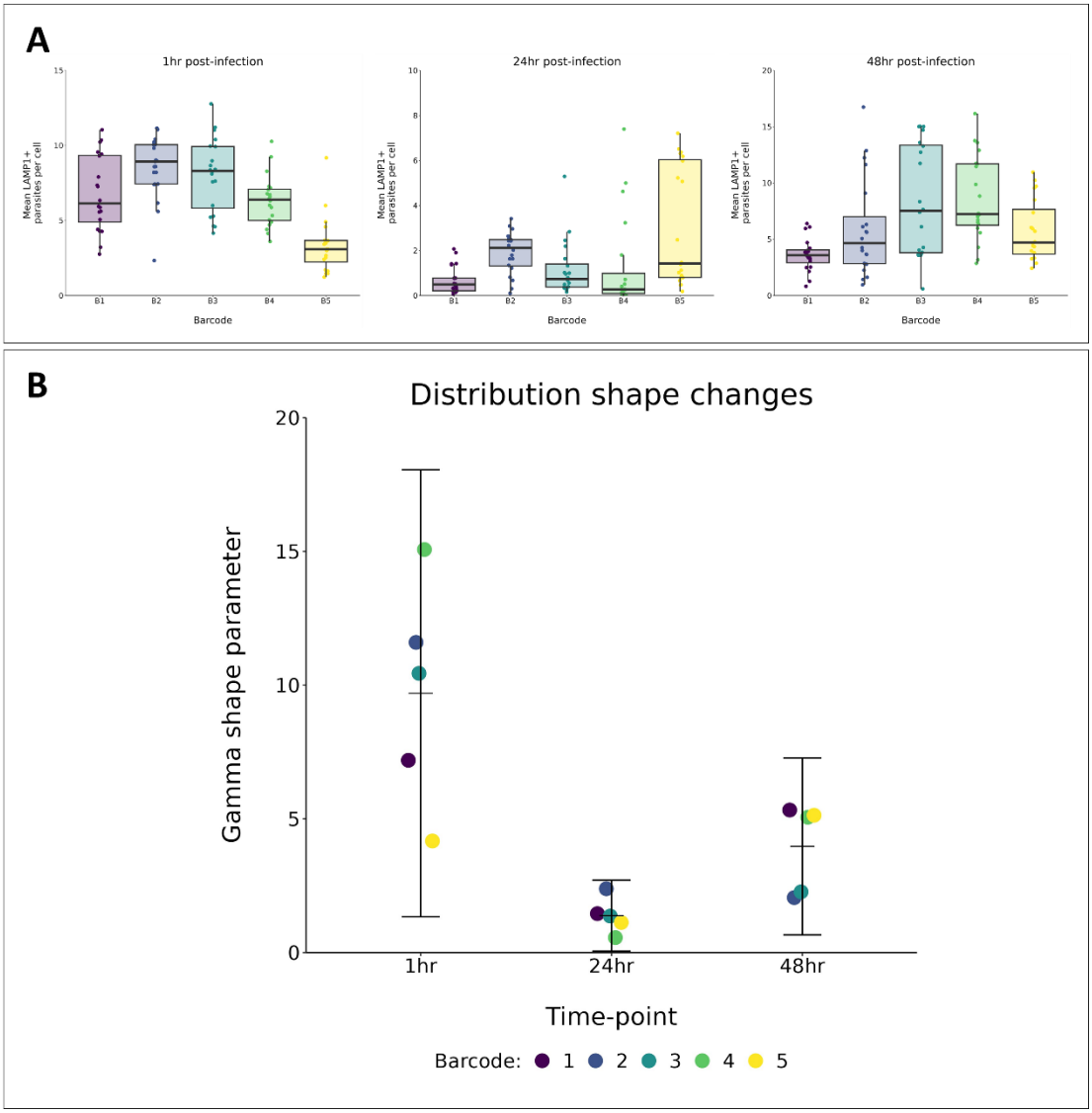


Figure 3.13: LAMP1 recruitment to parasite phagosome in response to infection with five individual barcoded *L. donovani* lines. A). Boxplots of total LAMP1+ parasites over total macrophage count, with each image (n=20 per barcode per time-point) represented as an individual dot. B). Shape parameter values for the Gamma distribution fitted to the mean LAMP1+ parasites per macrophage for each parasite line at each time-point. Colour indicates parasite line (B1-5). Error bars represent 2x standard deviation from the mean (black horizontal line). Colour indicates barcoded line (B1-5; see Appendix Table 1 for details) and is consistent across all panels. Boxplots show the median, 25th and 75th quartiles, with whiskers representing all data-point within 1.5x Interquartile Range of the median; points above or below the whiskers represent outlying data values.

Combined with the lambda values in Figure 3.11B, these data show that no barcoded line, of the five compared, is consistently exhibiting a greater ability to infect macrophages across the experiments and time-points assessed. This indicates

that the infectivity of barcoded lines is not affected by the barcode itself, although only 5 lines were tested here and it is possible other lines in the library may have fitness differences.

3.2.6 PILOT experiment *in vivo* to assess feasibility and optimise Barseq-PCR protocol

A PILOT experiment, using 10 barcoded parasite lines, to provide proof-of-concept, assess feasibility and optimise protocols for Barseq-STAMP was designed as shown in Figure 3.14. Five BALB/c mice were infected with a promastigote pool of B1-10 (see Appendix Table 1 for details) lines. Mice were killed at four weeks post-infection and spleen, liver and one lymph node were flash frozen for analysis of parasite barcode DNA. Bone marrow was also taken from femurs and tibias. Mouse 1 was killed two days prior to the other mice, to assess parasite burden and ascertain whether the infection had established, before culling other mice. Parasites were visible on the spleen and liver smear obtained, but an error in the staining protocol meant that the samples were not clear enough for counting by eye, so Mouse 1 is missing from the parasite burden analysis. The protocol for processing bone marrow samples was adapted in the period between Mouse 1 and the other mice being killed, to facilitate a higher DNA yield for Barseq amplification; this change to the protocol was successful in enabling detection of barcodes in bone marrow samples but Mouse 1 bone marrow could not be amplified by PCR and as such is missing from the analysis.

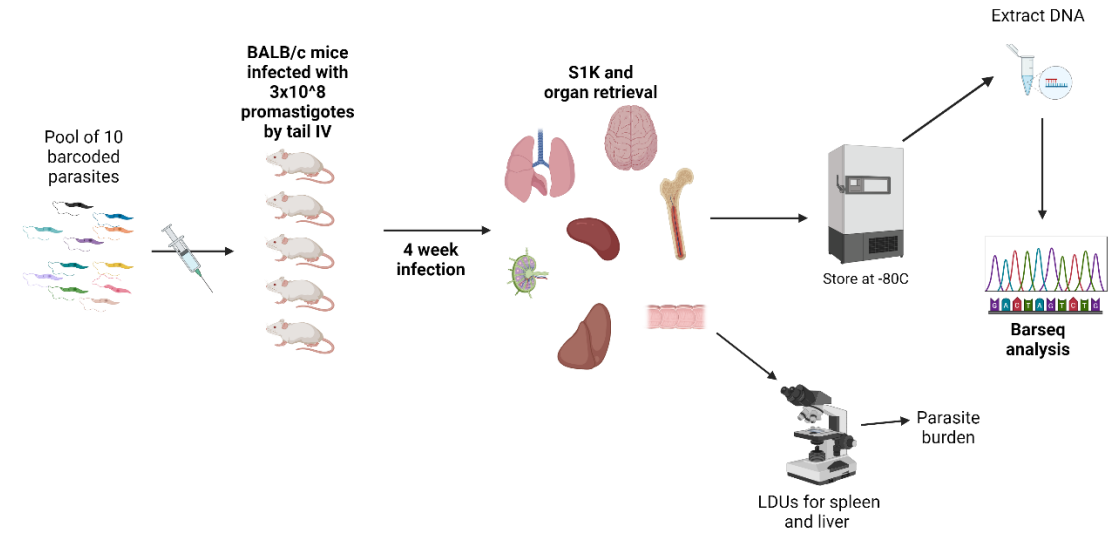


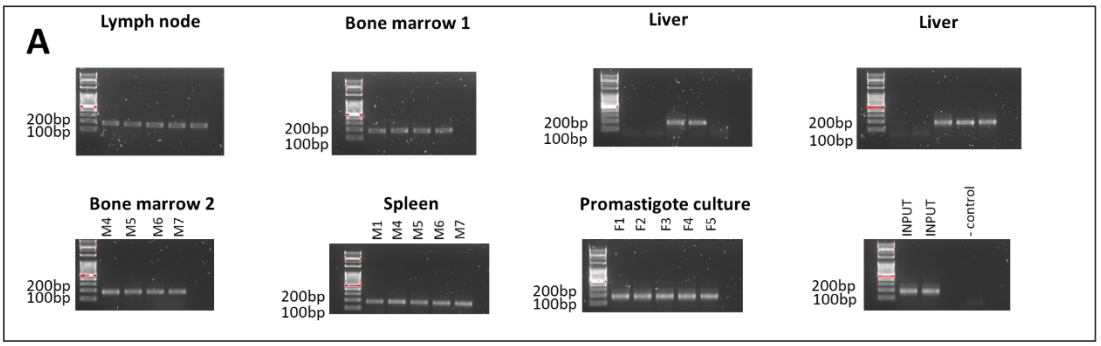
Figure 3.14: PILOT experiment using a 10 barcode pool in five 4 week infected BALB/c mice to validate Barseq-STAMP protocol. Schematic representation of the experimental protocol. Five mice were infected by tail IV with a pool of 10 barcoded promastigote LV9 T7 cas9 *L. donovani* parasites (3x10^8 parasites per mouse). Mice were culled at 4 weeks post-infection and various tissues were flash frozen and stored at -80 °C. Smears were also taken of liver and spleen. Cryo-sections of each tissue (except bone marrow) were also taken for further validation and analysis but were not subsequently used in this project.

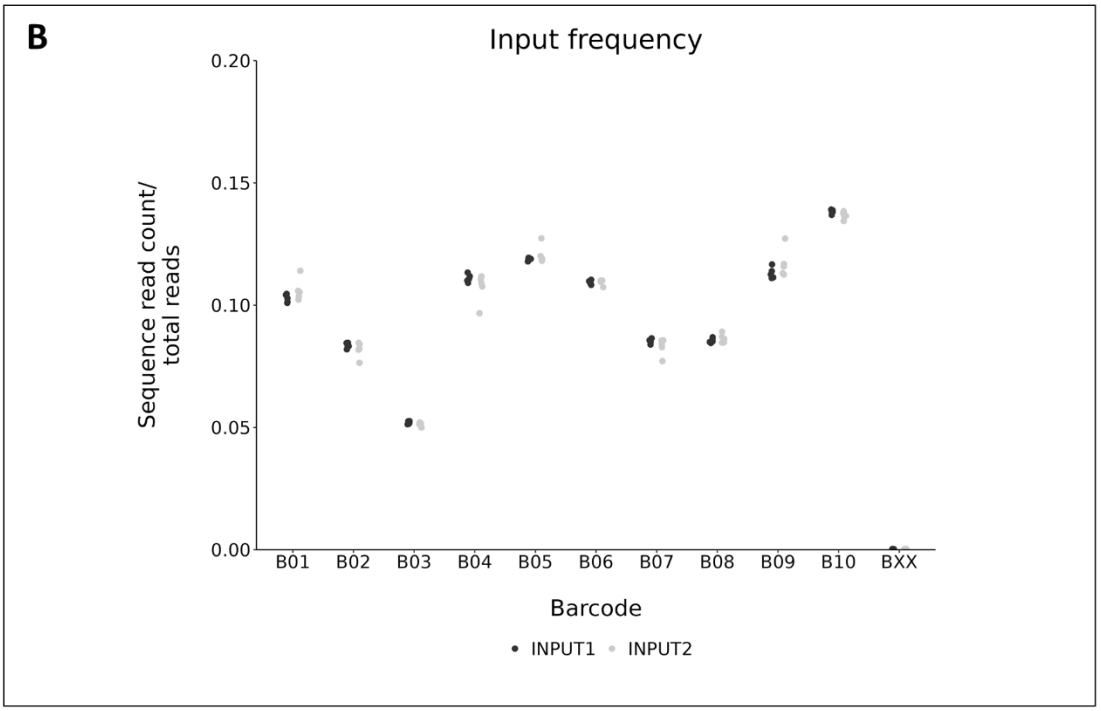
Figure 3.15A shows the PCR products for each sample sent for sequencing after Barseq amplification. Figure 3.15B shows the barcode reads (normalised to total reads present in the sample) for the input pool. This DNA was amplified by two separate PCR reactions and each sample from this was sequenced independently 5 times, to assess the replicability of the sequencing across samples; this showed that the PCR and sequencing is consistent in amplifying the barcodes.

The first data processing step extracts the 20 bp sequence following the amplification primer in all forward sequence reads, in order to detect any potential unexpected barcode sequences present. This led to the discovery of an unexpected barcode at a high abundance in most *in vivo* tissue samples: B59 (sequence: TGGATCTTAGGACGCAACAT). On average, this barcode comprised around 50% of the reads in tissue samples. However, it was only present below background levels in the input samples for this experiment. We initially suspected this was due to contamination during DNA extraction from tissues; however, further analysis of data from subsequent experiments using the whole library also detected this

barcode at higher than expected levels in other tissue samples, albeit at far lower total proportions than seen in the PILOT study, and not across all tissue samples. This indicates that the over-representation may be a result of increased virulence or propensity of this line to undergo clonal expansion *in vivo*. This barcode was subsequently removed from all analysis for both this PILOT experiment and for all other Barseq experiments included in this thesis.

Figure 3.15C shows the read frequency for each barcode (normalised to total reads) for each tissue, with colours indicating the individual mouse the tissue was taken from. It was observed in the lymph node samples that B1 seemed to be present at a higher abundance than the other barcodes in all of the samples obtained, but the same pattern was not observed in other tissues assessed.


A
Lymph node
Bone marrow 1
Liver
Liver
200bp
100bp
Bone marrow 2
M4
M5
M6
M7
Spleen
M1
M4
M5
M6
M7
Promastigote culture
F1
F2
F3
F4
F5
INPUT
INPUT
- control


B
Input frequency
Sequence read count/
total reads
0.20
0.15
0.10
0.05
0.00
B01
B02
B03
B04
B05
B06
B07
B08
B09
B10
BXX
Barcode
INPUT1
INPUT2

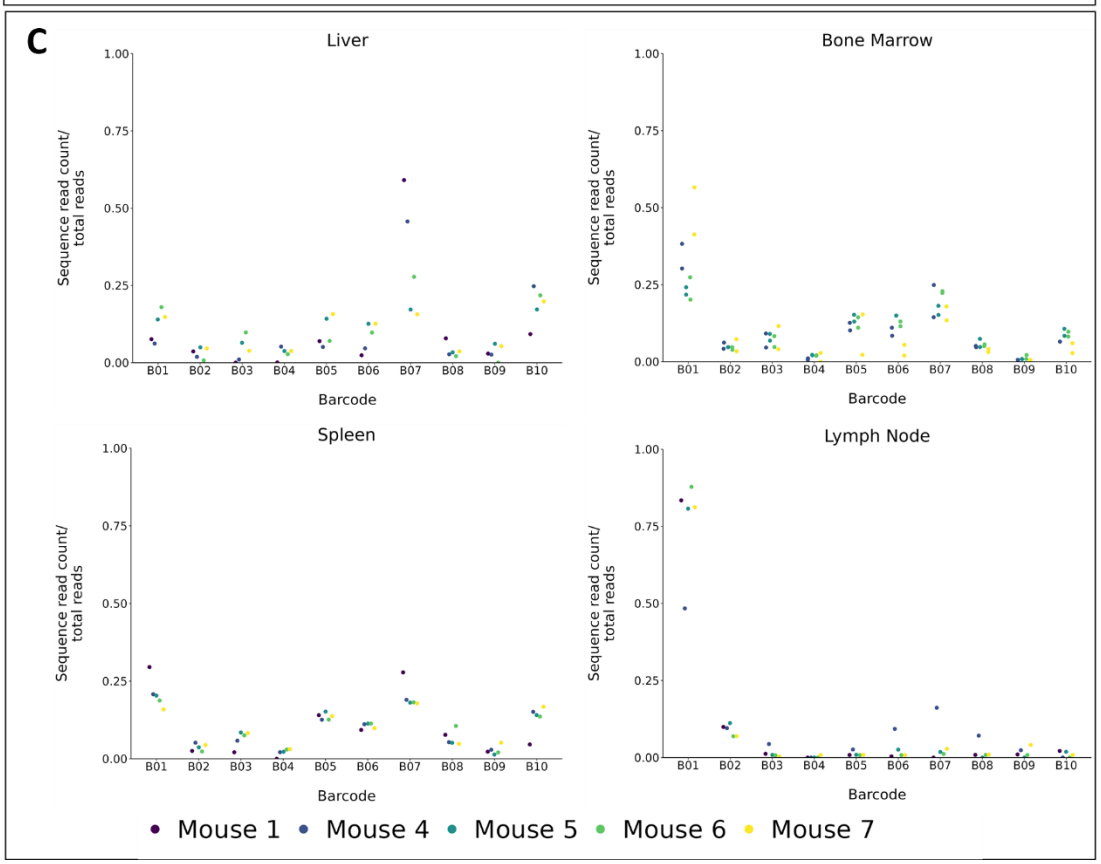

C
Liver
Bone Marrow
Spleen
Lymph Node
Sequence read count/
total reads
1.00
0.75
0.50
0.25
0.00
B01
B02
B03
B04
B05
B06
B07
B08
B09
B10
Barcode
Mouse 1
Mouse 4
Mouse 5
Mouse 6
Mouse 7

Figure 3.15: PCR and Illumina sequencing of barcodes in five 4 week infected BALB/c mice. A). PCR products run on a 2% agarose gel for each mouse tissue and input population. M1-7 indicates mouse ID. F1-5 refers to replicates of *in vitro* culture of the barcoded pool. B). Illumina sequencing results for two separate PILOT experiment input pool PCRs containing 10 barcoded LV9 T7 cas9 *L. donovani* lines. Each PCR was sequenced five times. BXX indicates proportion of barcode which was not B1-10 barcode in input. C). Illumina sequencing results (after other barcodes were excluded) for four mouse tissues (spleen, liver, lymph node and bone marrow) taken from BALB/c mice infected with this input pool for 4 weeks in a proof-of-concept PILOT experiment. Only four mice (M4, M5, M6 and M7) are shown for bone marrow, due to technical issues with M1 samples. Colours indicate individual mice.

These data show that the barcoded library can be used for *in vivo* work safely, that the Barseq-STAMP protocol works, and that the individual barcodes can be detected in multiple tissues in a BALB/c model.

3.2.7 Barseq-STAMP metrics in PILOT feasibility experiment

We then sought to ascertain that the barcode proportions for each tissue could be used to conduct population dynamics analysis on each tissue. Figure 3.16 shows the results of STAMP analysis to determine FP for each tissue in each mouse and GD for pairwise population comparisons in tissues within each mouse (Figure 3.10A; for details see Methods and Appendix Script 3). Generations (g) was set as 1 for FP calculations. Due to the low number of barcodes used, the estimates presented are not necessarily accurate and are simply an illustration of the technique *in vivo*.

The five mice showed consistently that the lymph nodes had the most restrictive bottleneck to parasite colonisation, with the lowest FP. Bone marrow also showed similar FP across all mice and within each mouse for the two separate sets of bones assessed. Spleens showed a relatively high FP across all mice, except Mouse 1, whilst livers were much more heterogeneous across the five mice. Mouse 5 and Mouse 7 had a much higher FP, and thus more permissive bottlenecks, in their livers than the other mice.

GD analysis showed that the spleen and liver were consistently similar, and both were highly similar to the input pool. The lymph nodes were the most unrelated to other tissues across all mice, with the highest GD across comparisons. The parasite burdens in spleen and liver (Figure 3.16B) did not appear to show a link to FP values in these tissues, with the lowest spleen burden (Mouse 7) actually having the highest FP in this tissue, but a small sample size means statistical analysis is not possible to confirm this relationship. Spleen and liver burdens showed the same pattern across mice overall, with the highest spleen LDU matching with the highest liver LDU and vice versa.

A visual representation of relative FP and population relatedness is shown in Figure 3.16C. Diversity is indicated by colour of the amastigotes in each circle. Similar colours in the same mouse indicate that populations are relatively similar to each other.

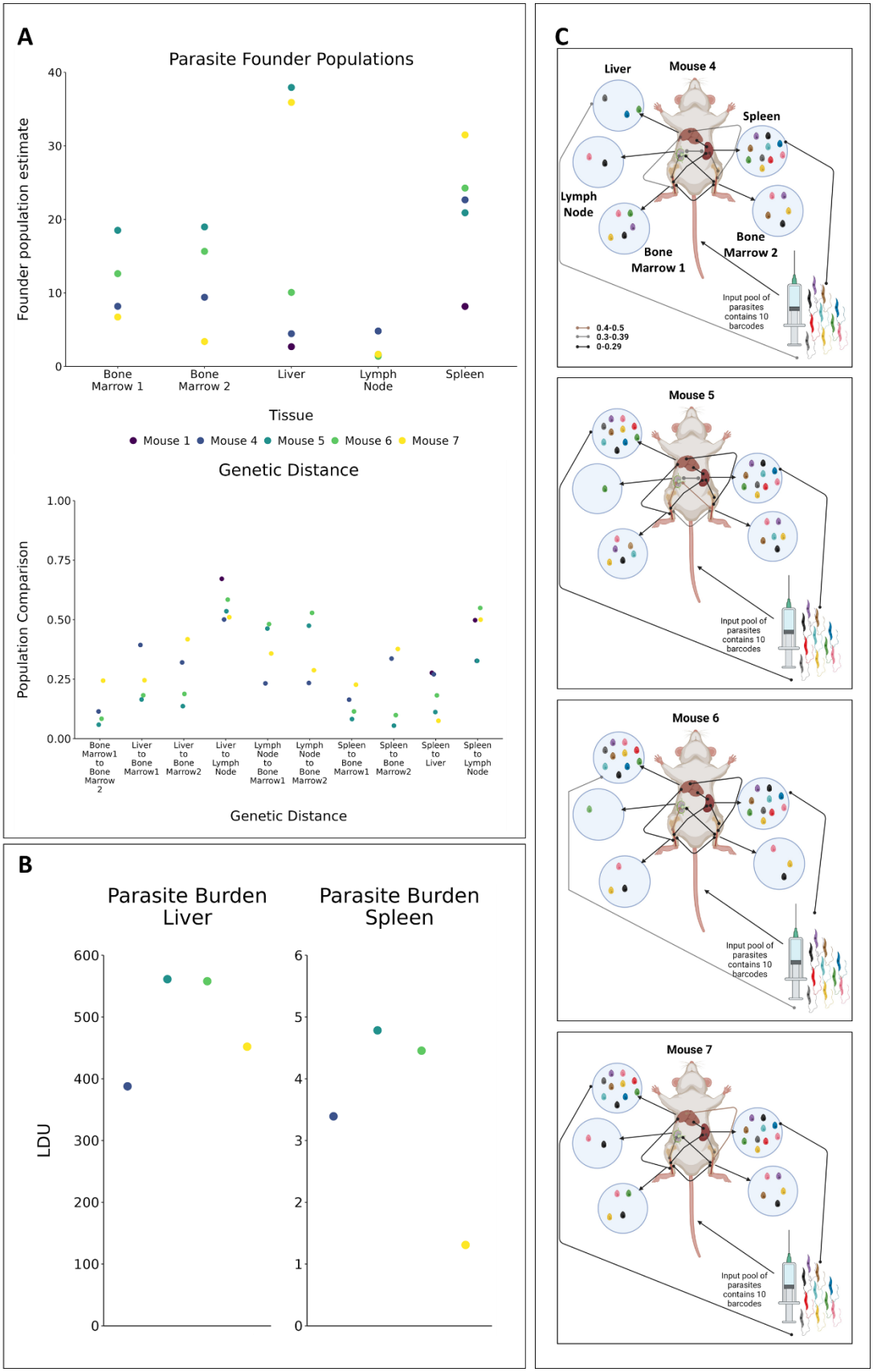


Figure 3.16: Barseq-STAMP metrics for Founder Population (bottleneck stringency) and Genetic Distance (population relatedness) for tissues obtained from five 4 week infected BALB/c mice. A). STAMP calculation of founder population (FP) and genetic distance (GD) from proof –of-concept PILOT experiment. Tissue samples from BALB/c mice infected for 4

weeks with a pool of ten barcoded LV9 T7 cas9 *L. donovani* promastigotes were processed via the Barseq-STAMP protocol to determine barcode frequencies compared to the input parasite pool. These frequencies were used to calculate FP values, indicating bottleneck restrictiveness, and GD values, indicating population relatedness between tissues. Larger FP values indicate a more permissive bottleneck, whilst smaller values indicate a more restrictive one. FPs were calculated using generations (g) as 1. Larger GD values indicate a more divergent pair of populations, whilst smaller values indicate a more similar pair of populations. Only four mice (M4, M5, M6 and M7) are shown for bone marrow, due to technical issues with M1 samples. B). Parasite burden (LDU) for liver and spleen. Four mice shown due to a technical error with staining of M1 sample. Organ weights for M4 were taken as the mean of the organ weights for the other three mice, due to technical error. Colours represent individual mice and are consistent across panels A and B. C). Diagrammatic representation of PILOT results for four mice. Circles represent the founder population for each tissue, with number of amastigotes representing relative population size and colours indicating relative diversity (not to scale). Colours do not indicate individual barcodes and are not comparable between different mice. Lines linking two populations indicate the GD value for that pair (only values below 0.5 are shown), as indicated by the colours in the Mouse 4 image. Panel C created in Biorender.

Together, this work describes the development of a neutral barcoded library of *L. donovani* parasites, and their successful use *in vivo* to assess parasite population dynamics in multiple tissues in a mouse model of VL.

3.3 Discussion:

The PILOT study described here provides proof-of-concept for the Barseq-STAMP technique, showing that we can amplify and sequence the barcoded parasites from a range of tissue types, and that these can be used to calculate both FP estimates and GD for the parasite populations in these tissues. Additionally, this study has enabled optimisation of the technique and troubleshooting of practical issues for future experiments.

Our simulations, combined with previously published work by other teams using similar numbers, provides evidence that around 100 barcodes is sufficient for calculating FP with reasonable accuracy (Chevée et al., 2024; Jiang et al., 2024;

Louie et al., 2019; Wincott et al., 2022; Zhang et al., 2017). However, this is dependent on the biology of the system in question; larger numbers of barcodes enable more detailed discrimination of more permissive bottlenecks (Abel et al., 2015b). Since our library had to be produced using multiple individual transfections, a limit was imposed on the feasible library size that could be generated here. To create a larger library it might be possible to utilise degenerate base primers, which would allow far greater diversity within a small number of transfections. However, the barcode sequences chosen were designed to fit stringent criteria, including GC content, lack of BLAST matches, hamming distance and low likelihood of hairpin structures. Truly random barcodes can exhibit problems with differences in PCR and sequencing efficiencies, due to repetitive sequences, hairpins and GC biases (Johnson et al., 2023). The locations of the degenerate bases in the barcode sequences would therefore need to be carefully selected in order to avoid issues with any of the above and enable our barcode extraction and merging criteria to be applied to the raw sequence reads. Degenerate bases may also reduce the efficiency of the Barseq-PCR protocol and this could require further optimisation, and may potentially limit the use of the same lines for other techniques such as Zombie (Askary et al., 2020) or community ecology analysis.

We found that double selection for lines containing two copies of the same barcode was not possible, as had been hoped initially, due to the subsequent finding that only one cas9-HYG repair cassette had integrated into the parental *L. donovani* T7 cas9 line. We attempted a pooled transfection approach, but this was unsuccessful. It is not clear why this did not work, but the technique likely required further optimisation and slightly altered conditions compared to the single transfection approach we subsequently used. Alternatively, there is the possibility that integration did occur but detection of this was hampered by tandem duplication events or integration into a different locus than the intended one.

To assess barcode neutrality, an *in vitro* growth assay was utilised. The data indicated that differences between lines in terms of growth rates were stochastic, whilst our infectivity studies likewise provided evidence of neutrality.

We also tested the infectivity of the individual lines using an *in vitro* infection model with BMDMs. Only five lines were assessed for infectivity, so it is possible other

lines would have shown a different result to that observed, but for *in vitro* growth in culture, all lines were included. Analysis involved fitting distributions and comparing the parameters for each line to the average parameters of this fit. However, another approach would have been to compare the individual infection burdens to those of the pooled population. Comparing to the parental line was not used as a strategy, on the basis that we were interested in variance between barcoded lines, rather than a difference in fitness compared to the parental line which would not be used in any of the *in vivo* experimental work.

We also used LAMP1 as a marker of lysosome maturation, to check how macrophages responded to infection by promastigotes. Previous studies have noted that promastigote infection with *L. donovani* can delay recruitment of this marker and interfere with endosome fusion (Scianimanico et al., 1999; Verma et al., 2017). Other assays could have studied the macrophage-parasite interaction in more depth by using other endosomal markers, such as Rab-GTPases. We could also have measured how the parasites impacted key macrophage responses, such as superoxide production, inflammasome activation or IFNγ responsiveness, which would have provided insight into whether the barcodes might alter how parasites interact with the immune response *in vivo*.

Optimisation of the Barseq-PCR for amplification of barcodes proved more difficult than expected, most likely due to the limited range of options for primer designs, as we had to generate a short PCR product around the barcode for Illumina sequencing. With hindsight, this primer design could have been optimised when designing the barcode insertion primers, prior to insertion of the barcodes. The optimal reverse primer could have been included in the cassette design, which could have reduced this optimisation process. However, due to the high abundance of mouse DNA present in comparison to parasite DNA in each sample, it is likely this still would have taken time to fully optimise for infected tissues.

Our PILOT study allowed the full protocol from infection through to analysis to be tested in a group of 5 female BALB/c mice. The Barseq-PCR was able to amplify barcodes from spleen, liver, lymph node and bone marrow tissue from these mice; these are key tissues known to be colonised by *Leishmania* parasites *in vivo* and in patients. Extracting parasite DNA from bone marrow proved slightly problematic,

but was optimised between the cull of Mouse 1 and the remaining mice. In future studies, this set of tissues can be widened to include other tissues with evidence of infection in VL, such as gut, lung, brain and skin (Abreu-Silva et al., 2003; Bispo et al., 2020; Doehl et al., 2017; Duarte et al., 1989; Laskay et al., 1995; Prasad et al., 1996; Melo et al., 2017; Vinuelas et al., 2001; Zijlstra, 2016). In particular, assessing founder populations in skin would provide valuable data on the potential mechanisms underlying PKDL development and transmission of parasites to sand flies.

The method for extracting the barcoded parasite DNA differed compared to that used for previous STAMP experiments in bacterial systems and in the *T. gondii* study (Bachta et al., 2020; Campbell et al., 2023; Hullahalli and Waldor, 2021; Louie et al., 2019; Wincott et al., 2022; Zhang et al., 2017); these studies extracted pathogens from tissues and grew them in culture before extracting DNA from the populations, whereas the method described above for *Leishmania* required extraction of parasite DNA directly from infected tissue samples. This meant that efficiency of barcode amplification was reduced for the *Leishmania* system, but also reduced risks of inadvertent genetic drift or selection effects from the *in vitro* culture stage of the experiments. Since *Leishmania* in culture grows as a promastigote, whilst in tissues *in vivo*, the parasites exist in the amastigote stage, it was felt that too many additional opportunities for selection effects were present to use the established method for *Leishmania*. Studies have previously indicated that aneuploidy is more common in promastigotes, compared to the amastigote parasite form and that this can occur rapidly during the transition between life stages (Dumetz et al., 2017; Negreira et al., 2022). This introduces the capacity for parasites to alter their genomes in divergent ways, introducing fitness differences, whilst direct extraction from tissues avoids this problem. Extraction of live *Leishmania* from infected tissue other than spleen is also inefficient, so it is unlikely this would have yielded a full cohort of the present barcodes, confounding the analysis of the true tissue bottlenecks.

The PILOT study enabled the development of the analytic pipeline for STAMP analysis. The FP for each tissue was quantified, as was the GD between tissues in each mouse. Because we used only 10 barcodes here, these values are unlikely to

accurately reflect the true tissue population dynamics, but future studies using the whole library can now follow on from this work. Additional population dynamics and community ecology analytic methods can also be performed using the same barcode abundance data to investigate the *in vivo* parasite populations in more depth.

Whilst our *in vitro* work indicated that there were no obvious differences between lines, *in vivo* parasite behaviour and growth is not necessarily well-reflected by *in vitro* experimental work. In particular, the immune response to *Leishmania* is highly complex and differs between tissues. As such, in our PILOT study, it was noted that the B1 line appeared to have a slight advantage in lymph node specifically. Possibly this was due to the restrictiveness of the lymph node bottleneck, meaning a small advantage in growth rate or virulence resulted in a larger impact in this tissue. This could also be an example of a priority effect, where a slight head-start in colonisation causes a greater subsequent impact on the population structure downstream. It could also result from differences in the proportion of metacyclics present between lines before infection. However, the other 9 barcodes were present in various tissues, indicating that B1 did not override the others via a particularly strong selective effect across a range of tissues. This impact could be diluted with the full library of barcodes in the input pool. We also found that one barcode from our library, B59, was behaving unexpectedly – sequencing of this line could be useful to determine whether it has any specific mutations which are increasing its virulence in mice.

Finally, the production of a neutral barcoded library with the addition of the T3 promotor sequence means that other applications may be possible using the same lines. For example, a subset could be used to analyse the spatial distribution of individual barcoded lines, as described in Askary et al., 2020, providing a view of tissue dissemination and parasite interactions at a cellular level. This would be of particular value in the skin, where it has been shown that parasites spread in a 'patchy' way in order to maximise uptake by sand flies (Doehl et al., 2017; Doehl et al., 2021). Imaging-based approaches have previously been used to study cancer metastasis and population dynamics (Quinn et al., 2021; Rovira-Clavé et al., 2022), and these types of analyses could be valuable in understanding how parasites are

disseminating between tissues, particularly with reference to skin pathology in PKDL. These approaches could also aid in understanding cutaneous *Leishmania* infections if applied to other parasite species. Other potential uses may include tracking particular lineages over time or monitoring individual populations of parasites after multiple or sequential infections with different lines, to understand how co-infection and re-infection affect population dynamics.

In summary, this work has shown that the generation of an isogenic barcoded line of LV9 *L. donovani* parasites is possible using a CRISPR genome editing approach, and that this library can be utilised for the Barseq-STAMP approach for assessing bottlenecks and dissemination dynamics *in vivo* in a mouse model of VL. Following the successful PILOT study, this approach can now be applied to examine tissue parasite populations in detail, to understand the temporal and spatial dynamics of colonisation. This tool can potentially be used in additional ways to further understand infection dynamics and to test hypotheses around host-parasite interactions.

Chapter 4: Colonisation and dissemination of *L. donovani* in an amastigote and promastigote infected C57BL/6J mouse model of VL.

4.1 Introduction:

Visceral leishmaniasis (VL) is a systemic disease of multiple organs, with parasites known to heavily infect the spleen and liver, causing characteristic pathology of hepatosplenomegaly in both patients and animal model systems (Lipoldová and Demant, 2006; Stanley and Engwerda, 2007). Parasites have also been shown to infect lymph nodes and bone marrow, and can present in the skin after re-activation post-cure as PKDL (Mukhopadhyay et al., 2014; Stanley and Engwerda, 2007). The extent to which *Leishmania* parasites can infiltrate other tissues is currently unclear, but there is evidence for parasite colonisation in the brain (Abreu-Silva et al., 2003; Prasad and Sen, 1996; Melo et al., 2017), lungs (Bispo et al., 2020; Duarte et al., 1989; Laskay et al., 1995) and gut (Chattopadhyay et al, 2020; Dos Santos et al., 2021; Laguna et al., 1994; Lewis et al., 2020) in human patients and animal models (Figure 4.1). Parasite colonisation and dissemination to and between tissues is not well-understood and is thought to encompass both host and parasite factors (Lipoldová and Demant, 2006), alongside possible roles for abiotic contributions such as UV exposure (Montes de Oca et al., 2023; Mukhopadhyay et al., 2014) and specific co-infections or treatments which result in immune suppression (Mukhopadhyay et al., 2014; Stanley and Engwerda, 2007; Zijlstra, 2014).

Leishmania parasites have a complex life cycle, which involves maturation from procyclic promastigote to the infectious metacyclic promastigote form as they migrate through the sand fly vector (Séguin and Descoteaux, 2016). Parasites are initially inoculated into the bite site and are then taken up via phagocytosis by various host cells recruited to or resident at this site (Séguin and Descoteaux, 2016). Within the host cell, they can disrupt the normal process of lysosomal degradation to enable their survival, including by differentiation into the non-flagellated amastigote life stage after 24-48 hours (Séguin and Descoteaux, 2016).

Different *in vivo* experimental model systems for VL exist, with the most common being the mouse model, including the commonly used C57BL/6J mouse (Figure 4.1). This inbred model system is well-established within immunological research generally, providing advantages in research in terms of manipulation of mouse genetics and immunological function and for ease of replicability (Loeuillet et al., 2016). Other animal models, such as primates, dogs and the Syrian hamster are sometimes used for their greater reflection of human disease pathology, but have drawbacks such as wider genetic diversity, necessitating increased sample sizes in experiments (Lewis et al., 2020; Loría-Cervera and Andrade-Narvaez, 2014; Pearson et al., 1992). Most research does not use sand fly biting for inoculation of parasites, due to difficulties with optimising parasite uptake, transmission, rearing and maintenance of sand flies and issues of biosafety (Michalsky et al., 2007; Volf and Volfova, 2011), so infections are usually done using intravenous or intraperitoneal injection (Ong et al., 2020) (Figure 4.1).

Additionally, different research teams often use different parasite stages within their experiments, but it is unclear to what extent these different approaches mirror natural infection dynamics. Spotin and Parvizi (2016) showed a possible difference between visceralisation of *L. major* dependent on parasite stage at infection in a small number of mice, whilst Ong et al (2020) showed differences in *L. infantum* parasite load between these two approaches, with promastigote infection yielding a reduced parasite burden when assessed via *in vivo* imaging; however, specific differences in disease pathology and tissue colonisation have not been assessed.

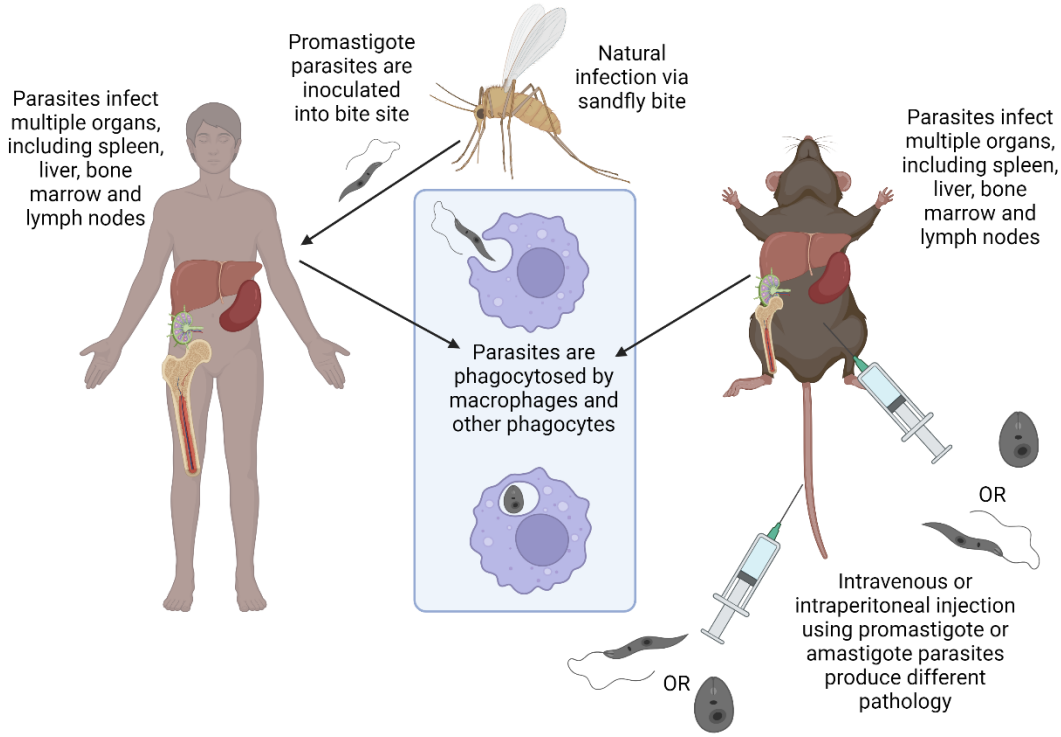


Figure 4.1: Natural infection versus artificial infection murine model for VL. In a natural infection, the sand fly vector bites a host and transmits promastigote-form *Leishmania* parasites into the skin. The parasites are phagocytosed by macrophages, monocytes and other phagocytic cells. Parasites are known to disseminate to liver, spleen, bone marrow and lymph nodes, along with some evidence for their presence in other tissues including lung, gut and brain. In the mouse model, the infection is usually done via intravenous injection into the tail vein, although intraperitoneal injection is also sometimes used. Amastigote stage parasites are most common, but promastigote parasites can also be injected. These methods produce some differences in pathologies and parasite loads in the few comparative studies undertaken, although parasites still infect the main organs of human disease using both infection methods. Figure created in Biorender.

The field of microbial ecology has expanded rapidly thanks to developments in high-throughput DNA sequencing, allowing detection of microbial species in both host tissues and the wider environment (Nemergut et al., 2013). These improvements have enabled the assessment of bottlenecks in microbial dispersal and pathogen spread, including via the insertion of artificial diversity in the form of genetic barcodes, for analysis of bottlenecks within *in vivo* models (Abel et al., 2015a; Abel et al., 2015b; Bachta et al., 2020; Campbell et al., 2023; Hotinger et al., 2024; Hullahalli and Waldor, 2021; Louie et al., 2019; Wincott et al., 2022; Zhang et al.,

2017). Developing an understanding of spatial and temporal bottlenecks by assessing different tissues and different time-points in an experimental model (Figure 4.2) provides indications of when and where stressors are applied to the pathogen, enabling the conception of hypotheses about important immunological events in disease progression. Additionally, these analyses can provide a more fundamental understanding of the mechanisms by which infiltration and survival of pathogens occur. For example, the barriers which result in population reduction can include the innate and adaptive immune response, the availability of suitable host cells, and tissue-specific conditions such as local pH or nutrient availability, as well as inter and intra-species competition.

Other community ecology methods to understand diversity and assembly processes in macro-level systems have been adopted and adapted by the microbiology community primarily to understand environmental microbial ecology (Nemergut et al., 2013). Diversity of a community can be parsed into three main types: alpha, beta and gamma (Figure 4.2). Alpha diversity is concerned with the number of "species" (or other relevant categories for the community in question) and their relative evenness within a defined population, such as a spatial region ("patch") or an experimental culture (Nemergut et al., 2013); it is often measured using metrics such as species richness and evenness, or metrics which encompass both concepts, like the Shannon Index (SI), which quantifies the uncertainty in identity of an individual in the population in question (Alberdi and Gilbert, 2019; Shannon, 1948). Beta diversity assesses the similarities between these patches in a pairwise manner, measuring whether different populations are present and their relative abundances in each locale to generate metrics of dissimilarity or distance, such as the Bray-Curtis and Raup-Crick metrics (Nemergut et al., 2013; Raup and Crick, 1979; Stegen et al., 2013; Stegen et al., 2015). Gamma diversity measures the diversity of the entire community of patches comprising the ecosystem in question and can be used to compare different, non-interacting, systems (Nemergut et al., 2013).

The processes which give rise to these different communities have, until relatively recently, been an exercise in theory for microbial ecosystems, but recent developments in the use of null models have allowed the concept of beta diversity to be adapted to determine how communities form (Figure 4.2). Whilst Sloan et al

(2006) proposed a neutral model of assembly, which was shown to fit some prokaryote populations well, Vellend (2010) and Hanson et al., (2012) described the four mechanisms shaping microbial communities as selection, dispersal, stochastic drift and mutation/diversification, providing a framework for considering the roles of non-stochastic processes in community composition (Nemergut et al., 2013). Comparing the compositional differences in terms of taxa abundances to that expected to occur due to chance can provide indications of deterministic effects such as selection, whilst comparisons of species incidence relative to overall abundance can parse out different stochastic influences shaping the system, namely drift versus dispersal mechanics (Figure 4.2) (Chase et al., 2011; Ning et al., 2020). Chase et al., 2011 developed the beta Raup Crick Index (βRC) to compare species incidence between two groups, using the earlier-established Raup-Crick index (Raup and Crick, 1979); this approach was further modified to create the Bray-Curtis modified Raup Crick Index (RC_{BRAY}) by Stegen et al. (2013 and 2015) to account for differences in abundance and different taxa relationships, comparing the differences observed to null model simulations, with further developments in the specific parameters for taxa relationships provided by Ning et al (2020) in their iCAMP protocol, allowing the estimation of the contributions of different processes in the assembly in microbial communities. The null model approach removes the influence of differences in alpha and gamma diversity, which can complicate interpretation of differences in beta diversity (Chase et al., 2011; Nemergut et al., 2013; Ning et al., 2020; Stegen et al., 2013; Stegen et al., 2015; Zhou and Ning, 2017).

Until recently, the application of these tools to understand within-host pathogen communities has been limited to studies of the microbiome (Burns et al., 2016; Debray et al., 2022) and pathogens with high levels of natural diversity, such as viruses (Anthony et al., 2015). However, the expansion of CRISPR-genome editing which enabled the development of STAMP can also be used to understand population diversity and assembly processes of artificial pathogen systems within hosts. A recent paper has combined both methods to determine bottlenecks, diversity and how populations are being built under different conditions in a plant-pathogen system (Jiang et al, 2024).

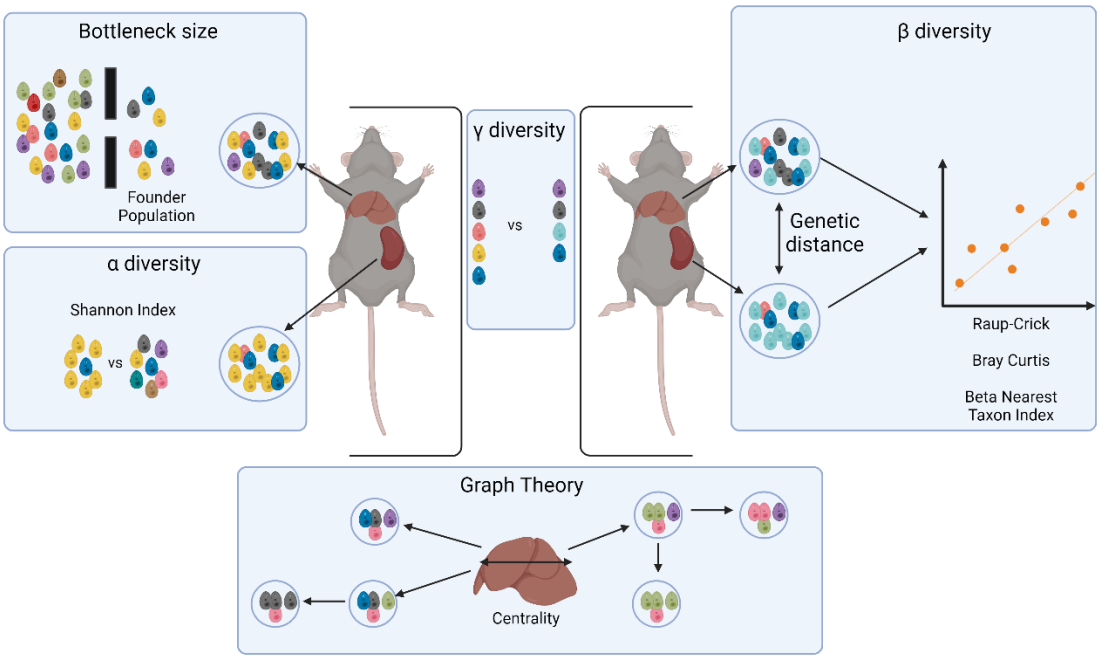


Figure 4.2: Illustration of the application of ecological analysis concepts to assess within-host pathogen dynamics in the context of a murine VL model. The murine model of VL involves systemic infection of multiple tissues, which allows the mouse to be considered as an ecological community composed of multiple parasite populations. The illustration shows this in the context of two tissues known to harbour parasites in murine and human infection - the liver and spleen - but the concept can be applied to any and indeed all tissues in the mouse model to fully understand the infection in terms of ecosystem dynamics. Bottleneck size is sensible to consider first, given that it applies an initial barrier to colonisation of tissues by parasites. Bottleneck size can be measured by proxy using founder population (FP); as bottlenecks become more restrictive and allow fewer parasites to pass, the FP becomes smaller and less diverse. Next, there is population diversity, which can act on a number of levels. Alpha diversity is diversity within a single population (in this case a tissue), and can be measured using various metrics which quantify “species” abundance, incidence and evenness. One popular metric which captures all three concepts is the Shannon Index (SI). A higher SI indicates that a tissue has a population consisting of a larger number of individual “species” and that they are relatively even in abundance, whilst a smaller value indicates few “species” or a highly skewed population with a few “species” dominating. Beta diversity is the relative difference in diversity of two populations - in this example the liver compared to the spleen. It can also be measured using various metrics, including: Genetic Distance (GD), which accounts for genetic drift; Raup-Crick, which assesses the probability of “species” co-occurring in sub-samples from both populations; Bray-Curtis, which measures the dissimilarity of the abundances of species in both populations, weighting for overall abundance such that rarer “species” are more important

if shared; and beta Nearest Taxon Index (bNTI), which measures the similarity of shared taxonomic (or other chosen clustering, such as by specific traits) groups in both samples and compares this to the level which would be expected to occur due to random chance. Gamma diversity quantifies the entire community diversity and can be used to compare two communities, such as a control mouse and a mouse under a treatment, such as antimicrobial drugging or gene knock-out. Graph theory is also a concept often utilised for spatial ecology and biogeography analysis of landscapes and epidemics, but can be applied to within-host parasite populations when assessing possible links and trajectories between different tissue populations. Tissues can be considered as vertices and paths between them as edges, measured as distance metrics or other conceptual approaches. In the illustration above, the liver can be conceptualised as a connecting vertex between parasite populations in other tissues within a mouse, acting as a bridge between the populations of parasites to the left and right; directed graphs can provide indications of directed movement whilst undirected graphs give insight into connections between vertexes which may flow in either or both direction. This can enable prediction of minimum spanning trees (MSTs) and insight into dissemination patterns, as well as pointing to key tissues which may be fundamental to network stability, and thus key targets for clearance efforts. Figure created in Biorender.

Artificial systems have some advantages over the use of natural diversity in terms of community ecology approaches. Firstly, genetic barcoding allows for neutral individuals to be distinguished from each other, which enables parsing of trait-based selection from other deterministic processes, such as priority effects and historical contingency, which may be deterministic but not grounded in a trait-specific advantage (Zhou and Ning, 2017). Conversely, the initial neutrality means that potentially advantageous adaptations arising can be spotted and their influence quantified (Jasinska et al., 2020). Additionally, the number of unique individuals can be manipulated to be similar to that of traditional macro-level communities; most diversity quantification methods were designed for use in settings with around 100 species or fewer and as such there are problems when adapting these approaches for communities consisting of far more individuals, which includes most microbial samples derived via sequencing (Gotelli and Ulrich, 2012; Zhou and Ning, 2017). Artificial barcode approaches can circumvent this problem and allow more powerful analyses of community dynamics.

Whilst network theory has been widely applied to understand various biological phenomena, including taxa interactions (Faust and Raes, 2012; Xiao et al., 2017), genetic signalling pathways (Badia-i-Mompel et al., 2023) and biodiversity analysis (Gonzalez et al., 2011), including for environmental distribution of microbes (Nemergut et al., 2013), it has not yet been applied to understand within-host spatial relatedness of microbial population at different sites. Considering the host as a meta-community of interconnected patches can allow the application of macro-biological landscape ecology and biogeography theory, established in conservation and habitat loss studies, to microbial systems (Figure 4.2) (Hanson et al., 2012; Nemergut et al., 2013). This can help to decipher the connectedness of different tissue pathogen populations, enabling analysis of the key tissues and routes grounding overall network stability.

The aim of this study was to utilise community ecology and population dynamics approaches such as those described above to understand how parasites colonise and disseminate between different tissues after infection with promastigote or amastigote *L. donovani* parasites, using the C57BL/6J murine model of VL. The previously generated barcoded library allows these approaches to be used to determine bottlenecks, relatedness and assembly dynamics of populations within key tissues of interest over a 28 day infection. The overarching hypothesis was that spatial and temporal changes in tissue bottlenecks would differ between the two models and that studying the diversity, dissemination and assembly of the resulting populations over time would provide insights into parasite behaviour in different microenvironments, generating evidence for further testable hypotheses about parasite-host interactions *in vivo*.

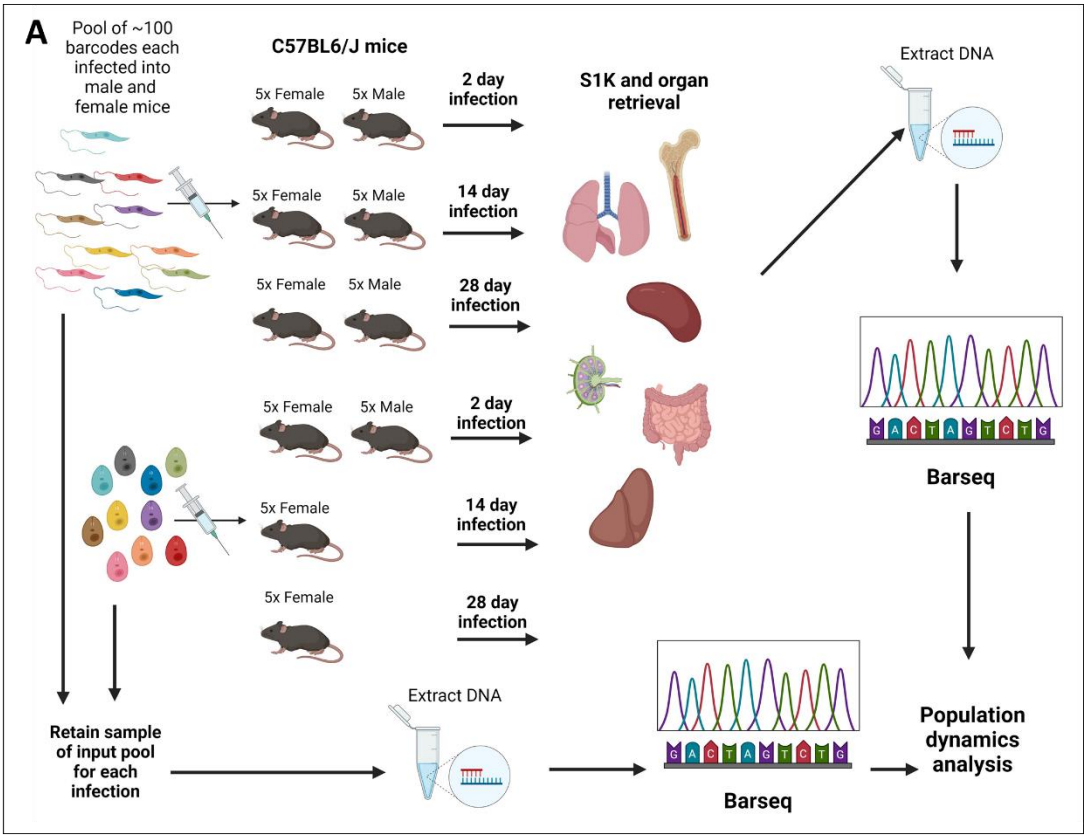
4.2 Results:

4.2.1 Establishing a promastigote and amastigote *L. donovani* infection model in C57BL/6J mice

In order to compare the population dynamics of parasites within host tissues, we utilised two infection models. Figure 4.3A shows the experimental protocol undertaken. The promastigote infection (PRO) used 1x10^8 stationary phase

promastigote LV9 T7 cas9 *L. donovani* parasites, consisting of the full barcoded library, infected via a tail intravenous injection, whilst the amastigote infection (AMA) used the same infection protocol, but with 5x10^7 amastigote parasites immediately after harvest from the spleens of B6.*Rag2*$^{-/-}$ CD45.1 mice. Mice were killed at either 2 days, 14 days or 28 days post-infection (p.i) and tissues were flash frozen for subsequent DNA extraction, Illumina sequencing via the Barseq protocol and downstream analysis. Five male and five female C57BL/6J mice were infected at each time-point in the PRO group and also for the 2 days p.i AMA group, whilst the 14 day and 28 day p.i AMA groups consisted of five female mice each. It should be noted that the 2 day infection time-point for the promastigote infection was a separate infection to the two later time-points, and only liver, spleen, two sets of femurs and tibias, and a single lymph node were taken from these mice. For other time-points, the same tissues, plus an additional lymph node, gut and lung were also taken and stored in the same manner. At 28 days p.i, three lobes within each liver were frozen individually to allow intra-organ comparisons. In the PRO infection model, it was found that some mice exhibited what appeared to be an anaphylactic reaction shortly after injection with promastigote parasites; this was not the case when a lower dose of promastigote parasites was given, and as such, follow-up promastigote infection experiments were performed with 5x10^7 parasites per mouse.

Parasite burdens (Figure 4.3B and D) and hepatosplenomegaly (Figure 4.3C and E) for 14 and 28 days p.i mice from both PRO and AMA groups exhibited a significant increase in burden (MW_{AMA}: p=0.0079 and MW_{PRO}: p=0.00013) and organ weight in the spleen (MW_{AMA}: p=0.0159 and MW_{PRO}: p=0.00018), but only in organ weight in the liver (MW_{AMA}: p=0.0357 and MW_{PRO}: p=0.0003) for both infection groups.


A
Pool of ~100 barcodes each infected into male and female mice
C57BL6/J mice
5x Female
5x Male
2 day infection
14 day infection
28 day infection
S1K and organ retrieval
Extract DNA
Barseq
Retain sample of input pool for each infection
Population dynamics analysis

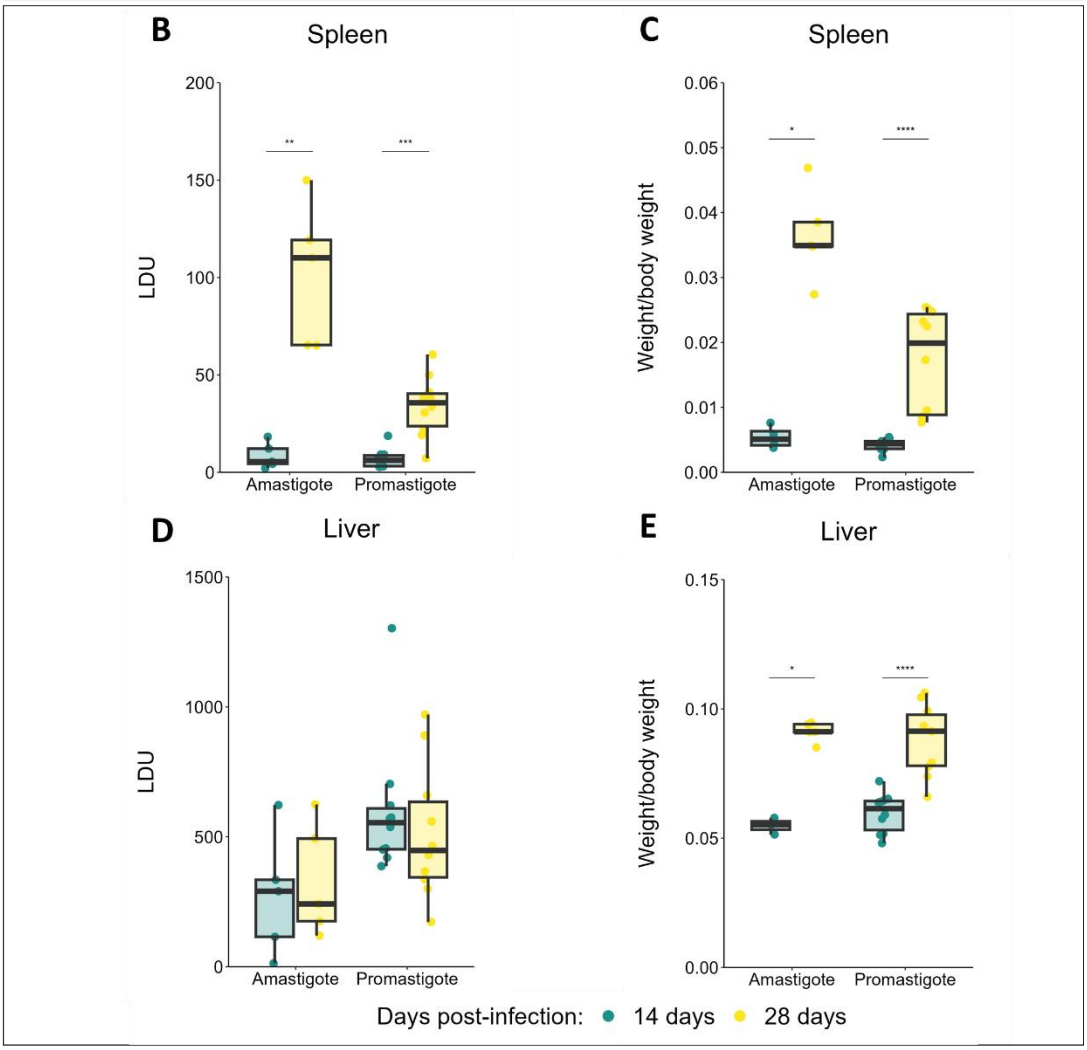

B
Spleen
LDU
C
Weight/body weight
D
Liver
E
Amastigote
Promastigote
Days post-infection: 14 days 28 days

Figure 4.3: Establishing a promastigote (PRO) and an amastigote (AMA) infection model for *L. donovani* infection of C57BL/6J mice. A). Schematic illustration of the experimental protocol for the PRO and AMA infection set-up. Mice were infected intravenously with 1x10^8 barcoded LV9 T7 cas9 *L. donovani* stationary phase promastigote parasites (PRO) or 5x10^7 amastigotes (AMA) consisting of a library of 102 isogenic 20bp barcodes. AMA mice used parasites in the amastigote life stage, after harvest from B6.*Rag2*$^{-/-}$ CD45.1 spleens. PRO mice: n=5 males and 5 females for all time-points. AMA: n=5 males and 5 females for day 2, and 5 females each for day 14 and day 28. Mice were killed at time-points indicated and tissues were flash frozen on dry ice, then stored at -80 °C before subsequent DNA extraction, Illumina sequencing, Barseq analysis and population dynamics studies, alongside a sample of the input parasite pool from each infection. B and C). Spleen parasite burdens calculated as LDU (B) and spleen weight as a proportion of total body weight (C) in PRO and AMA mice at 14 (green) and 28 (yellow) days p.i. Each point is an individual mouse. One AMA mouse spleen weight is missing due to technical error, but was assumed to be the mean of the experimental group for LDU calculation. D and E). Liver parasite burdens calculated as LDU (D) and liver weight as a proportion of total body weight (E) in PRO and AMA mice at 14 (green) and 28 (yellow) days p.i. Two AMA mice liver weights are missing due to technical error, but were assumed to be the mean of the experimental group for LDU calculation. Mann-Whitney-Wilcoxon test. Significance of p-value: * p<0.05; **p<0.01; *** p<0.001: **** p<0.0001. Boxplots show the data median, 25th and 75th quartiles, with whiskers representing all data-point within 1.5x Interquartile Range of the median; points above or below the whiskers represent outlying data values. Panel A created in Biorender.

4.2.2 Parasites undergo a bottleneck to colonisation and the liver is an early colonisation site in both promastigote and amastigote VL infection models

To quantify the relative spatial and temporal bottlenecks to colonisation following promastigote or amastigote infection, the FPs in each tissue of each mouse at each time-point in PRO and AMA mice were calculated, using the STAMP method described by Abel et al., 2015 (see Methods for equations). Our hypothesis was that there would be tissue-specific differences in bottlenecks, which may also differ by the infection model (i.e. dependent on the life stage at infection).

Figure 4.4 shows the spatial diversity and change over the timecourse of infection in FP in the liver, spleen, bone marrow, lymph nodes, gut and lung for AMA and PRO

models. AMA mice showed early colonisation of the liver and spleen, with FP below 11 for all other tissues at 2 days p.i. The spleen showed no significant difference between mice infected for 2 days and those infected for 28 days, indicating that the increase in LDU observed between mice killed at day 14 and at day 28 is likely to be predominantly a result of replication of the initial colonising population, rather than establishment of new individuals. In mice infected for 28 days, there was an approximately six-fold increase in liver FP compared to mice infected for only 2 days (MW: p=0.014) (Figure 4.4).

In the 14 days p.i mice, the calculated FP in the bone marrow was highly variable (2 - 644) and not significantly different from the day 2 p.i mice. In mice infected for 28 days, calculated bone marrow FP was more tightly clustered and was significantly different from mice infected for 2 days (MW: p<0.0001).

Lymph node FP was approximately seven-fold higher in mice infected for 28 days compared to those infected for 2 days (MW: p=0.005), whereas lung FP was greater in mice infected for 14 days, compared to those infected for 2 days although remaining very low in both cases (MW: p=0.038). The gut showed no significant difference between infection time points. No statistically significant differences were observed between male and female AMA mice at 2 days p.i, except in lungs (p=0.0080) (Figure 4.5A), but no comparisons are available for the two later time-points.

In the PRO mice, a broader variance in colonisation was observed (Figure 4.4). The liver had the highest FP of all tissues at 2 days p.i and this was four-fold higher when mice were infected for 28 days (MW_{2-28}: p=0.009). The spleen showed a significantly higherFP in mice infected for 14 or 28 days compared to those infected for 2 days (MW_{2-14}: p=0.006 and MW_{2-28}: p<0.0001) and between mice killed on days 14 and 28 p.i (MW_{14-28}: p=0.009). At day 28 spleen FP was approximately four-fold lower than that of the liver. A significant increase in FP across all time-points was also observed in the lymph nodes, with an approximately 10-fold rise in FP between mice infected for 2 days and those infected for 28 days (MW_{2-28}: p<0.0001). In the bone marrow, the FP was also significantly higher at in mice infected for 14 and 28 days, compared to mice infected for 2 days (MW_{2-14}: p<0.0001 and MW_{2-28}:

$p<0.0001$), with a three-fold increase in average FP at day 28 p.i compared to at day 2.

For the lung and gut, no samples were evaluated at 2 days p.i, but both tissues showed evidence of colonisation at day 14, with FP higher than any other tissue except the liver. However, FP did not significantly change between mice infected for 14 and 28 days in either tissue.

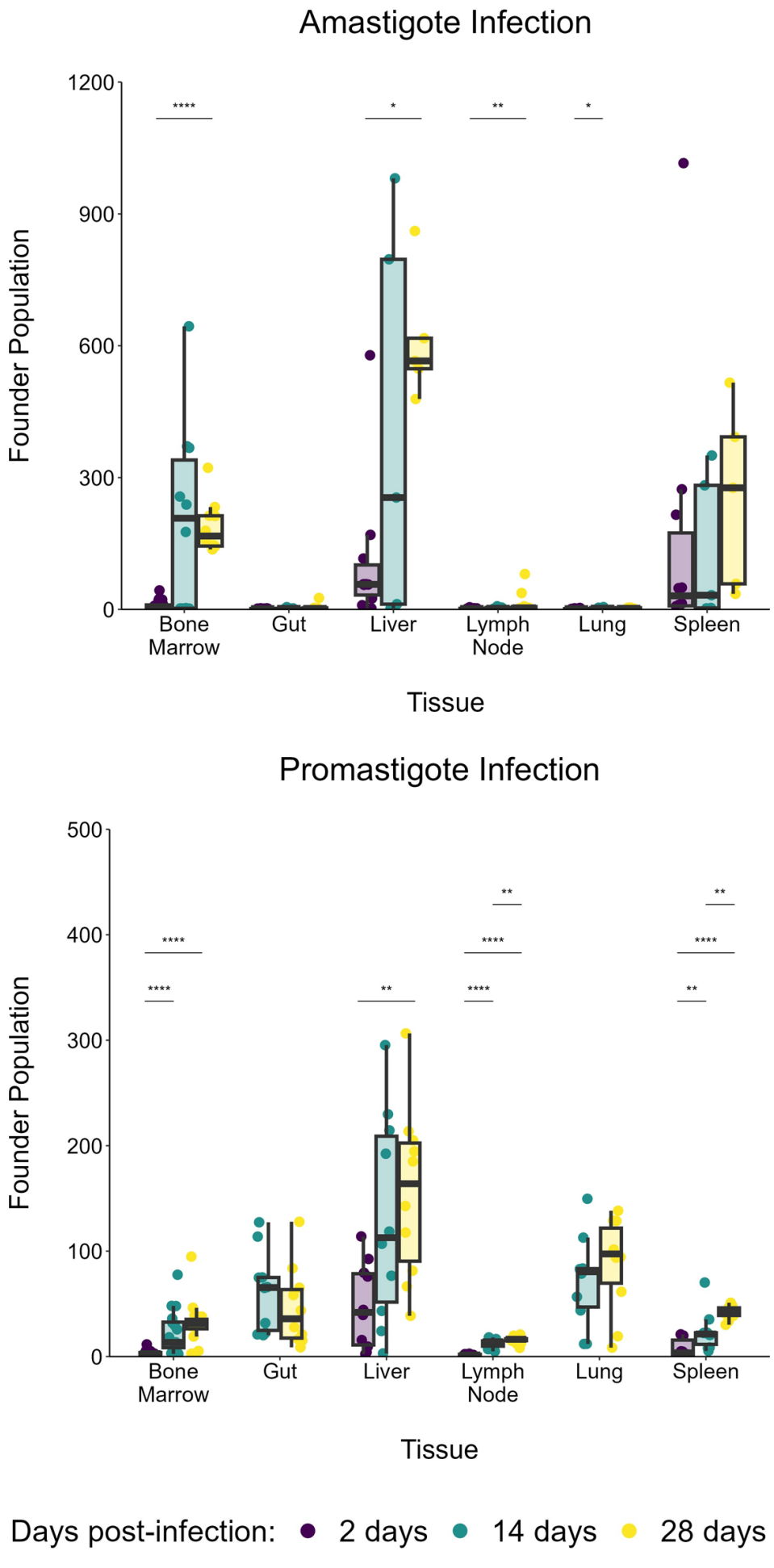
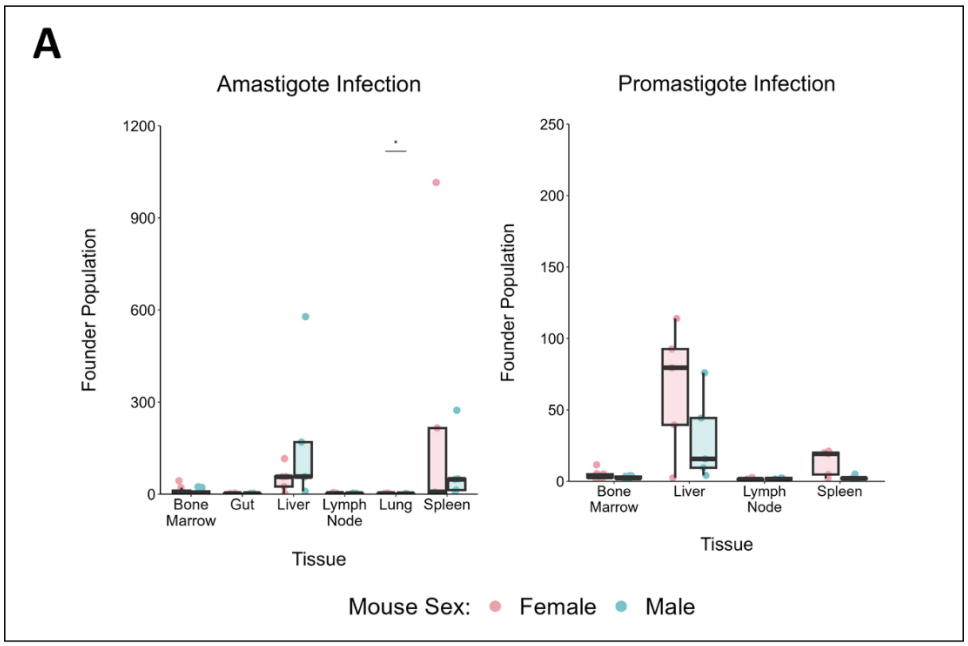

Amastigote Infection
Founder Population
1200
900
600
300
0
****
*
**
*
Bone Marrow
Gut
Liver
Lymph Node
Lung
Spleen
Tissue
Promastigote Infection
Founder Population
500
400
300
200
100
0
****
****
**
****
****
**
**
****
**
Bone Marrow
Gut
Liver
Lymph Node
Lung
Spleen
Tissue
Days post-infection:
2 days
14 days
28 days

Figure 4.4: Bottlenecks to colonisation in different tissues in C57BL/6J mice infected with amastigote and promastigote *L. donovani* parasites at 2 days, 14 days and 28 days post-infection. Founder population sizes, calculated by STAMP (Abel et al., 2015) in an amastigote (AMA) and promastigote (PRO) infection at 2 days, 14 days and 28 days post-infection. For PRO infections, n= 5 males and 5 females at each time-point. For AMA infections, n=5 males and 5 females at 2 days, and 5 females only at each of 14 and 28 days. Colour scheme indicates time-point. Each data point represents an individual tissue from each mouse. Mann-Whitney-Wilcoxon test, with Bonferroni correction for time-point comparisons. Significance of (adjusted if applicable) p-value: * $p<0.05$; ** $p<0.01$; *** $p<0.001$: **** $p<0.0001$. Boxplots show the data median, 25^{th} and 75^{th} quartiles, with whiskers representing all data-point within 1.5x Interquartile Range of the median; points above or below the whiskers represent outlying data values.

When split by sex, similar patterns of change were seen over time in PRO mice, with no significant differences in FP observed (Figure 4.5B).


A
Amastigote Infection
Promastigote Infection
Founder Population
1200
900
600
300
0
250
200
150
100
50
Bone Marrow
Gut
Liver
Lymph Node
Lung
Spleen
Tissue
Mouse Sex: Female Male

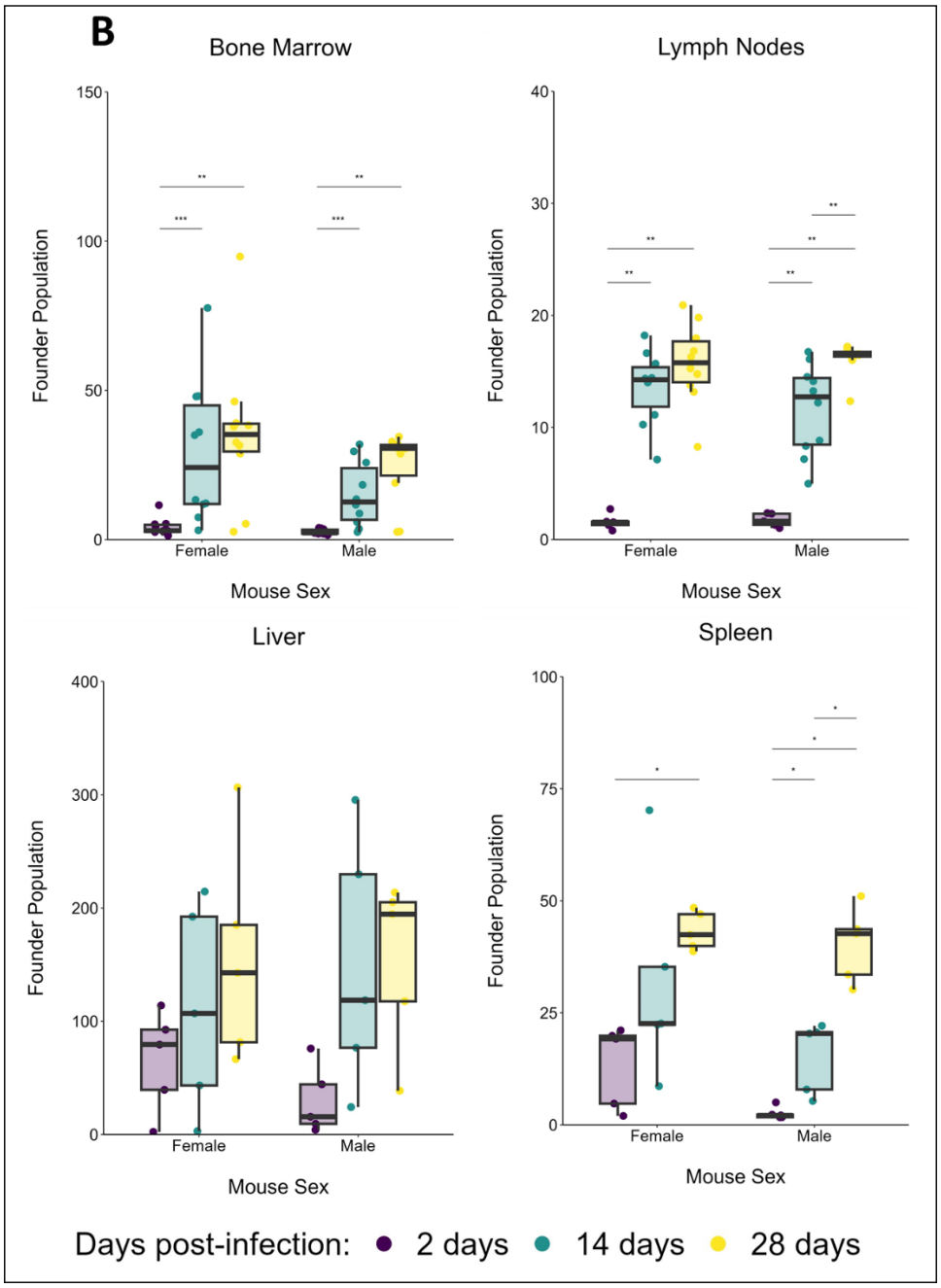

B
Bone Marrow
Lymph Nodes
Liver
Spleen
Founder Population
Female
Male
Mouse Sex
Days post-infection: 2 days 14 days 28 days

Figure 4.5: Bottlenecks to colonisation in different tissues in C57BL/6J mice, split by sex after infection with amastigote and promastigote *L. donovani* parasites at 2 days, 14 days and 28 days post-infection. A). Comparison of founder population size in male and female PRO and AMA mice at 2 days post-infection. B). Sex split comparisons of founder populations in PRO mice at 2 days, 14 days and 28 days p.i, for bone marrow, lymph node, liver and spleen. For PRO infections, n= 5 males and 5 females at each time-point. For AMA infections, n=5 males and 5 females at 2 days. Colour scheme indicates mouse sex in panel A and time-point in panel B. Each data point represents an individual tissue from each mouse. Mann-Whitney-Wilcoxon test, with Bonferroni correction for time-point comparisons. Significance of (adjusted if applicable) p-value: * $p<0.05$; ** $p<0.01$; *** $p<0.001$: **** $p<0.0001$. Boxplots show the data median, 25^{th} and 75^{th} quartiles, with whiskers representing all data-point within 1.5x Interquartile Range of the median; points above or below the whiskers represent outlying data values.

In summary, these data indicate that the liver is relatively permissive to both promastigote and amastigote infection compared to other tissues. In AMA mice, the spleen and bone marrow were also relatively permissive, with very restrictive bottlenecks in the lung, gut and lymph nodes. In contrast, PRO mice had proportionally greater increases in colonisation over time in most tissues, with the lung and the gut being surprisingly permissive. Sex appears to have little influence over tissue colonisation in C57BL/6J mice.

4.2.3 Parasite life stage on infection impacts the relationship between bottleneck size and subsequent population diversity

To investigate whether the diversity of the populations varied between tissues and over time, we assessed alpha-diversity using the SI, with barcode counts representing species abundances. A higher SI indicates a diverse and relatively even distribution of “species”, whilst a lower value can indicate a low total “species” number or a highly skewed distribution, with some individuals comprising a high proportion of the population and others present being very few in number. Figure 4.6 shows the SI at 2, 14 and 28 days p.i in each tissue for AMA and PRO mice.

AMA mice showed a relatively high diversity in all tissues except the liver at 2 days p.i. Diversity in the lung, lymph node and gut was consistent over the 28 day infection, but diversity decreased significantly in bone marrow ($MW_{2\text{-}14}$: p=0.008 and $MW_{2\text{-}28}$: p<0.0001) and liver ($MW_{2\text{-}28}$: p=0.002), indicating a skew in the numbers of each barcoded line over time. No significant changes in SI in the spleen were observed.

In PRO mice, the liver had a relatively high SI at all time-points, indicating that the diversity and evenness of the population did not change significantly over the course of infection, despite the increase in FP over time. Other tissues initially showed lower diversity at 2 days p.i, compared to liver, but diversity increased over time in the bone marrow ($MW_{2\text{-}14}$: p<0.0001 and $MW_{2\text{-}28}$: p<0.0001), spleen ($MW_{2\text{-}14}$: p=0.027) and lymph nodes ($MW_{2\text{-}14}$: p<0.0001 and $MW_{2=28}$: p<0.0001 and $MW_{14\text{-}28}$: p=0.024). The gut and the lung did not have data for day 2, but showed a consistently high diversity, similar to liver, at 14 and 28 days p.i.

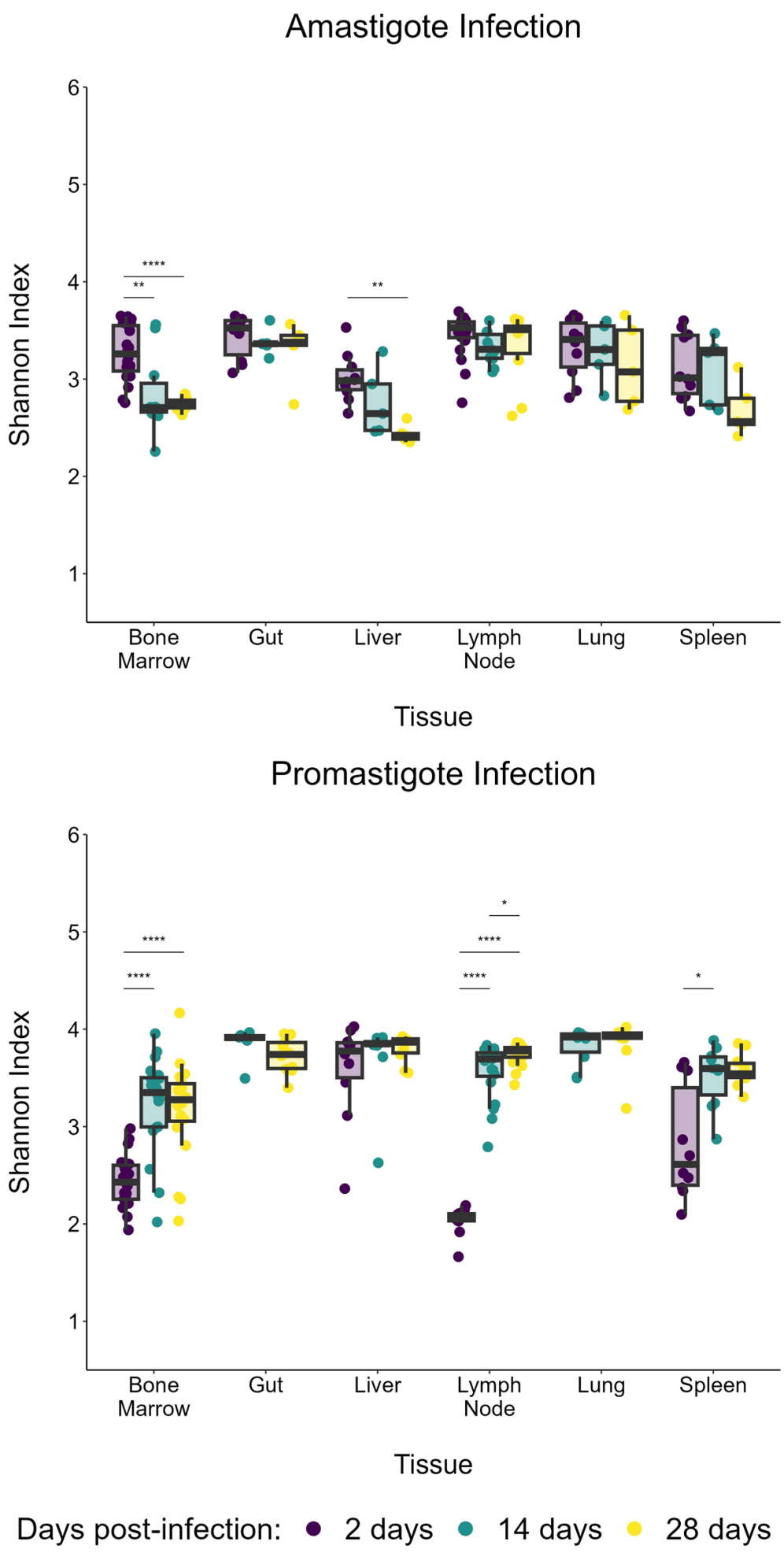

Amastigote Infection
Shannon Index
6
5
4
3
2
1
****
**
**
Bone Marrow
Gut
Liver
Lymph Node
Lung
Spleen
Tissue
Promastigote Infection
Shannon Index
6
5
4
3
2
1
****
****
****
*
*
Bone Marrow
Gut
Liver
Lymph Node
Lung
Spleen
Tissue
Days post-infection:
2 days
14 days
28 days

Figure 4.6: Population diversity quantified by Shannon Index in different tissues in C57BL/6J mice infected with amastigote or promastigote *L. donovani* parasites at 2 days, 14 days and 28 days post-infection. A). Shannon Index in an amastigote (AMA) and promastigote (PRO) infection at 2 days, 14 days and 28 days post-infection. For PRO infections, n= 5 males and 5 females at each time-point. For AMA infections, n=5 males and 5 females at 2 days, and 5 females only at each of 14 and 28 days. Colour scheme indicates time-point. Each data point represents an individual tissue from each mouse. Mann-Whitney-Wilcoxon test, with Bonferroni correction for time-point comparisons. Significance of (adjusted if applicable) p-value: * p<0.05; **p<0.01; *** p<0.001: **** p<0.0001. Boxplots show the data median, 25th and 75th quartiles, with whiskers representing all data-point within 1.5x Interquartile Range of the median; points above or below the whiskers represent outlying data values.

At 2 days p.i no significant differences were found between male and female mice for any time-points in either infection model (Figure 4.7A). Diversity in the spleen increased from day 2 to day 14 and day 28 in male PRO mice ($MW_{2\text{-}14}$: p=0.024 and $MW_{2\text{-}28}$: p=0.024), but remained similar in female PRO mice, with a far wider range of values at 2 days p.i in the females compared to the males (Figure 4.7B). The trajectories over time for other tissues were similar in both sexes.

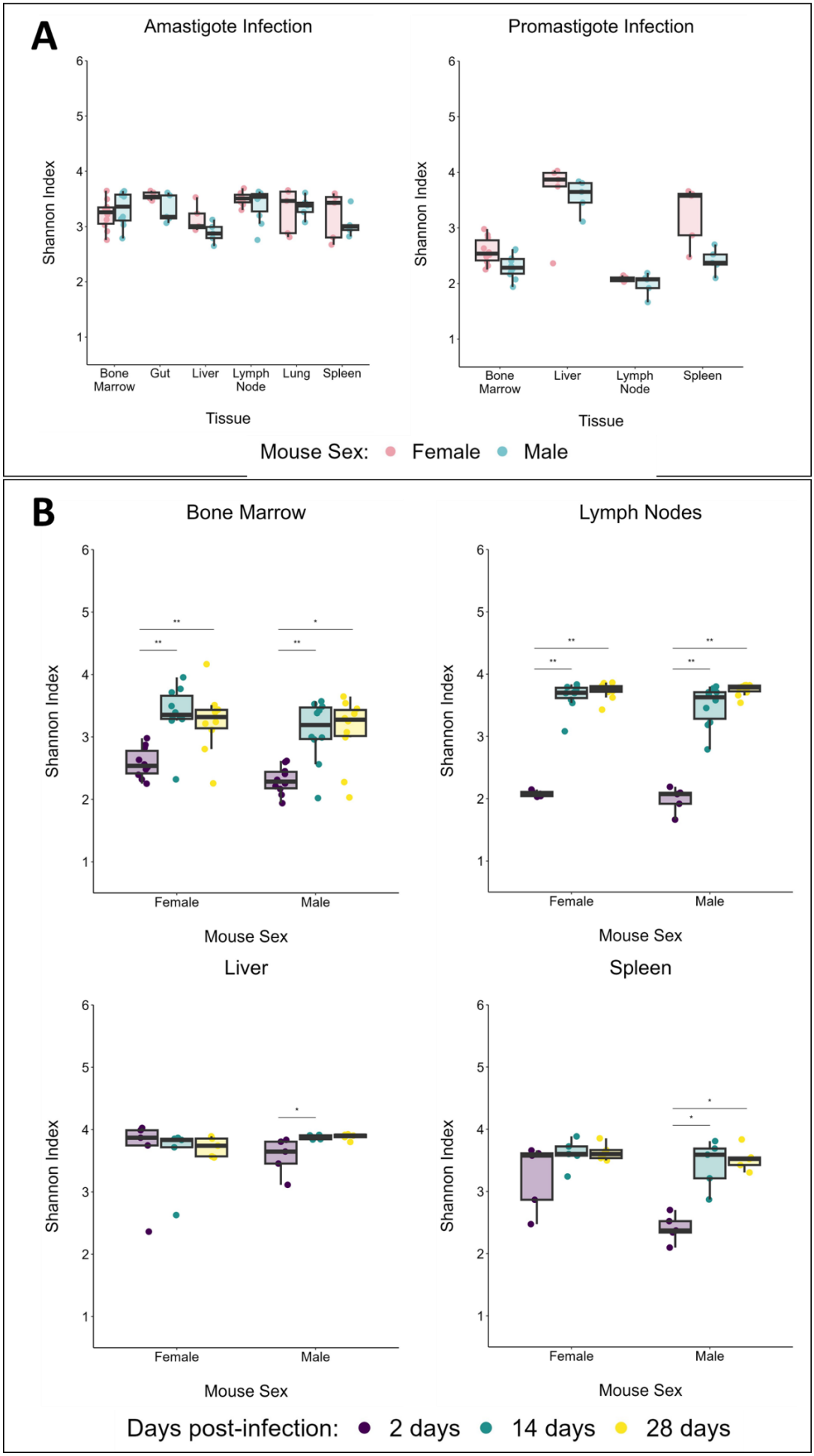

A
Amastigote Infection
Promastigote Infection
Shannon Index
Bone Marrow
Gut
Liver
Lymph Node
Lung
Spleen
Tissue
Mouse Sex: Female Male
B
Bone Marrow
Lymph Nodes
Liver
Spleen
Female
Male
Mouse Sex
Days post-infection: 2 days 14 days 28 days

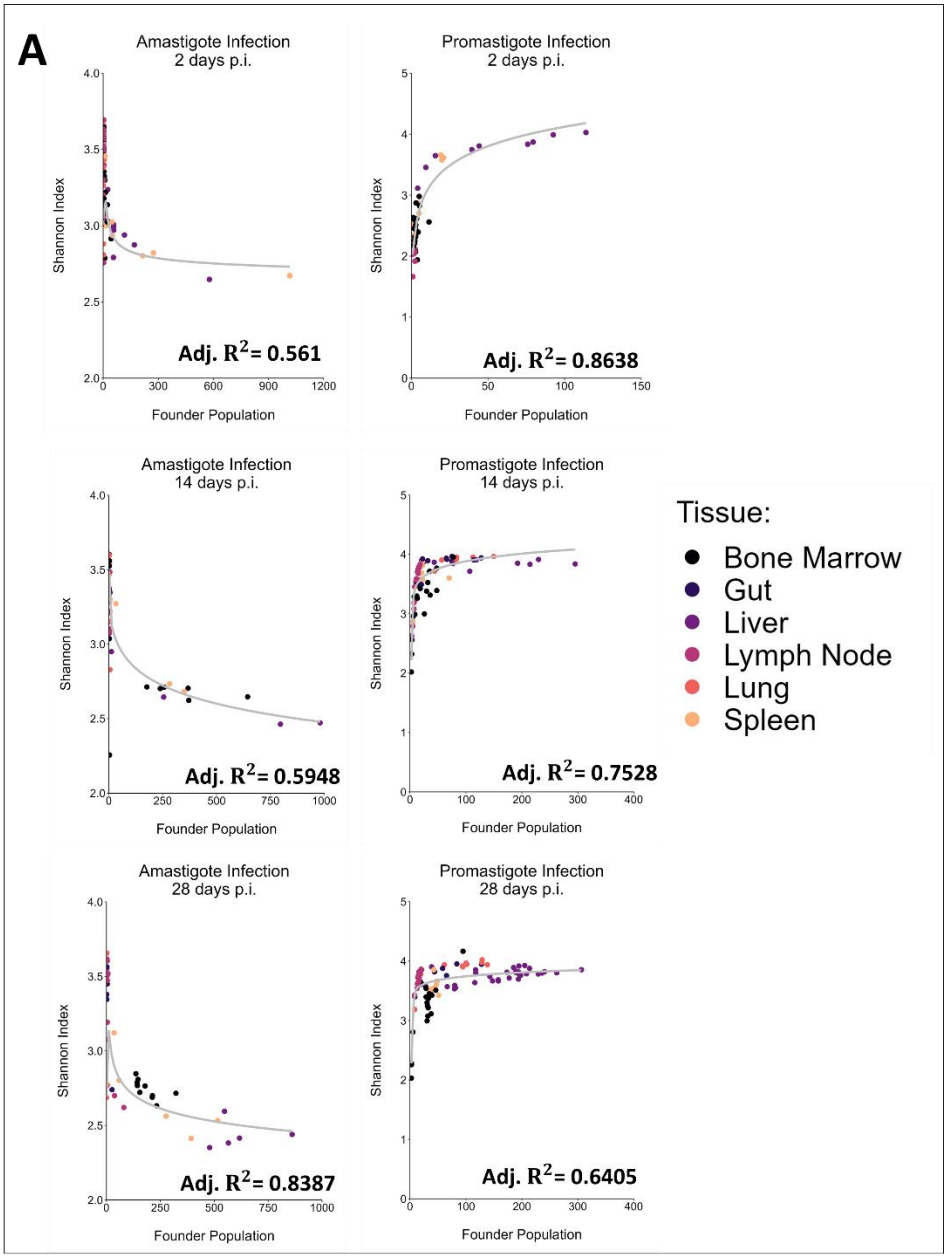
Figure 4.7: Population diversity quantified by Shannon Index in different tissues in C57BL/6J mice, split by sex, after infection with amastigote or promastigote *L. donovani* parasites at 2 days, 14 days and 28 days post-infection. A). Comparison of Shannon Index size in male and female AMA and PRO mice infection at 2 days post-infection. B). Sex split comparisons of Shannon Index change over time for PRO mice for bone marrow, lymph node, liver and spleen. Colour scheme indicates mouse sex in panel A and time-point in panel B. For PRO infections, n= 5 males and 5 females at each time-point. For AMA infections, n=5 males and 5 females at 2 days. Each data point represents an individual tissue from each mouse. Mann-Whitney-Wilcoxon test, with Bonferroni correction for time-point comparisons. Significance of (adjusted if applicable) p-value: * $p<0.05$; ** $p<0.01$; *** $p<0.001$: **** $p<0.0001$. Boxplots show the data median, 25th and 75th quartiles, with whiskers representing all data-point within 1.5x Interquartile Range of the median; points above or below the whiskers represent outlying data values.

To understand the relationship between bottleneck size and population diversity, we compared FP and SI to each other for each infection model. Following Jiang et al., 2024, we hypothesised that bottleneck size would be related to diversity, such that as FP increased so would diversity, at an exponentially decaying rate that fits an asymptotic relationship.

Figure 4.8A shows the SI plotted against the FP value for each tissue in AMA and PRO mice. PRO mice exhibited the anticipated asymptotic relationship, with increased FP resulting in increased diversity at all time-points, and a reduction in this effect as FP increased. Surprisingly, AMA mice showed the opposite relationship, with increased FP being associated with a reduced overall diversity, but the reduction in impact of FP increases as FP rose was consistent across both infection models.

A linear model fitted to the data from each infection model at each time-point (Figure 4.8B), with the formula $y_{PRO}=\log(x)*e^{-x}$ or $y_{AMA}=\log(x)*e^{-1/x}$, where x is FP and y is SI, was significant in both infection models at all time-points ($p<0.0001$ in all cases). The parameters fitted at each time-point varied and resulted in the limit of diversity at higher FPs decreasing over time in both models. Adjusted R-squared values are shown for each model fit. The models indicate that a higher FP earlier in infection results in a higher overall diversity, following both promastigote and

amastigote infections, and that addition of new individuals later in infection will result in a lowering of the overall possible diversity which can be achieved in the infected tissue (Figure 4.8B).


A
Amastigote Infection
2 days p.i.
Shannon Index
Founder Population
Adj. R2= 0.561
Promastigote Infection
2 days p.i.
Shannon Index
Founder Population
Adj. R2= 0.8638
Amastigote Infection
14 days p.i.
Shannon Index
Founder Population
Adj. R2= 0.5948
Promastigote Infection
14 days p.i.
Shannon Index
Founder Population
Adj. R2= 0.7528
Amastigote Infection
28 days p.i.
Shannon Index
Founder Population
Adj. R2= 0.8387
Promastigote Infection
28 days p.i.
Shannon Index
Founder Population
Adj. R2= 0.6405
Tissue:
Bone Marrow
Gut
Liver
Lymph Node
Lung
Spleen

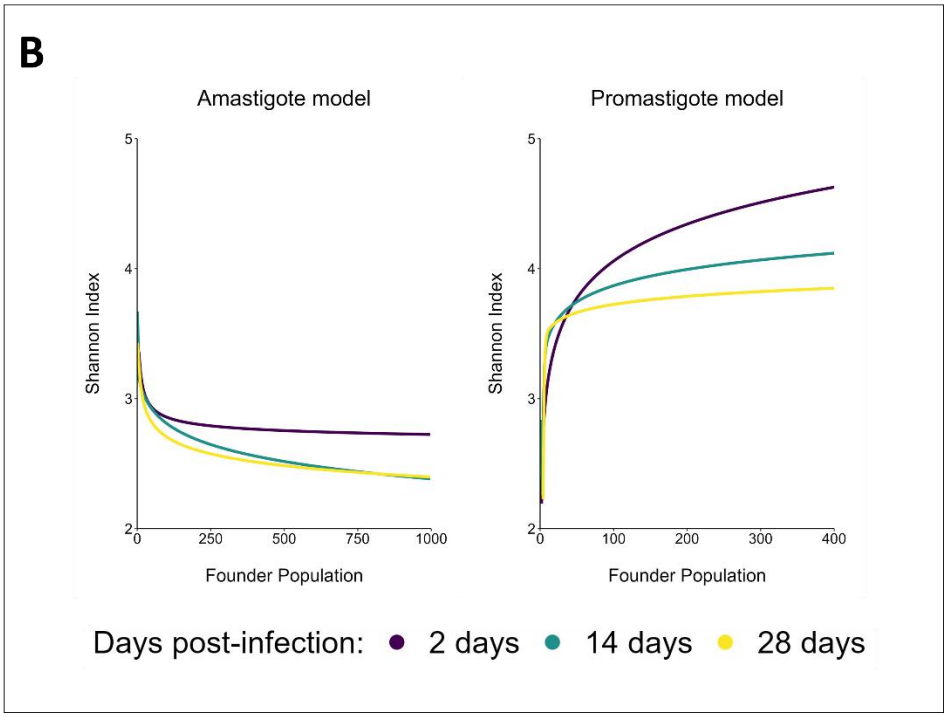

B
Amastigote model
Shannon Index
Founder Population
Promastigote model
Shannon Index
Founder Population
Days post-infection:
2 days
14 days
28 days

Figure 4.8: The relationship between bottleneck size and population diversity after infection with amastigote or promastigote *L. donovani* parasites. A). The relationship between Founder Population and Shannon Index at 2, 14 and 28 days p.i in amastigote (AMA) and promastigote (PRO) infected mouse tissues. Grey lines indicate the linear model fitted to the data (see Methods for details), using the formula $y_{PRO}=\log(x)*e^{-x}$ or $y_{AMA}=\log(x)*e^{-1/x}$, with the model R^2 for each fit indicated on the plot. All fits had $p<0.0001$. B). The linear models fitted for the relationship between Shannon Index and Founder Population in AMA and PRO mouse tissues. Colour indicates tissue in A. Colour indicates day post-infection in B.

These data reveal the intriguing finding that *L. donovani* behaviour *in vivo* differs starkly at the population level, even at 4 weeks p.i, when different parasite life stages are used for the initial infection.

4.2.4 Tissue-seeding dynamics differ depending on parasite stage at infection

Next, the barcode proportions in each tissue were compared in a pairwise manner to assess their relatedness via genetic distance (GD), as described in Abel et al., 2015 (see Methods for details). GD is one measure of beta diversity, providing information on how similar two populations are and as applied here, changes in GD between tissues over time should provide an indication of the extent of parasite dissemination.

Interestingly, AMA mice exhibited a pattern characterising two sub-groups of tissues in terms of their GD values (Figure 4.9A). The GD between bone marrow, spleen and liver was consistently low compared to gut, lung and lymph nodes. Notably, lymph node, gut and lung GDs were also lower when compared with tissues, suggesting that parasites disseminate in differing circuits comprising the two tissue groups.

Comparing day 2 and day 28, AMA mice showed a significant increase in relatedness (reduced GD) over time for bone marrow to spleen ($MW_{2\text{-}28}$: $p=0.0008$) and bone marrow to liver comparisons ($MW_{2\text{-}28}$: $p=0.0006$), but not for any other pairwise tissue comparisons. In fact, GD increased between bone marrow to gut and bone

marrow to lymph node, likely as a result of genetic drift in non-mixing populations. GD did not change over time for any other paired tissue comparisons.

In contrast, in PRO mice (Figure 4.9B) most tissues had relatively similar GDs to all other tissues, with the exception of the lymph node to liver interaction at the earliest time-point, which was on average far higher than other tissues paired with lymph nodes at the same time-point, indicating a lack of relatedness in these tissues early in infection. This GD decreased such that by days 14 and 28, it was similar to that of liver to bone marrow and liver to spleen.

Between day 2 and 28, PRO mice showed a significant increase in relatedness for liver to spleen ($MW_{2\text{-}28}$: p=0.003), bone marrow ($MW_{2\text{-}14}$: p=0.015 and $MW_{2\text{-}28}$: p=0.002) and lymph nodes ($MW_{2\text{-}14}$: p=0.004 and $MW_{2\text{-}28}$: p=0.004), over the course of infection. No other pairs became more closely related over time; the lymph nodes overall became less related to both bone marrow and spleen, whilst no other tissues showed a significant change in GD to other tissues over time.

There was no change in GD for any lung or gut comparisons in PRO mice, but gut and lung appeared to be highly related to each other, with low GD for a pairwise comparison of the two. They also appeared to be closer to the liver overall than to bone marrow, lymph node or spleen, indicating a potential gut-lung-liver relationship in PRO mice.

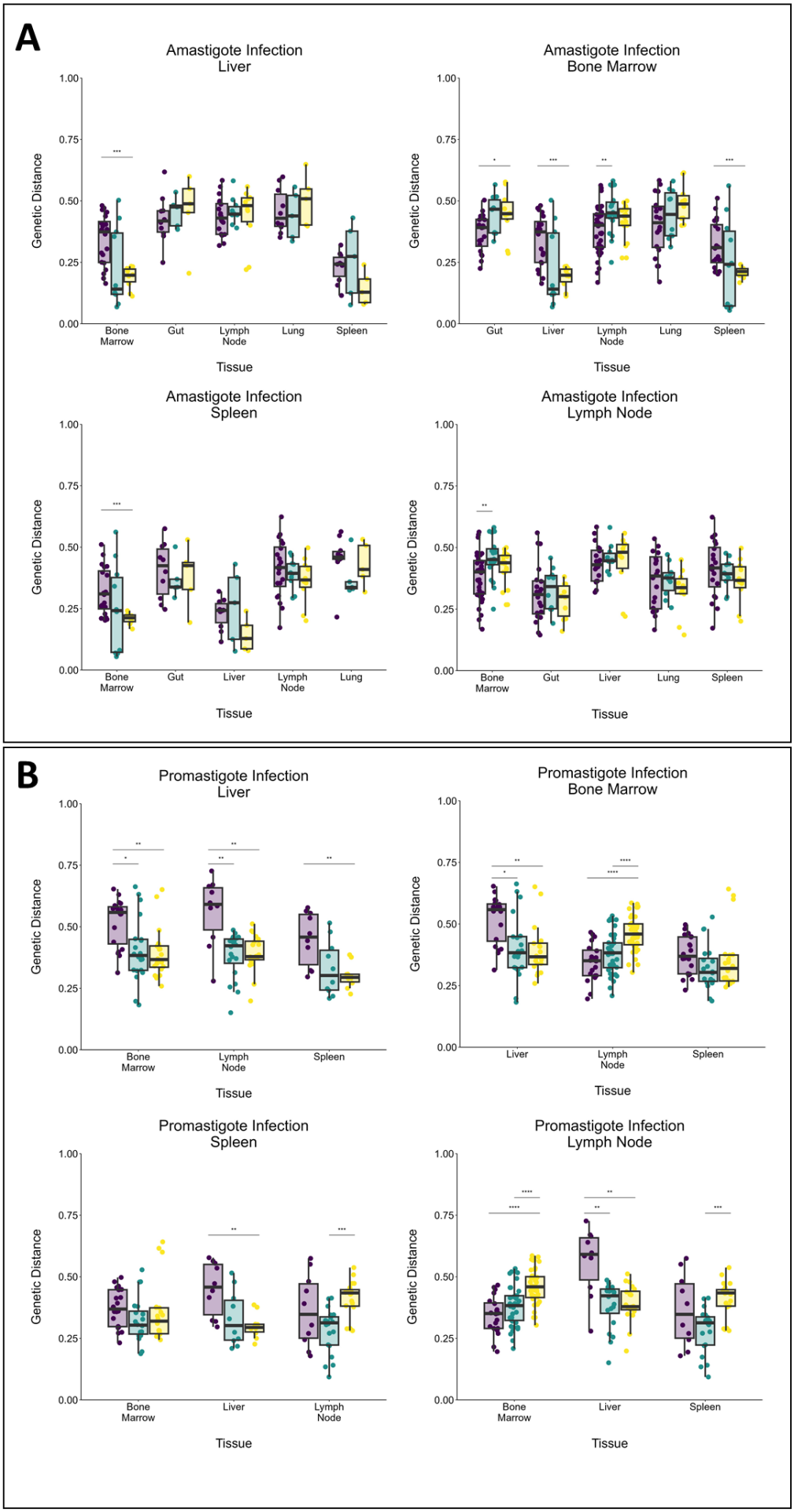

A
Amastigote Infection
Liver
Amastigote Infection
Bone Marrow
Amastigote Infection
Spleen
Amastigote Infection
Lymph Node
Genetic Distance
1.00
0.75
0.50
0.25
0.00
Bone Marrow
Gut
Liver
Lymph Node
Lung
Spleen
Tissue
B
Promastigote Infection
Liver
Promastigote Infection
Bone Marrow
Promastigote Infection
Spleen
Promastigote Infection
Lymph Node
Genetic Distance
Tissue
Days post-infection: 2 days 14 days 28 days

Figure 4.9: Genetic distance analysis of amastigote or promastigote *L. donovani* infection parasite populations in different tissues in C57BL/6J mice at 2, 14 and 28 days p.i. A). Genetic distance, calculated as described in Abel et al., 2015 using the equations originally defined in Cavalli-Sforza and Edwards, 1967, for pairwise comparisons of each tissue in AMA mice for 2 day (n=10), 14 day (n=5) and 28 (n=5). B). Genetic distance for pairwise comparisons of each tissue in PRO mice infected for 2 days (n=10), 14 days (n=10) and 28 days (n=10). Colour indicates days post-infection. Each data point represents a pairwise comparison between two tissues. As such, the reverse pairwise comparison, in the corresponding switched direction plot (i.e. the x-axis tissue is the title tissue and vice versa), is identical. Mann-Whitney-Wilcoxon test with Bonferroni correction. Significance of adjusted p-value: * p<0.05; **p<0.01; *** p<0.001: **** p<0.0001. Boxplots show the data median, 25th and 75th quartiles, with whiskers representing all data-point within 1.5x Interquartile Range of the median; points above or below the whiskers represent outlying data values.

Heterogeneity in relatedness of different tissues was observed between individual mice, but some trends emerged when mean GD values for each comparison were plotted as heat maps for each experimental group (Figure 4.10). At day 2 in AMA mice, there was an overall higher relatedness between tissues in female mice compared to male mice (Figure 4.10A). However, over time female mouse tissues became less related in many pairs. Of note, lung and gut became less related to liver, spleen and bone marrow.

Tissues in PRO mice were less related at 2 days p.i than AMA mice, particularly in liver relationships to other tissues, and this was particularly pronounced in female mice (Figure 4.10B). Over time, relatedness increased in both sexes. The lung, gut and liver were highly related at 14 and 28 days p.i in both sexes. The lymph nodes were highly related to gut and lung across both models and at all time-points for which data for all three tissues was available.

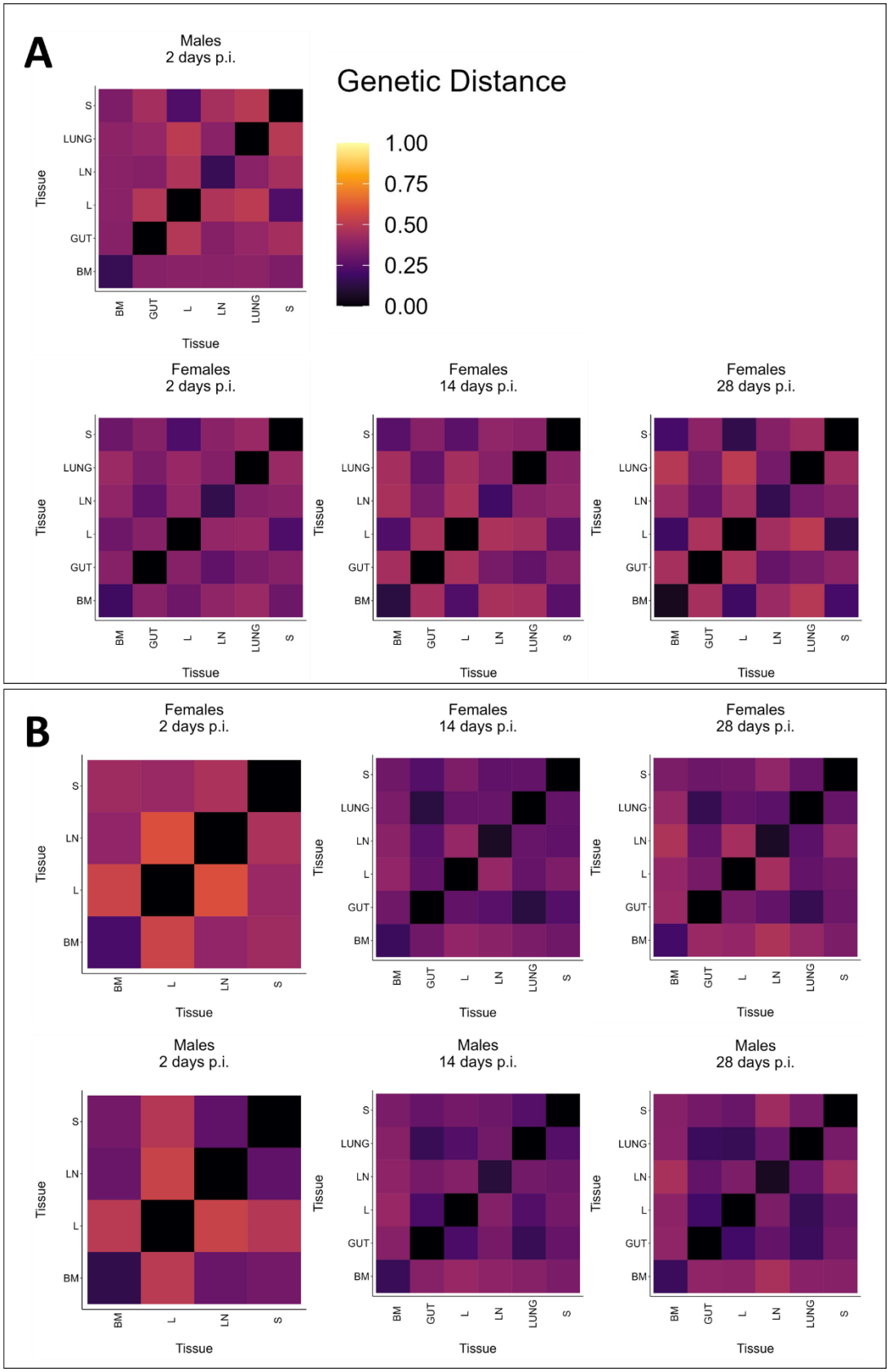


Figure 4.10: Genetic distances between tissue parasite populations for the average amastigote or promastigote *L. donovani* infected mice at 2, 14 or 28 days p.i. A). Heat maps showing the average (mean) genetic distance for each pairwise tissue comparison in male and female amastigote-infected (AMA) mice at 2 days, and female mice only at 14

days and 28 days p.i. B). Heat maps showing the average (mean) genetic distance for each pairwise tissue comparison in male and female promastigote-infected (PRO) mice at 2 days, 14 days and 28 days p.i. For all experimental groups, n=5. All bone marrow data comprises two sets of bone marrow per mouse, and as such the GD is not 0 for these comparisons. The same applies for lymph nodes of all groups except day 2 p.i PRO mice, which comprise one lymph node per mouse only. Abbreviations: S = Spleen; L = Liver; LN = Lymph Node; BM = Bone Marrow. Colour scale is the same for all heat maps. Relatedness can be considered as 1-GD, so lower GD scores indicate higher relatedness.

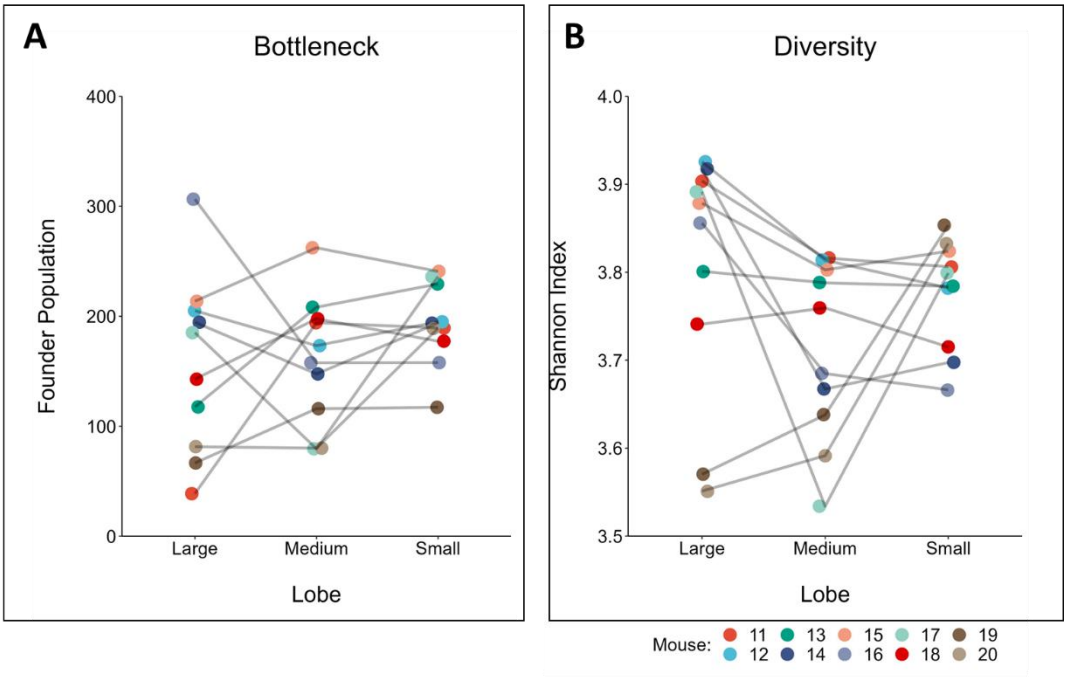
Overall, the findings here provide evidence for the importance of the liver in promastigote infections, with indications that it may be seeding parasites to other tissues, evidenced by the increasing relatedness over time to liver in the promastigote-infected mice. The data also provide support for a possible liver-lung-gut axis in the promastigote, but not amastigote, infections. In contrast, amastigote infections appear to exhibit a partitioning of parasite movement such that liver, bone marrow and spleen form one network and the other tissues a second network, with little interaction between them.

4.2.5 Liver lobes show consistent colonisation and population diversity after promastigote infection

We have noted small differences between individual lymph nodes and bone marrow samples in individual mice, but the possibility of localised populations in different regions of a tissue has not been studied before for *L. donovani*. Previous studies have found heterogeneity in blood supply both to and from the individual liver lobes in rodents, which could influence liver colonisation patterns (Malarkey et al., 2005; Sänger et al., 2015). We therefore examined GD between separate liver lobes from the same mouse, to determine whether parasite populations might be seeded via different colonisation events, with little intermixing (i.e. there would be high GD between liver lobes).

Analysis of three liver lobes in PRO mice at day 28 p.i indicated no significant difference in FP (Figure 4.11A) or SI (Figure 4.11B). For comparisons of each lobe to each other lobe and to the other tissues from the same mouse, little difference in

GD was noted. There was a relatively low GD between the three lobes themselves. Few significant differences were observed in any comparisons (medium lobe to the other two liver lobes MW: p=0.043; lymph node to the two larger lobes MW: p=0.001).



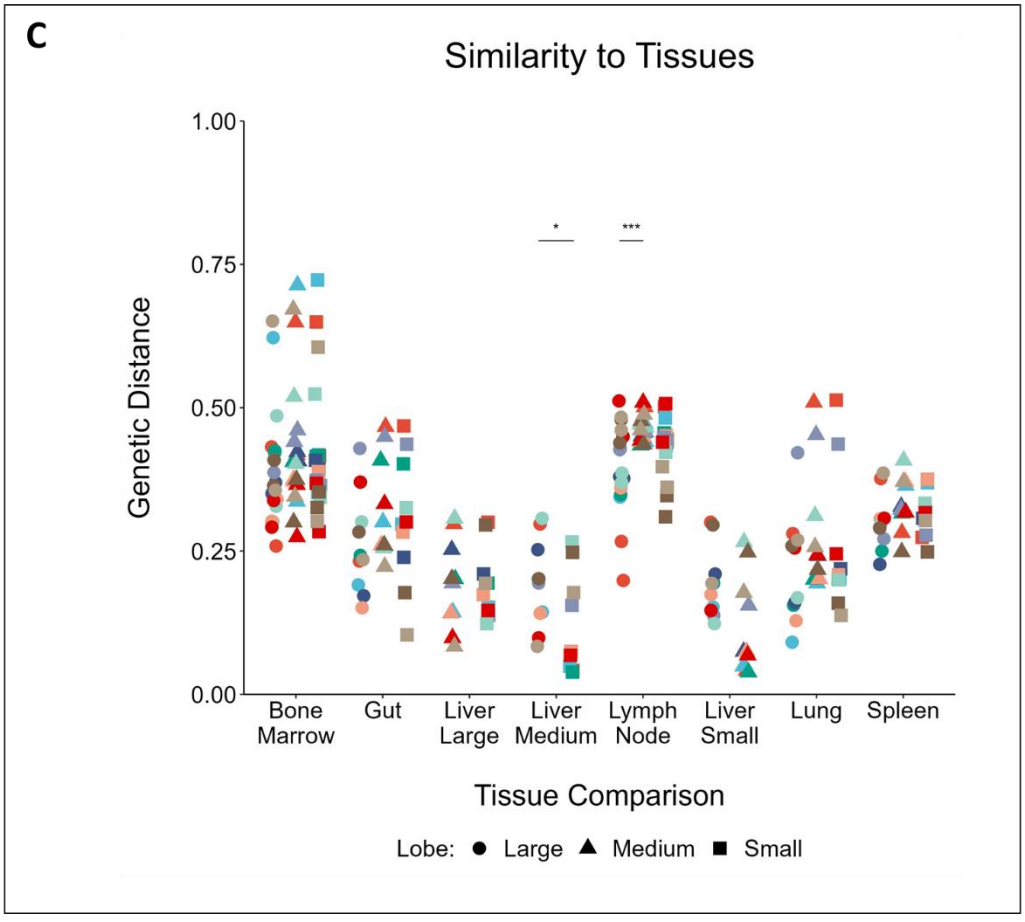


Figure 4.11: Population dynamics within individual liver lobes of promastigote *L. donovani* infected mice at 28 days p.i. A). Founder population size for each lobe of the liver for 10 mice at 28 days p.i. B). Shannon Index for each lobe of the liver for 10 mice at 28 days p.i C). Genetic distance for pairwise tissue comparisons with each liver lobe for 10 individual mice at 28 days p.i. Shape indicates the liver lobe being compared to the tissue on the x-axis. Colours indicate each mouse and are consistent across all panels. Grey lines

link lobes from the same mouse. Mann-Whitney-Wilcoxon test (with Bonferroni correction for genetic distances). Significance of (adjusted if applicable) p-value: * $p<0.05$; ** $p<0.01$; *** $p<0.001$: **** $p<0.0001$.

The data overall indicate that parasite populations were relatively homogenous across the liver. This provides confidence that small samples are representative of parasite populations within larger tissues, where the entire organ cannot be sequenced for practical reasons.

4.2.6 Network analysis reveals differential centrality of tissues in different model systems

To further parse out likely dissemination paths and possible connections between different tissues, MSTs were constructed for each mouse, using GD as the basis for distances and a maximum likelihood estimate method, along with Prim's algorithm for determining the connected vertices. A Louvain clustering algorithm was applied to provide indications of compartmentalisation within the network. Exemplar network diagrams are shown for one female AMA and one female PRO mouse at 14 days p.i in Figure 4.12A. Other MST networks are available in Appendix Figure 1.

Qualitatively, PRO mice showed evidence of increasing compartmentalisation over time. The number of PRO mouse networks exhibiting fewer than 3 clusters decreased between days 14 and 28. The spleen and at least one bone marrow population frequently appeared in the same cluster, as did liver and lung or liver and gut (or all three together). In AMA mice, one mouse network had fewer than three clusters at 14 days p.i, whilst two mice did at 28 days p.i., meaning compartmentalisation did not increase between these time-points. Liver appeared to regularly cluster with bone marrow in these mice, across all time-points.

Each individual network was then quantified using two metrics: harmonic centrality (Freeman, 1979; Marchiori and Latora, 2000) and betweenness centrality (Brandes et al., 2000). The concept of both is illustrated in Figure 4.12B. Briefly, the harmonic centrality of a vertex within a network is a measure of the total information which flows through a vertex. For each vertex, it is calculated as the inverse of the mean

path length to every other vertex in the network. In this case, it is used to represent the amount of parasite diversity passing through the relevant tissue in the mouse, with a higher value indicating a tissue through which a high diversity of parasites passes on its way to other tissues; in this way it provides a measure of connectedness for the tissue. The overall balance of these values across the network can indicate whether the network exhibits asymmetry, which has been linked to directional movement and source-sink dynamics in landscape ecology (Gonzalez et al., 2011).

Betweenness centrality is a related concept, which counts the number of shortest paths which pass through the vertex. The higher this count, the more central to the network the vertex is. In this case, it provides a numeric indication of the likeliest paths via which parasites are traversing the mouse to reach each tissue. Tissues with no betweenness centrality (no shortest paths) may still exhibit high harmonic centrality, if they are overall closely related to multiple other tissues via an intermediary tissue.

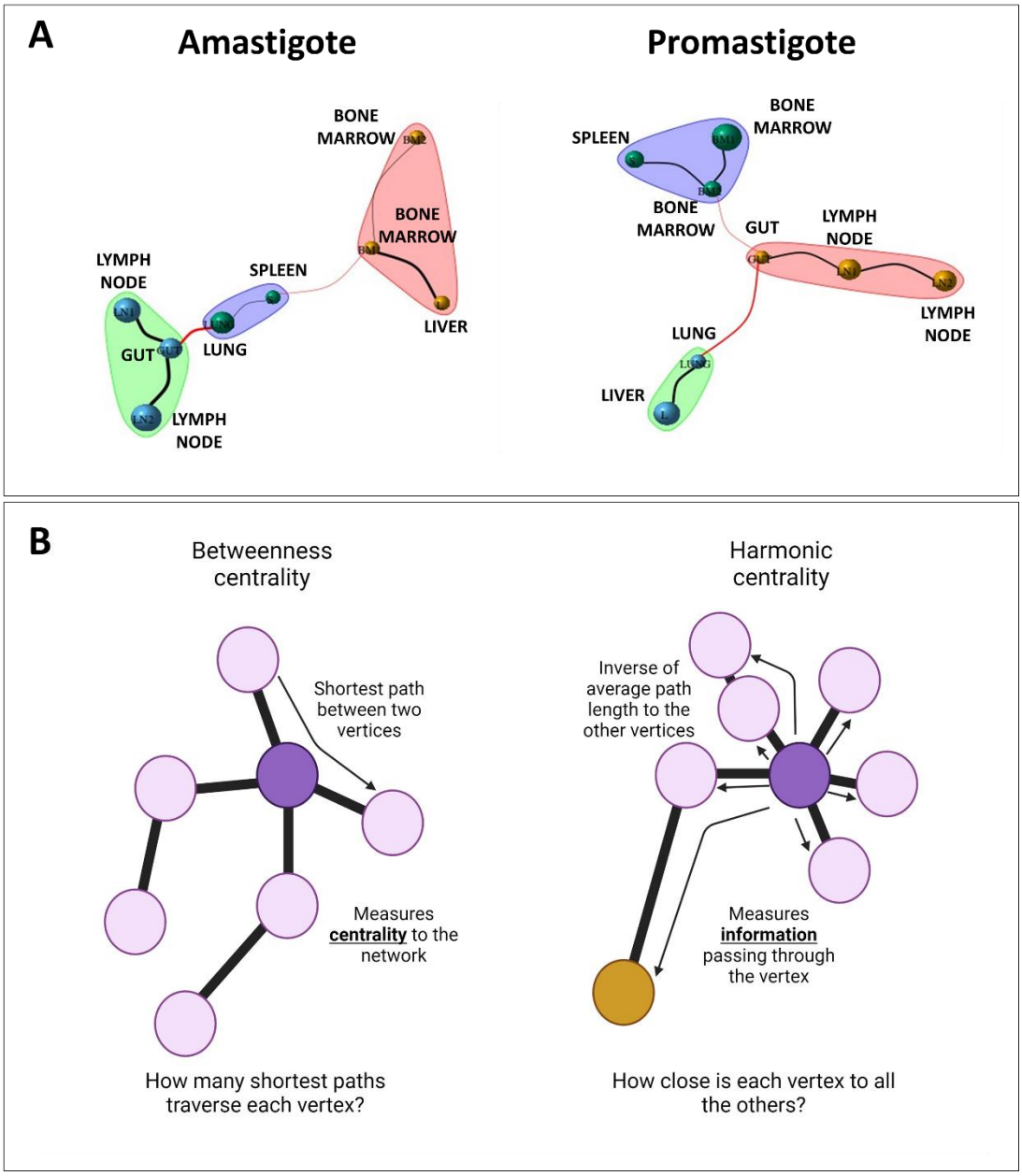


Figure 4.12: Network analysis of tissues in amastigote or promastigote *L. donovani* infected mice to model parasite population relatedness via fitting of a minimum spanning tree (MST). A). Exemplar networks for one female PRO (M4) and one female AMA (M14) mouse at 14 days p.i., calculated using the iGraph R package to generate an MST of tissue relatedness, using genetic distance (GD) as the basis, with Prim's algorithm and Louvain clustering. Cluster colours are arbitrary. Black lines indicate relatedness within clusters whilst red lines indicate relatedness between clusters. Line thickness is proportional to relatedness (1-GD). Tissues are labelled as: GUT, LUNG, L (LIVER), S (SPLEEN), LN1 or 2 (LYMPH NODE), BM1 or 2 (BONE MARROW). B). Illustration of the concepts of harmonic centrality and betweenness centrality. Betweenness centrality measures the number of shortest paths which cross a particular vertex in a network, and provides a metric of how central to the network the vertex is. Harmonic centrality measures the information which

passes through a vertex and provides a metric of how close a vertex is to the other vertices in the network on average. Dark purple indicates the vertex for which a high betweenness centrality or harmonic centrality value applies; light purple indicates other vertices in the network; gold indicates a vertex which, in comparison to the dark purple one, has a low harmonic centrality (it is not close to all other vertices in the network). Panel B created in Biorender.

Figure 4.16 shows the calculated harmonic centrality of each tissue at each time-point in AMA (Figure 4.13A) and PRO (Figure 4.13B) mice. In the AMA model, the harmonic centrality of each tissue is similar at 2 days p.i, indicating similar parasite diversity traversing each tissue; this is evidence of low asymmetry in the network. By day 14, this balance has shifted and the bone marrow shows a higher average centrality compared to other tissues, but with heterogeneity between mice. By day 28, the network is highly asymmetrical, dominated by the bone marrow, liver and spleen compared to other tissues.

In PRO mice, a different trend was observed (Figure 4.13B). At 2 days p.i, the harmonic centrality is again broadly similar across tissues, but at days 14 and 28, there is a higher harmonic centrality in the gut, lymph nodes and lung compared to other tissues. Nevertheless, the difference between upper and lower values is less pronounced than in the AMA networks, indicating lower overall asymmetry in PRO mice.

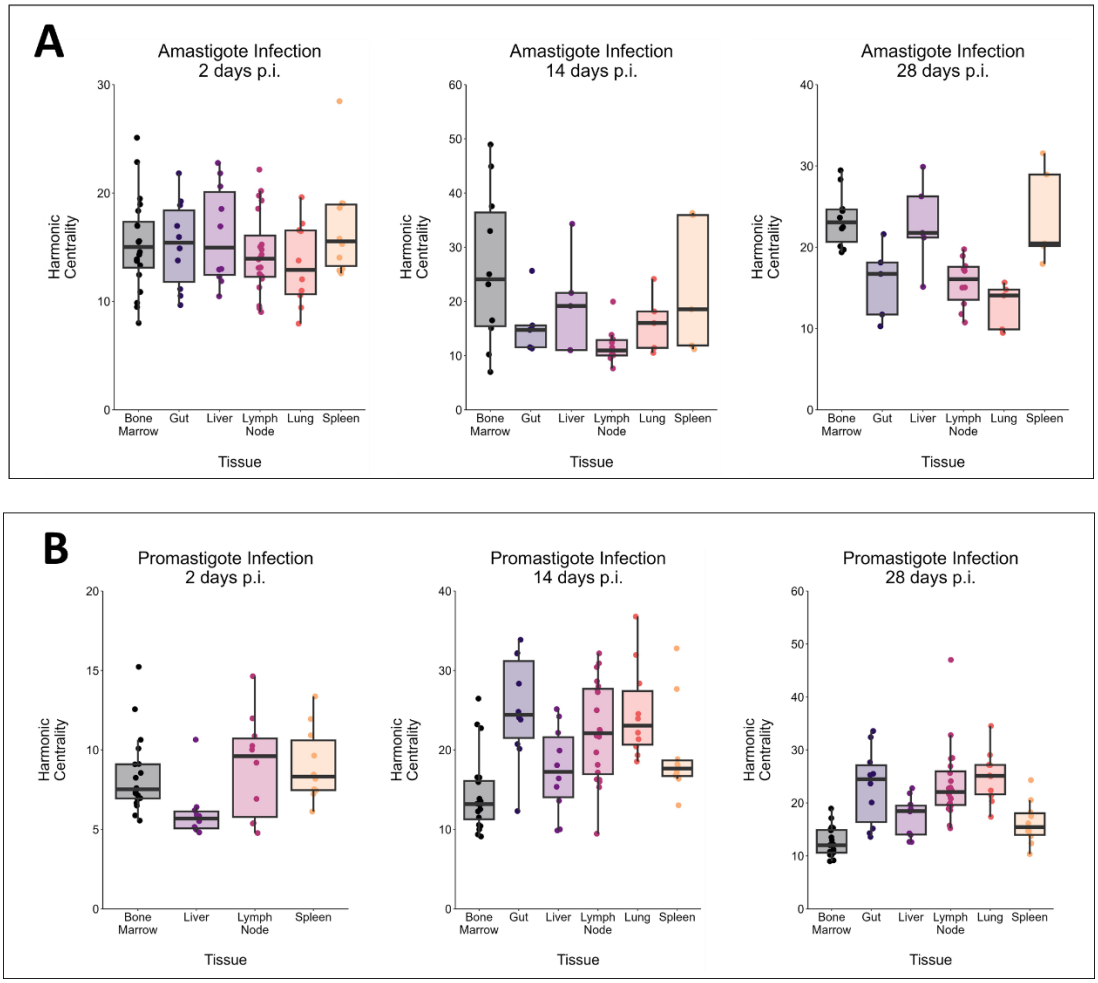


Figure 4.13: Connectedness of parasite populations in different tissues after amastigote or promastigote *L. donovani* infection of male and female C57BL/6J mice at 2, 14 and 28 days p.i. A). Pooled harmonic centrality of each tissue for AMA mouse networks. n=5 males and 5 females at 2 days p.i, and n=5 females at 14 and 28 days p.i. B). Pooled harmonic centrality of each tissue for PRO mouse networks. n=5 males and 5 females at all time-points. Colours indicate tissue and each individual mouse tissue is represented as a single data-point. Boxplots show the data median, 25th and 75th quartiles, with whiskers representing all data-point within 1.5x Interquartile Range of the median; points above or below the whiskers represent outlying data values.

The betweenness centrality distribution of each tissue at each time-point is shown as a violin plot in Figure 4.14A and B for AMA and PRO mice respectively. In AMA mice, the distributions show a wide base at 2 days p.i for lymph node and lung, indicating a bias towards lower centrality, but these distributions narrow in the later time-points, with a greater spread of centrality values. Conversely, the liver becomes more skewed towards lower centrality at 14 days p.i, then at 28 days p.i, it

becomes more widely distributed again. The spleen in AMA mice shows an interesting progression from a wide range of values at 2 days p.i, to a very narrow range at 28 days p.i, with a seemingly very central position in all networks at that time-point. PRO mice show a trajectory in which bone marrow seems to become less likely to be central over time, gradually becoming wider at the base of the distribution.

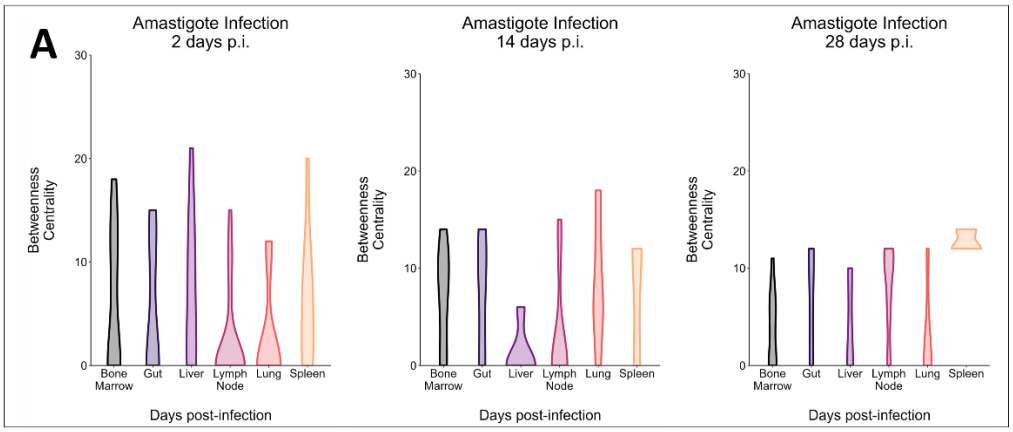
In order to assess the overall contribution of each tissue to the total betweenness centrality, stacked bar plots showing each tissues' betweenness centrality as a proportion of the total betweenness centrality value (total shortest network paths) for each mouse are shown in Figure 4.14C. In AMA mice, there is heterogeneity between mice at 2 days p.i, with some mice showing a high proportion of centrality contributed by the bone marrow, and others showing a more even split between multiple tissues. Two mice show no contribution from bone marrow, with one of these indicating that only the liver contributes any shortest paths. Another mouse shows 75% of the shortest paths are via the spleen, with six out of the total ten mice showing contributions from spleen at this time-point. The gut, lung and lymph nodes appear to have a sporadic contribution, each featuring in fewer than half of mice at this time-point.

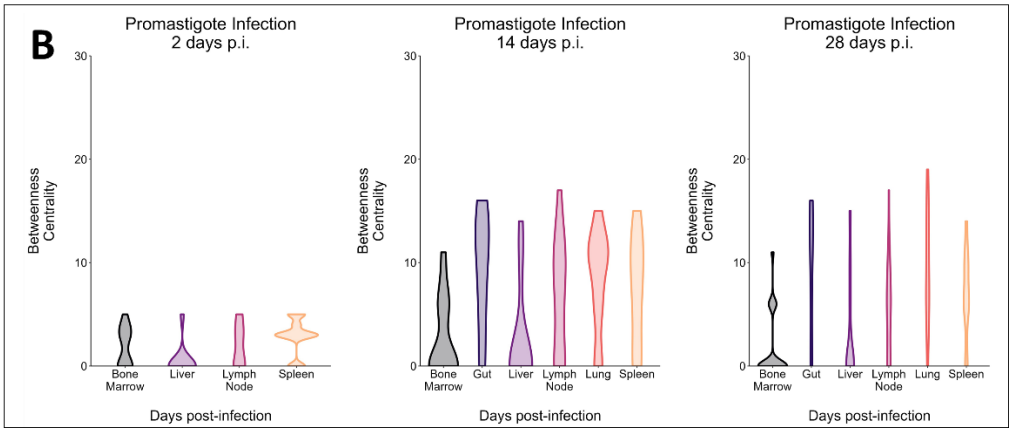
At 14 days p.i, four of five AMA mice show some contribution from the lung, whilst three show contributions of gut and spleen respectively. All mice show bone marrow at relatively even proportions of around 25-40% and a single mouse shows a liver contribution to its shortest paths. At 28 days p.i, the bone marrow contribution, while still present in all mice, has reduced to around 10%. The lymph nodes have increased in contribution, as has the spleen; both contribute shortest paths in all mice, in contrast to the liver, which contributes in only two, and only at around 10% in these mice.

In PRO mice, the bone marrow is more heterogeneous, with seven of ten mice at 2 days p.i indicating it contributes to shortest paths, at widely varying proportions. Spleen contributes in almost all mice at this time-point, whilst liver contributes to only a single mouse, but makes up over 60% of the total shortest paths in this individual. Lymph nodes appear sporadically, with half of mice showing a large

contribution and the others showing none for this tissue. Gut and lung were not included in this network so cannot be accounted for.

At 14 days p.i and 28 days p.i, PRO mice distributions of tissue contributors appeared broadly similar. The liver contributes in two and three mice respectively, whilst the spleen contributes in seven and eight respectively, at relatively even proportions across both time-points. The lymph nodes appear to be important at this time-point, with a large contribution to shortest paths in all mice at both time-points, whilst bone marrow contributions amount to a smaller proportion across a few mice only. Lung and gut both appear in most mice, with eight and seven showing gut contributions and eight and nine showing lung contributions at 14 and 28 days p.i respectively.


A
Amastigote Infection
2 days p.i.
Amastigote Infection
14 days p.i.
Amastigote Infection
28 days p.i.
Betweenness Centrality
Bone Marrow
Gut
Liver
Lymph Node
Lung
Spleen
Days post-infection


B
Promastigote Infection
2 days p.i.
Promastigote Infection
14 days p.i.
Promastigote Infection
28 days p.i.
Betweenness Centrality
Bone Marrow
Gut
Liver
Lymph Node
Lung
Spleen
Days post-infection

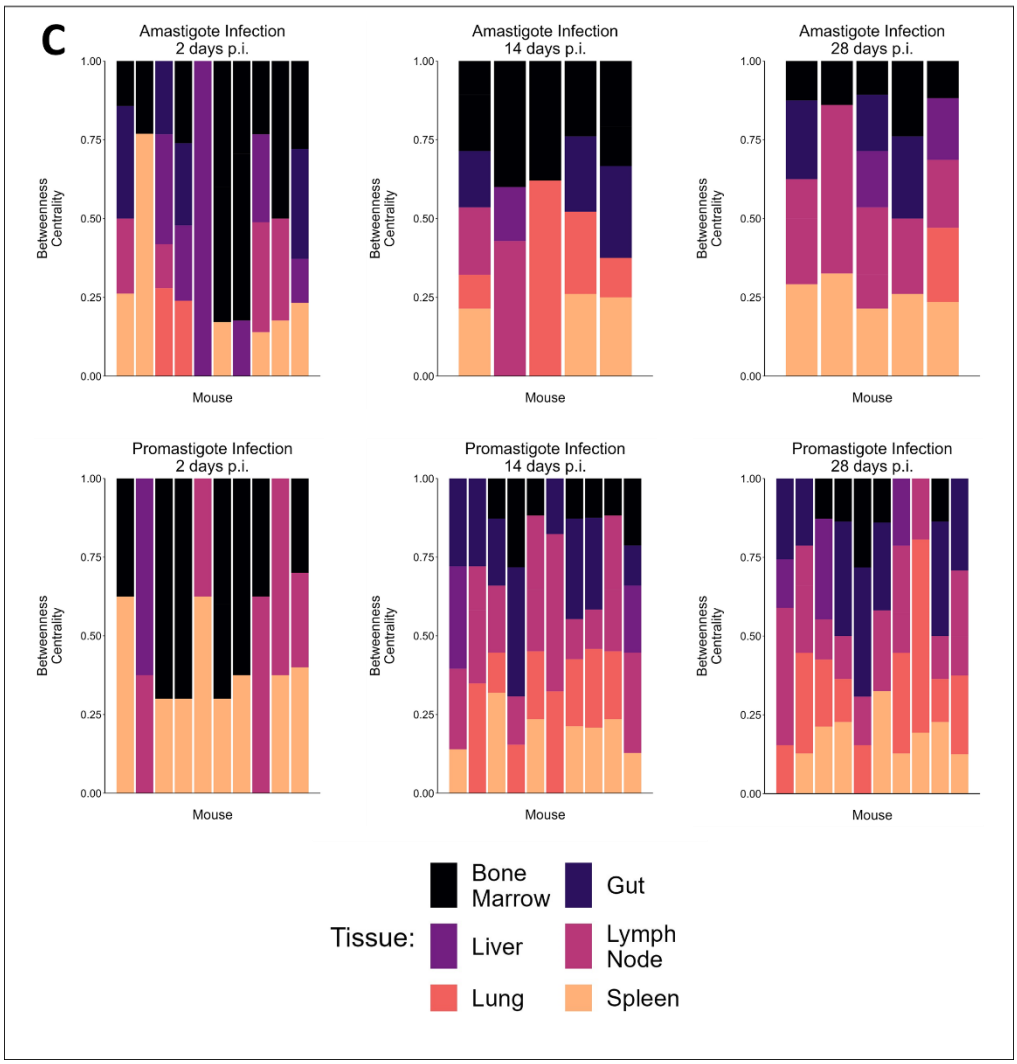

C
Amastigote Infection
2 days p.i.
Amastigote Infection
14 days p.i.
Amastigote Infection
28 days p.i.
Promastigote Infection
2 days p.i.
Promastigote Infection
14 days p.i.
Promastigote Infection
28 days p.i.
Betweenness Centrality
Mouse
Tissue:
Bone Marrow
Gut
Liver
Lymph Node
Lung
Spleen

Figure 4.14: Betweenness centrality of different tissues after amastigote or promastigote *L. donovani* infection of male and female C57BL/6J mice at 2, 14 and 28 days p.i. A). Violin plot visualising betweenness centrality for each tissue for AMA mouse networks. B). Violin plot visualising betweenness centrality for each tissue for PRO mouse networks. C).Stacked bar plots showing the proportion of total betweenness centrality of the network contributed by each tissue type for each AMA or PRO mouse at 2 days, 14 days and 28 days p.i. For AMA mice, at 2 days p.i, n-5 males and 5 females, and at 14 and 28 days p.i, n=5 females for each. For PRO mice, n=5 males and 5 females at all time-points. At 2 days p.i, PRO mice networks included only bone marrow, liver, spleen and a single lymph node. At all other time-points, all tissues are included, with two each of bone marrow and lymph node. Colour indicates tissue.

Overall, these data indicate that following both promastigote and amastigote infection, there is heterogeneity in parasite colonisation routes early on. These parasite trajectories become more homogeneous across different mice in both infections but this appears to occur earlier in promastigote-infected mice. Both infection models show a trend towards increased asymmetry over time, but with different dominant tissues. There is evidence for importance of the bone marrow in early dissemination for both infection models, but bone marrow appears to become less important in promastigote-infected mice over time. The gut and lung are likely to be important in providing dissemination routes between tissues in promastigote, but not amastigote, infected mice.

4.2.7 Assembly processes shaping parasite populations *in vivo* differ between promastigote and amastigote infection models

To examine the dominant assembly processes contributing to the population structure in each tissue, a null model-based analytical pipeline assessing beta diversity of pairwise tissues within each mouse was used. The protocol followed for this analysis is shown in Figure 4.15A.

Since barcode abundance in our pooled input library differed, we first created a phylogeny to use for testing for evidence of priority effects or historical

contingency, using the assumption that highly abundant barcodes would be the first to reach each tissue.

To generate the pseudo –phylogeny used for the binning and clustering process, a Canberra distance algorithm was applied to the input sample sequence read counts for each individual experiment, with the resulting pseudo-phylogenetic trees shown in Figure 4.15B. The tree compositions appear different across three separate experiments, indicating that the barcode insertions were not influencing this relatedness or the growth of the parasites *in vitro* prior to infection into mice.

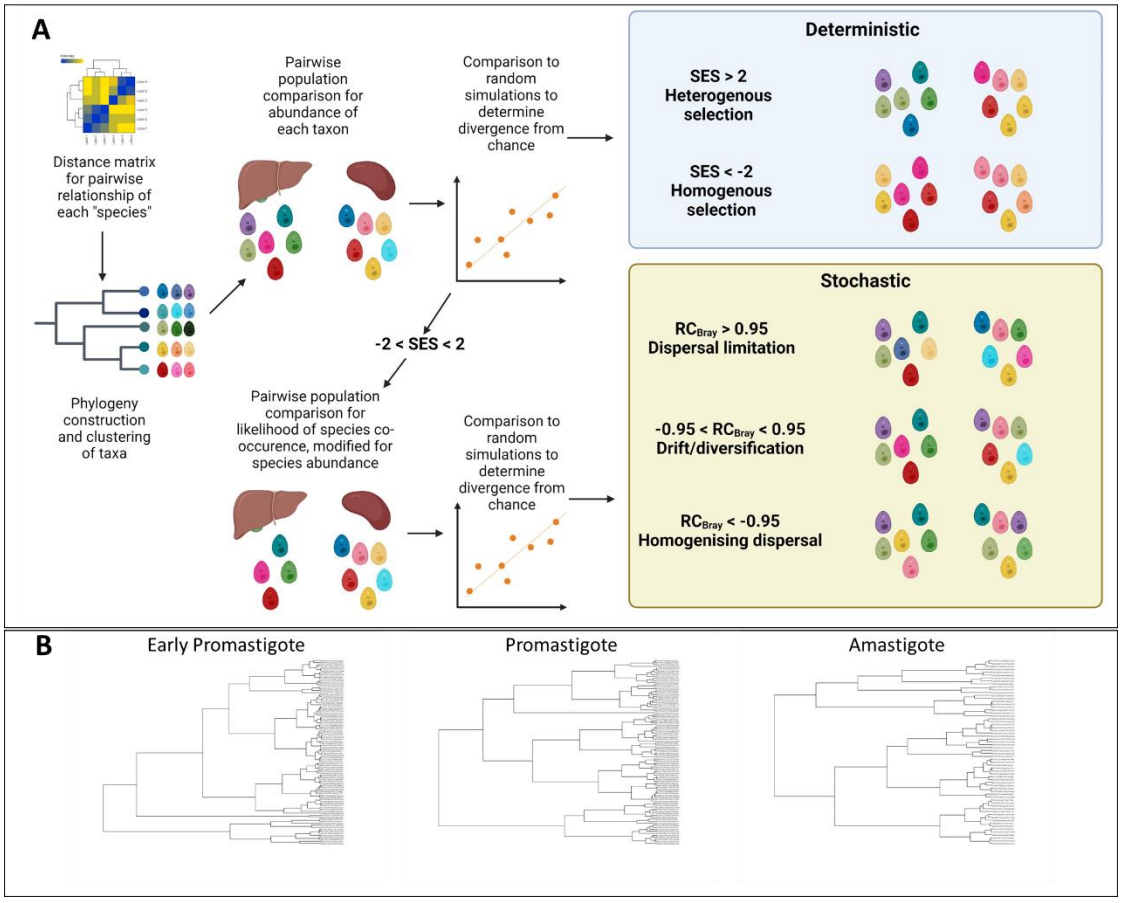


Figure 4.15: Computational analysis of assembly processes in microbial community structuring. A). Illustration of the computational methodology for determining the dominant assembly processes contributing to the structure of a microbial community. For details, see Methods. B). The computed pseudo-phylogenies for each experiment, using the INPUT barcode counts to calculate the distances via a Canberra algorithm. Pseudo-phylogenetic trees for the two day (left) and 14 and 28 day (centre) PRO infections and the AMA infections (right), calculated using the diversity of barcodes in the mean of INPUT sample sequence reads. Panel A created in Biorender.

Figure 4.16 and 4.17 show the resulting values of bNTI (Figure 4.16A and 4.17A) and RC_{BRAY} (Figure 4.16B and 4.17B) for each tissue to tissue comparison in AMA and PRO mice respectively. The relative proportions of pairwise interactions showing a dominance of each assembly process, organised by tissue, is shown in Figure 4.18.

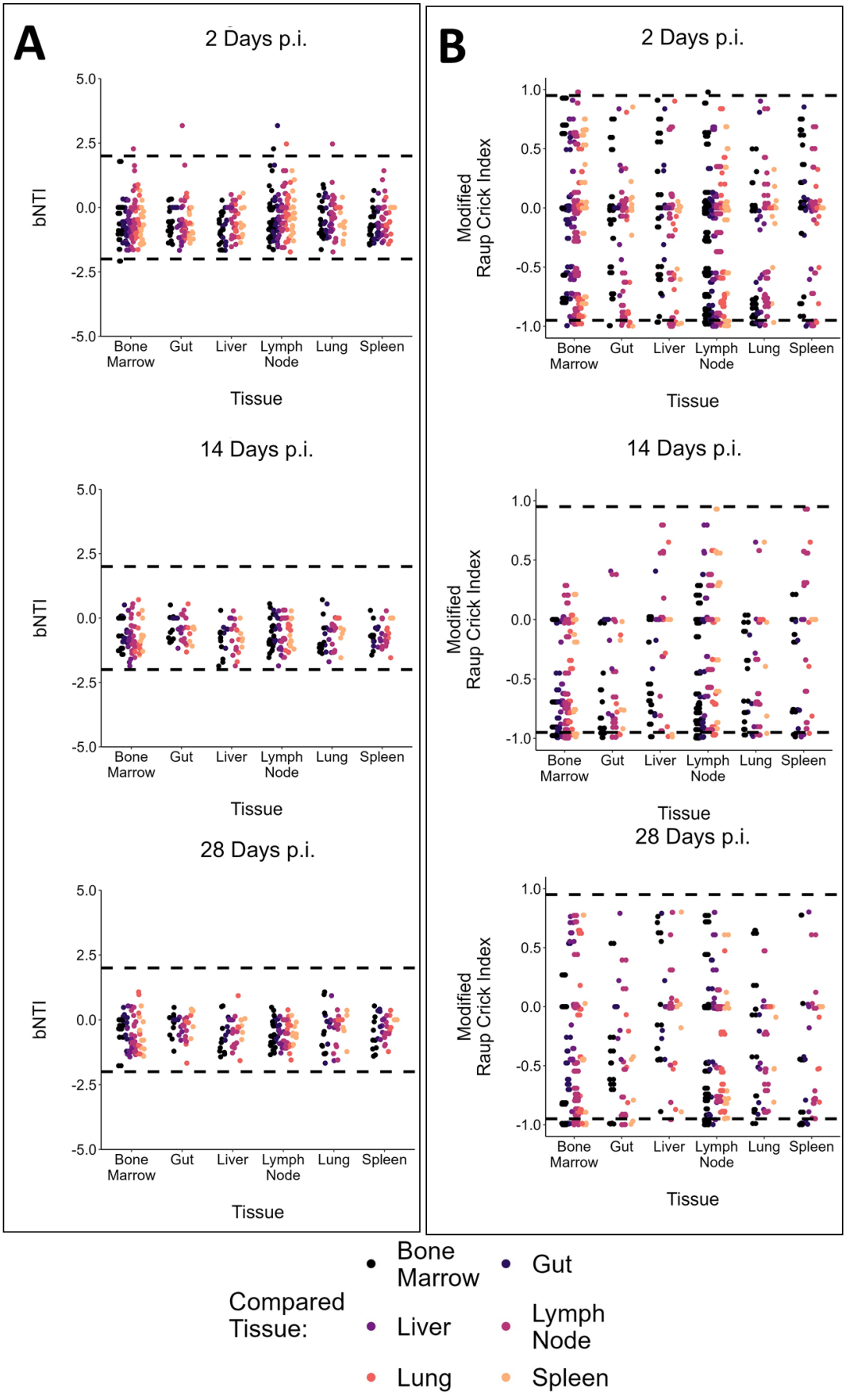


Figure 4.16: Beta diversity of parasite populations in different tissues after amastigote *L. donovani* infection of male and female C57BL/6J mice at 2, 14 and 28 days p.i. A). bNTI

values for pairwise comparisons of each tissue parasite population in amastigote-infected (AMA) mice. Dotted lines indicate bNTI = 2 or -2, the cut off for stochastic versus deterministic processes. B). Bray-Curtis modified Raup-Crick Index for pairwise comparisons of the interactions falling within the interval -2<bNTI<2. Dotted lines indicate RC_{BRAY} = 0.95 or -0.95, the cut off for dispersal limitation or homogenising dispersal respectively. Data points represent a single pairwise comparison of two tissues within the same mouse. Colours indicate the tissue compared with the one on the x-axis. As such, each data point is displayed twice in the plot. n=5 males and 5 females at 2 days p.i, and 5 females at 14 and 28 days p.i.

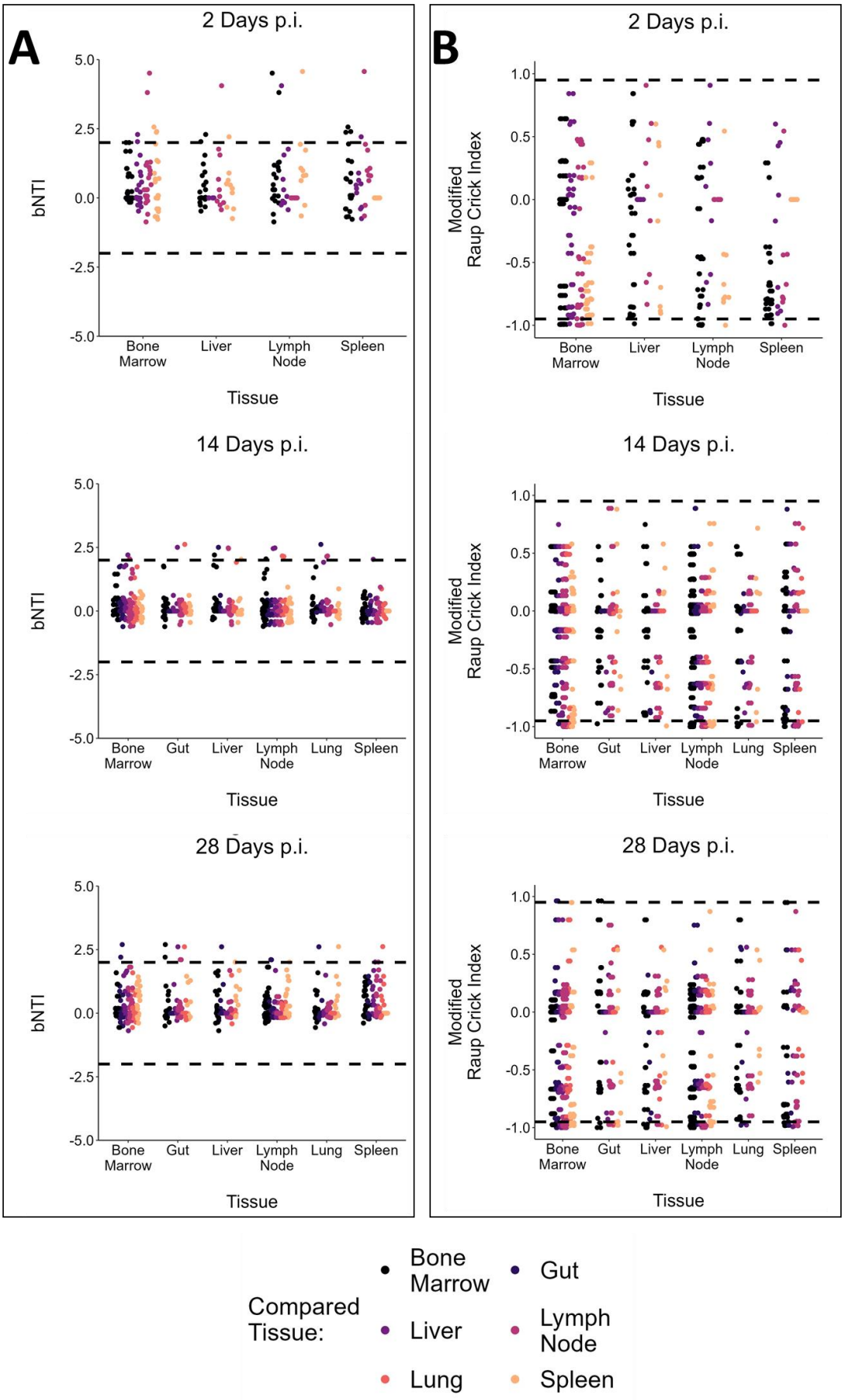


Figure 4.17: Beta diversity of parasite populations in different tissues after promastigote *L. donovani* infection of male and female C57BL/6J mice at 2, 14 and 28 days p.i. A). bNTI

values for pairwise comparisons of each tissue parasite population in promastigote-infected (PRO) mice. Dotted lines indicate bNTI = 2 or -2, the cut off for stochastic versus deterministic processes. B). Bray-Curtis modified Raup-Crick Index for pairwise comparisons of the interactions falling within the interval -2<bNTI<2. Dotted lines indicate RC_{BRAY} = 0.95 or -0.95, the cut off for dispersal limitation or homogenising dispersal respectively. Data points represent a single pairwise comparison of two tissues within the same mouse. Colours indicate the tissue compared with the one on the x-axis. As such, each data point is displayed twice in the plot. Due to scaling for visualisation purposes, a single value of bNTI = -99 is not shown; this was for the M15 GUT:SPLEEN interaction. n=5 males and 5 females at all time-points. A second lymph node, gut and lung were not taken for the 2 days p.i mice.

Drift was identified as the dominant force shaping the assembly of all tissues in both infection models, indicating that most of the diversity observed is due to random parasite movement and ongoing bottlenecks, rather than active selection of individuals. Homogenising dispersal was the second most dominant process in almost all tissues at all time-points, with the exception of the early PRO tissues. The contribution of homogenising dispersal increased over time in the PRO mice in almost all tissues, except the lung, whilst in AMA mice, it increased in the bone marrow, gut and spleen. It decreased in influence in the liver and lung of AMA mice and remained consistent in the lymph nodes over the 28 day infection. Overall, there was very little contribution of dispersal limitation in any tissues, indicating that barriers (immunological, physical or otherwise) were not preventing parasite dispersal between tissues. A very small number of comparisons in the early amastigote infection and the late promastigote infection showed evidence of limited dispersal.

The apparent contributions made by heterogeneous selection were limited to a small number of individual mice and demonstrated no clear pattern across infection time or tissue. Similarly, homogenous selection was observed in a total of two tissue to tissue interactions across all samples compared. It is therefore likely that this is due to statistical noise and not a true indication of a role for deterministic processes in this dataset.

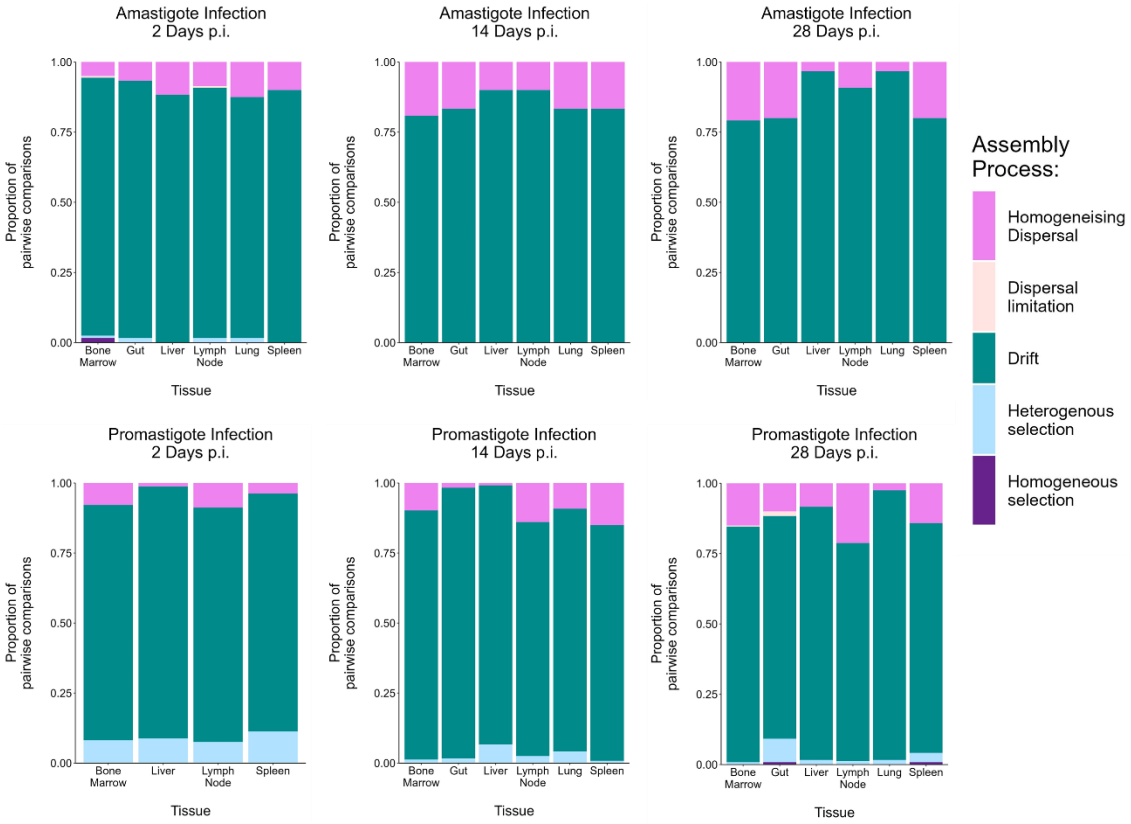


Figure 4.18: Proportion of interactions dominated by each assembly process, grouped by tissue, after amastigote or promastigote *L. donovani* infection of male and female C57BL/6J mice. A). Stacked bar plots illustrating proportion of comparisons, split by tissue, showing evidence of dominance by each method of assembly in the amastigote-infected (AMA) mice at 2 days, 14 days and 28 days p.i. n=5 males and 5 females at 2 days p.i, and 5 females at each of 14 and 28 days p.i. B). Stacked bar plots illustrating proportion of comparisons, split by tissue, showing evidence of dominance by each method of assembly in the promastigote-infected (PRO) mice at 2 days, 14 days and 28 days p.i Colour key is identical for both panels and indicates the dominant assembly process for the proportion of pairwise comparisons assessed which included the tissue on the x-axis; due to the comparisons being pairwise, each individual data value will be present twice - once for each tissue in the pair. n=5 males and 5 females for all time-points.

Overall, this analysis provides evidence for a long-term bottlenecking process in all tissues assessed over a 28 day infection, after infection with either life cycle stage. Additionally, it indicates that drift was the dominant assembly process, with homogenising dispersal increasing over time in most tissues after a promastigote infection, whilst it increases in the bone marrow, spleen and gut after amastigote

infection. There appear to be no strong barriers blocking dissemination between tissues.

4.3 Discussion:

The work presented here has integrated a genetic barcoding approach with high-throughput Barseq and ecological analysis methods to understand how *L. donovani* parasites colonise and disseminate in an *in vivo* mouse model of VL, using two different life cycle stages for the initial infection. This has revealed important differences in the patterns of dissemination between tissues after intravenous injection with amastigote or promastigote stage parasites, two widely used methods for experimental VL infections. Notably, tissue-specific colonisation appears to markedly differ between the two, as do routes of dissemination between tissues over a 4 week infection. Additionally, parasite diversity exhibits a strikingly different relationship to bottleneck size, dependent on life stage at infection.

Bottleneck assessment via FP quantification revealed the spatial and temporal bottlenecks within both infection models. The liver appears to be a key early colonisation site in promastigote and amastigote infections, with the highest average FP of all tissues in both cases. However, the ability of parasites to colonise the spleen was similarly high in AMA mice, whilst the spleen FP was far lower than the liver FP in PRO mice. The lung and gut showed a distinct divergence between the two infection models, with PRO mice having a relatively high FP, whilst very little colonisation occurred at any time-point in AMA mice in these tissues. However, parasite burden was not measured in these tissues – it is plausible that a small amastigote population expanded to constitute a burden of comparative size to that in the promastigote infection.

We found that in PRO mice, colonisation appeared to increase over the timecourse of infection in most tissues, which indicates that an initial strong bottleneck may prevent promastigote establishment, but this bottleneck then relaxes over time. In AMA mice, a similar trend was seen. This increase in permissiveness to colonisation could be a result of actions of the innate immune system imposing an initial toxic

barrier, which then dissipates over time, or could alternatively be due to an early-acting resource limit, such as host cell availability.

A striking finding was that, when assessing alpha diversity using SI, the asymptotic relationship between FP and SI was inverted when comparing PRO and AMA mice. In PRO mice, the asymptotic curve followed the anticipated trajectory (Jiang et al., 2024), with a higher FP resulting in increased diversity, but with a reduced impact for each new individual added to the population. The curve levelled off at a maximum diversity which was related to the logarithm of FP and decreased over time such that the maximum diversity was lower at 28 days than at 2 days p.i. This intuitively makes sense, since earlier founders would be expected to have less competition compared to late-comers. This along with the shorter time in which to replicate would therefore lead to fewer individuals of the later colonising lines.

In contrast, in AMA mice, the impact of new individuals was reduced where the FP was already high. There was a lower limit for diversity, which also increased over time such that a higher FP later in infection was not able to increase diversity as much as the founders at an earlier stage. This finding points towards an unbalanced population, with some barcodes making up a larger proportion of the total population and others being present in very low numbers. One explanation for this effect in amastigote – but not promastigote – infections is that some parasites may evolve advantageous mutations which are selected for whilst in passage in B6.*Rag2*$^{-/-}$ CD45.1 mice; these could enable more rapid replication and allow some lines to outcompete others. Promastigote parasites would be less likely to develop these advantageous mutations since there is no selective pressure for them in *in vitro* culture.

Other differences could arise during passage *in vivo* (e.g. epigenetic alterations which alter how genes are expressed or affect the translation of specific proteins), which could also alter parasite phenotype. Indeed, Dirkx et al. (2022) noted that amastigotes can rapidly evolve via transcriptional regulation after entry into host cells in the bone marrow. Another explanation could have been a role for priority effects or historical contingency in amastigote infection – if earlier parasites modified the environment such that later parasites are either advantaged or

disadvantaged, this could result in the uneven growth observed. However, our assembly process analysis indicates that this is unlikely to be the case here.

A further possible explanation is that the *ex vivo* amastigotes are primed to enter into a quiescent state, having previously entered this dormant state while in passage, or indeed they may remain quiescent during extraction from the B6.*Rag2*$^{-/-}$ CD45.1 spleen and subsequent infection process. Competition for resources will be higher in populations with higher FPs, which could induce more parasites to enter this dormancy state, which requires less metabolic input to maintain.

Previous studies with *L. donovani* have found evidence that parasites infecting bone marrow HSCs can be induced to enter quiescence and that this dormancy remains even when parasites are removed from the original environment (Dirkx et al., 2022; Dirkx et al., 2024). Likewise, studies of *L. mexicana* have indicated that quiescent parasites are present in skin granulomas (Kloehn et al., 2015; Kloehn et al., 2021) and that quiescence can be induced *in vitro* in *L. braziliensis* clinical isolates (Jara et al., 2023). The tissue specificity observed here is interesting in that only the bone marrow and liver show a significant decrease in diversity over time. This implies that either something specific about those microenvironments or the cell types infected in them may induce quiescence, or that quiescent parasites are trafficked to those tissues preferentially. Understanding which of these applies could give insight into parasite dissemination, as it is not known whether quiescent parasites are capable of tissue-specific trafficking, as well as providing further data on the mechanisms regulating quiescence entry.

Overall, this analysis points to tissue-specific differences in parasite seeding, dependent on life stage at infection, which persist up to 4 weeks, despite the fact that all parasites will either die or differentiate to amastigote form within 2 days of initial infection. Therefore, these data demonstrate that early events are important in the determination of later tissue dissemination patterns in VL.

Prior literature shows amastigote infections tend to produce higher parasite burdens (Ong et al, 2020), but most direct comparisons between life stages have been performed *in vitro*. Abu-Dayyeh et al (2010) reported that *L. mexicana* amastigotes and promastigotes induced different signalling pathways in macrophages, whilst other work has found differences in phagocytosis between life

stages, particularly in terms of the host cell receptors utilised for entry (Argueta-Donohué et al., 2015; Ueno and Wilson, 2012). Complement fixation, susceptibility to complement-mediated lysis and opsonisation-induced phagocytosis are also known to differ between promastigotes and amastigotes (Franke et al., 1985; Mosser and Edelson, 1984; Tirado et al., 2021). Life stage differences in surface expression markers involved in both phagocytosis and survival within host cells, such as LPG and GP63, are also well described (Desjardins and Descoteaux, 1997; Desjardins and Descoteaux, 1998; Lodge et al., 2006; Scianimanico et al., 1999; Verma et al., 2017). A recent study profiled the differential gene expression of bone marrow macrophages after infection with promastigote and amastigote *L. donovani* shortly after infection (Chaparro et al., 2022); their findings point towards divergent reprogramming of host cells dependent on parasite life stage. Host cell signalling in key pathways involved in the induction of the adaptive immune response and for responding to effectors such as Type 1 IFNs was affected differently after amastigote or promastigote infection, pointing towards a reprogramming mechanism for downstream effects triggered by the initial events of infection. Another study has also noted differences in the host cell translatome after infection with both life stages, with many alterations in pathways linked to increased parasite survival (Chaparro et al., 2020). However, few direct *in vivo* comparisons of the two infection models have been reported, meaning knowledge of how the innate response differs, and how this could alter the progression of disease in the longer term, is unknown. Further single cell transcriptomic analysis at the tissue level might assist in understanding the contribution of these early responses to later dissemination trajectories *in vivo*.

One key early event in leishmaniasis is the recruitment of phagocytic host cells, which is in part mediated by the complement system. The potential involvement of the complement system in early dissemination in PRO mice is discussed in Chapter 5. Tissue staining or flow cytometry for macrophage, monocyte, neutrophil and parasite markers in the liver, spleen and bone marrow would be valuable to understand whether there are differences in recruitment and parasite uptake by these cells early after amastigote or promastigote infection. Analysis of phagocyte-recruiting chemokines and complement components, such as CXCL-family attractants and C3a and C5a (Metzemaekers et al., 2020) would also be of interest

in establishing how early reprogramming of specific cell types might influence later tissue-specific patterns. Antibody-mediated depletion of different host cell types in a mouse model could provide greater clarity on their specific role, but has been problematic to establish thus far (Carlsen et al., 2015).

The quantification of GD for tissue to tissue comparisons in each infection model established how closely related different parasite populations were within each individual mouse. Again, key differences were seen between the two life stage infections. After promastigote infection, the lymph node, spleen and bone marrow all became more closely related to the liver over time, whilst after amastigote infection, the bone marrow became more closely related to the spleen and liver. No other tissues showed increased relatedness over the 4 week experiments.

These data provide evidence of possible seeding of parasites from the promastigote-infected liver to other tissues over the 28 day infection. In amastigote infection, seeding appears to be predominantly from the spleen to the liver and bone marrow. One possibility is that parasites within circulating inflammatory monocytes traffic between tissues as a by-product of their CCR2-mediated recruitment. It has been shown that *L. donovani* parasites can infect monocytes and that CCR2+ monocyte depletion results in reduced lymph node, spleen and liver parasite burdens after amastigote infection, but burdens in other tissues were not measured in this study (Terrazas et al., 2017). Parasite survival within monocytes has also been shown to be dependent on both tissue origin and exposure to signals such as IFNγ and *L. donovani* antigens (Glennie et al., 2017; Romano et al., 2017; Romano et al., 2021; Terrazas et al., 2017). Perhaps early reprogramming by parasites leads to differences in monocyte activation status and receptor expression, with downstream effects on migration and parasite killing activity, creating the specific patterns of dissemination observed here.

Parasites in the liver are predominantly found within granulomas after the early stages of infection; however, to our knowledge no studies have assessed specific differences in granuloma formation kinetics or composition between promastigote and amastigote-instigated infections. Therefore, it is possible that early events stemming from the innate immune response to each life stage can trigger a ‘looser’ granuloma, which is more permissive to parasite escape, in the case of

promastigote infection. Establishing how granulomas differ between the two models could enlighten on why parasite migration from the liver seems to occur more readily after promastigote infection. Using an anti-CCR2 antibody to block trafficking, following infection with the barcoded library, could provide evidence for (or refute) monocyte trafficking as a dissemination method from the liver, which in the long term may enable the development of potential immunotherapies to block dissemination.

The gut, lung and liver populations all appeared highly related following promastigote infection. The gut has previously been hypothesised to contribute to pathology and detrimental outcomes in end stage disease, using a hamster model of VL (Lewis et al, 2020); in their work, they showed a close linkage of gut with the liver. The proposed mechanism involved 'leakage' of commensal organisms into the liver, which then became pathogenic, leading to detrimental health outcomes. However, it is possible this observation was a result of wide-spread immune dysregulation in end stage disease. Whilst data in mice is lacking, the presence of parasites in the gut and lung of HIV-positive VL patients also points towards a possible role for this organ – HIV-positive patients are more likely to die or suffer relapse post-treatment due to secondary infections (Franssen et al., 2021; Zijlstra, 2014). The possible relevance of the gut-lung-liver relationship after promastigote infection may be important to study in more depth.

We additionally investigated the relatedness of different liver lobes to each other and to other tissues, to validate our experimental approach of using only a small segment of this tissue. The rodent liver is noted to exhibit heterogeneity between individuals in terms of blood supply to each lobe, and different lobes may be more closely linked to the supply of other tissues (Malarkey et al., 2005; Sänger et al., 2015). Our findings were reassuring in that we found similar FPs and relatively low GDs between lobes. We also noted few differences in GD to other tissues. However, understanding how the spatial distribution of parasites varies within the liver could be useful in more fully understanding parasite-host interactions in this complex organ. For example, examining how parasite diversity relates to blood vessel locations, and how it links to tissue gradients in mineral distribution, specific cell types or in conjunction with other disease states (Malarkey et al., 2005) could

reveal important information about within tissue dissemination that is as yet unknown.

Network theory has been utilised to understand the microbiome in a few studies, but is more commonly used for investigating aspects of environmental microbiology, including disease transmission between individuals. Here, it was applied to understand how parasite populations in different tissues relate to each other and how the overall network structure contributes to the stability of the parasite community. The asymmetry of the networks increased over time in both promastigote and amastigote infection systems. This is an indication that specific tissues which act as 'sinks' are highly dependent on others ('sources') in the network to remain linked – if these key 'source' tissues are removed from the network, parasites may be blocked from disseminating to other tissues. Finding these key tissues and connections using an iterative removal of vertices and edges (Gonzalez et al., 2011) could establish the key tissues or routes to target with therapies for destabilising the dissemination network; another adaptation of the analysis could establish evidence for multiple versus single routes between key tissues, which may assist in developing further hypotheses on parasite trafficking mechanisms.

AMA mouse networks indicated the importance of the spleen, bone marrow and lymph nodes as tissues of high betweenness, and it was noted that a clear division was present between the main network of bone marrow, spleen and liver, and a sub-network, with far lower diversity, in the lymph nodes, gut and lung. The lymph node appears to be the key linker, imposing a restrictive bottleneck to the colonisation of the second network. In PRO networks, by contrast, the gut and lung appear to play a key role, with a consistent betweenness contribution across most mice at later time-points. The liver is notably absent as a highly connective tissue, which may indicate that populations from the liver use the lung and gut as 'corridors' to disseminate parasites to other tissues in the network. It is unclear why these tissues would act as a conduit. It is possible this connection is due to a transient population of infected host cells flowing through in the bloodstream, rather than becoming resident in the tissue – parasite movement may be greater in PRO mice, compared to AMA mice, and measuring blood parasite diversity would

provide further clarity on this. Another possibility is that incoming infected monocytes from the liver become resident and parasites can then infect other monocytes, which subsequently leave and travel onwards to other tissues. Why this apparently only occurs in a promastigote infection is unclear, but may relate to the specific phenotype and tissue origin of the host cells involved. It could also relate to differences in parasite activity (i.e. quiescence), which may trigger different cellular responses, or reduce the propensity to infect new host cells.

The fact that the barcode library consisted of a relatively small number of barcodes meant that in our study there was reduced resolution of more permissive bottlenecks (Abel et al., 2015b), but as the FPs observed were relatively small, this was acceptable for our experimental models. It would be important to have a larger library in infection models where the bottlenecks are expected to be more permissive, to discriminate between higher FP values accurately. One key advantage of the smaller barcode library was the possibility to apply community ecology analytic approaches, which are more accurate for groups of up to approximately 100 species (Gotelli and Ulrich, 2012; Zhou and Ning, 2017). Thus, it was possible to treat each barcode as a 'species' and use these approaches to estimate the relative contributions of different assembly processes to tissue parasite populations. Since the barcoded lines were neutral and the population sizes were small, stochastic processes were anticipated to be more important for population structure, with a particular emphasis on drift. The analysis additionally allowed determination of whether priority effects and historical contingency were involved as a mechanism for amastigote diversity reduction.

The findings were mostly as expected, with a high proportion of the community being assembled via drift. Drift is a strong dominating force in populations undergoing severe bottlenecks (Abel et al., 2015b; Weaver et al., 2021); the indication that drift was a major process at all time-points implies that continual bottlenecks are being applied to all tissue populations, which would be dominated by different aspects of the innate and adaptive immune response as time progresses. Although mutations in the library other than the barcode insertions cannot be ruled out, as the lines were not each fully genome sequenced, there did not appear to be a great impact for selective processes overall, which provides

supporting evidence for neutrality of the barcoded library. Homogeneous selection was almost entirely absent, whilst heterogeneous selection appeared to play a role only in a small subset of individual mice – this indicates that priority effects were not strong enough to overcome drift in most cases and were not the underlying reason for the amastigote infection SI-FP relationship observation.

Homogenising dispersal was the second most prominent process, and increased over time in some key tissues including the lymph nodes in PRO mice, and the bone marrow and spleen in AMA mice. This indicates that there are low barriers to parasite movement in these tissues. It is important to note that dispersal effects must be very strong to overcome the impact of drift, especially for small populations, so the absence of a dominant contribution does not indicate that no homogenising dispersal occurred in a specific tissue – only that it was weaker than the role played by drift. Transcriptomics, proteomics and metabolomics approaches could be used to further study these processes – correlating tissue parameters such as the presence of key cell types, metabolites, cytokines, chemokines or factors such as pH could determine which are the likely key contributing factors for the populations of different tissues (Stegen et al., 2013). This approach could also be beneficial to apply to other studies; using a similar method to evaluate assembly processes for libraries of tagged mutants, such as that used in Baker et al (2021), could provide a greater understanding of the underlying roles of different genes. By generating a pseudo-phylogeny based on the similarity of hypothesised roles of different genes, improved insight into the importance of specific pathways for parasite colonisation and fitness is possible.

Based on the findings described, it appears that bottlenecks are both an ongoing process over a 4 week *L. donovani* infection, and that tissue-specific bottlenecks differ depending on parasite stage at infection. Further, they also support the hypothesis that early events, within the first two days of infection, can affect bottlenecks and wider dissemination patterns between tissues over the longer term. Figure 4.19 illustrates some examples of ongoing bottlenecks *in vivo* and how initial events could contribute to the differences between AMA and PRO mouse tissue bottlenecks.

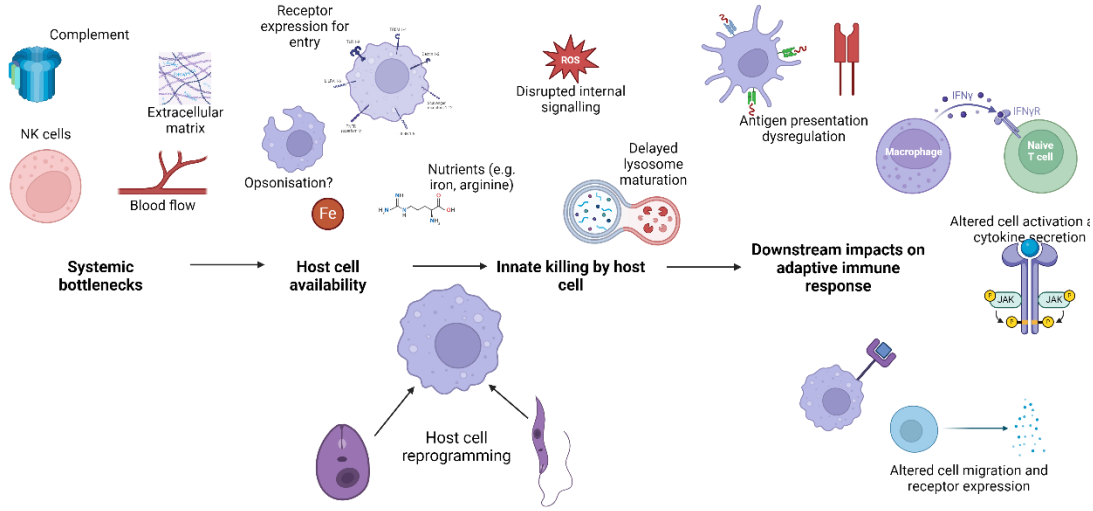


Figure 4.19: Initial bottlenecks and their possible long term impacts in amastigote and promastigote *L. donovani* infection. Amastiote and promastigote parasites face multiple obstacles to colonising tissues, including systemic bottlenecks and limits on host cell availability. This might be tissue specific, in that suitable receptors may differ between tissues. Parasites may also reprogramme host cells such that they can increase their survival, but these changes can alter cell phenotype such that there are downstream effects on the adaptive immune response, which may explain why early events are important in determining longer term dissemination patterns *in vivo*. Figure created in Biorender.

Figure 4.20 shows the hypothesised dissemination routes of AMA and PRO parasites, over the course of infection, and some possible mechanisms underlying them. Briefly, our hypothesis is that in AMA mice (Figure 4.20A), initial colonisation occurs predominately in the liver and spleen, with small numbers of parasites seeded elsewhere. This results in recruitment of inflammatory monocytes from the bone marrow, which initiate granuloma formation. Parasites appear to traffic from the spleen to the bone marrow, where they can become quiescent within HSCs (Dirkx et al., 2022; Dirkx et al., 2024; Karagiannis et al., 2023); amastigotes are more prone to this, having previously been (or remained) in a quiescent state. HSCs have been noted to be activated and induced to migrate in response to liver damage (Baldridge et al., 2010; Cotterell et al., 2000; Dalakas et al., 2005), which may occur here, explaining the apparent increase in population skew (decreased SI) over time in these two AMA tissues, but not in others assessed. Alternatively, parasites may more frequently enter quiescence-inducing host cells in these two tissues.

In PRO mice (Figure 4.20B), the liver is again a key early colonisation site, but one hypothesis for their dissimilarity to AMA mice is that differences in granuloma structure and kinetics mean parasites can escape in infected monocytes and are trafficked to other tissues in response to CCR2 and other chemokine signals.

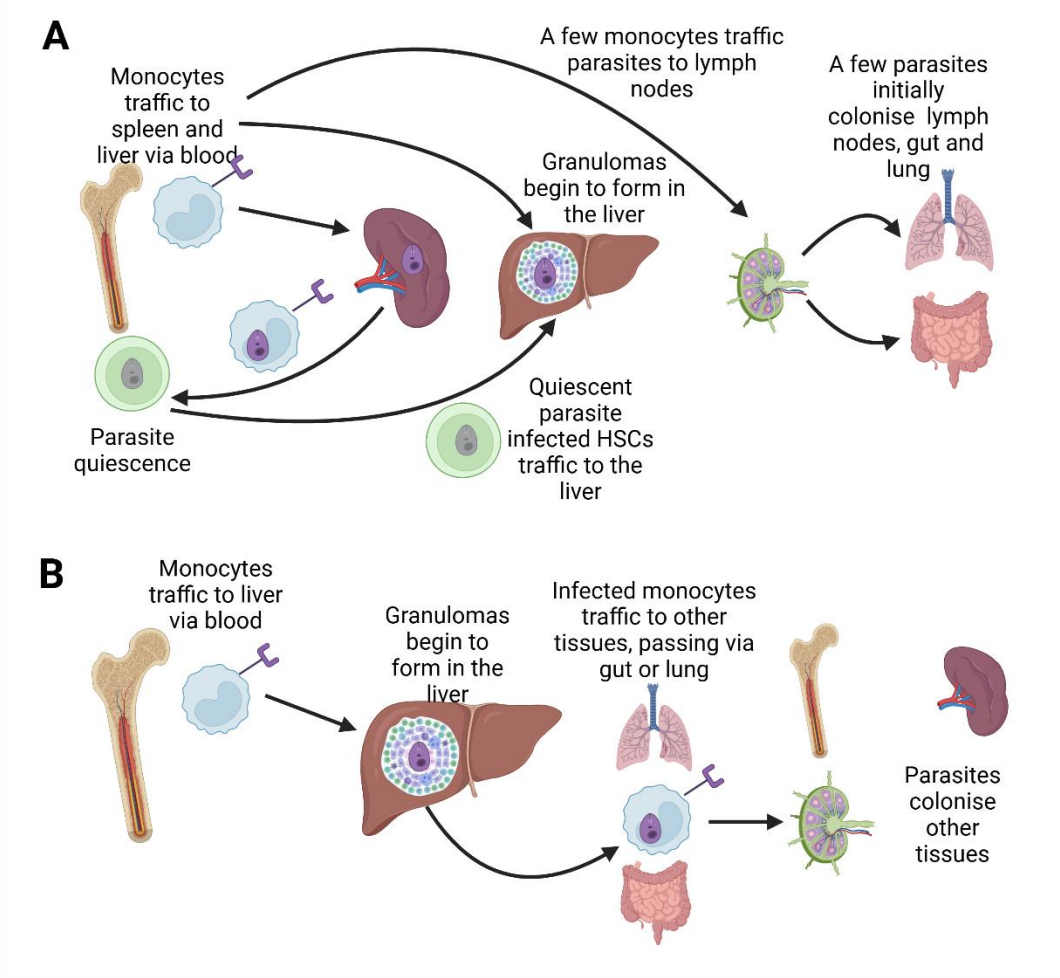


Figure 4.20: Proposed mechanisms of *L. donovani* parasite dissemination patterns in C57BL/6J mice. A). In AMA mice, uninfected monocytes are trafficked to the higly infected spleen, where they become infected and subsequently traffic back to the bone marrow. In the bone marrow, parasites may be more liable to enter quiescence within HSCs. A key question is whether quiescent parasites can be transported to the liver in HSCs or monocytes (or indeed both) from the bone marrow, or whether the liver can induce quiescence. A small number of parasites colonise the lymph nodes, gut and lung, which exhibit high relatedness, indicating that parasites may circulate between these tissues. This could be a function of the type of host cells present and the chemokines they are attracted by. B). In PRO mice, monocytes are attracted to the liver, where a 'looser' granuloma structure means parasites are able to leave and are carried in these cells to other tissues. This may include parasites becoming resident in the gut or lung in a transient manner. Figure created in Biorender.

Figure 4.21 illustrates some key questions which are raised by these data and the hypotheses proposed.

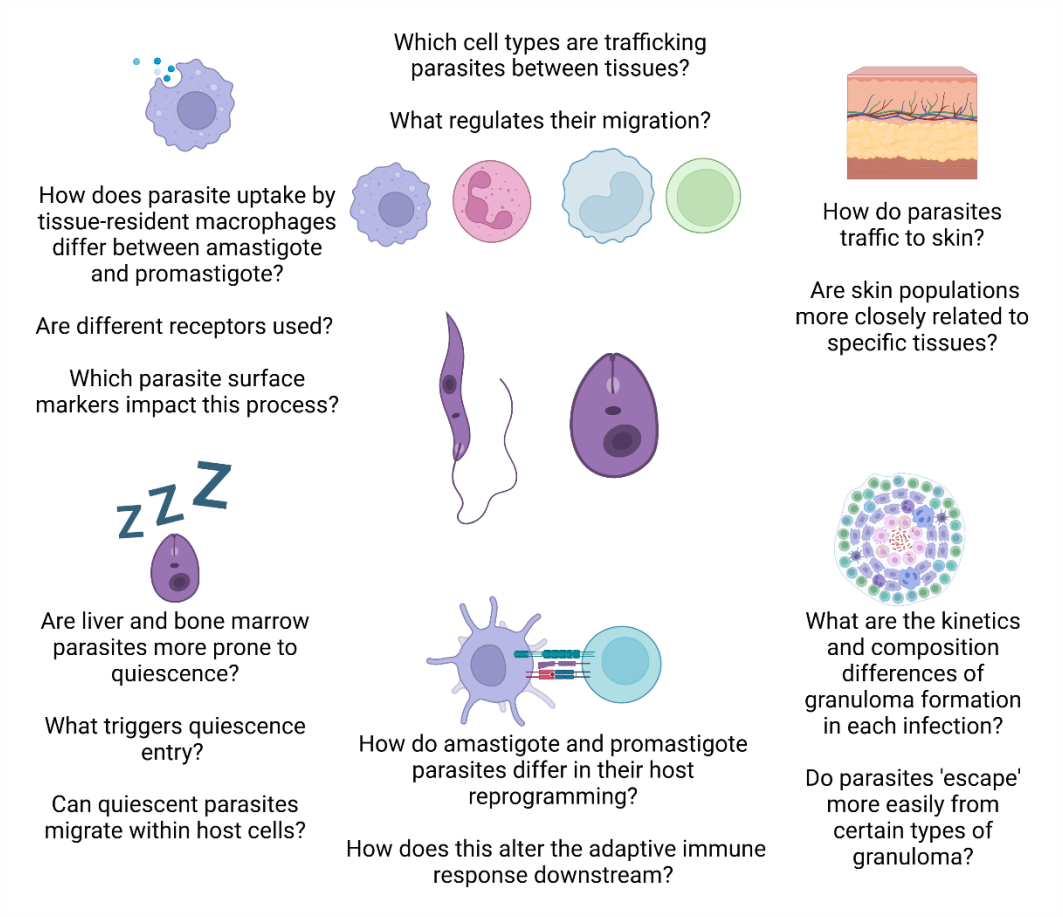


Figure 4.21: Some key questions to address on dissemination in promastigote and amastigote *L. donovani* infection. Figure created in Biorender.

Bottleneck events are important to understand not just within the host but also in terms of transmission of pathogens. This is of particular relevance in vector borne diseases (Forrester et al., 2012; Weaver et al., 2021), where the repeated bottlenecks of vector-host-vector transmission can affect pathogen evolution and thus disease patterns on a wider scale. Therefore, an important follow-up experiment would require assessment of the bottlenecks and diversity dynamics of *Leishmania* within the sand fly, in sand fly to host transmission and in host to sand fly transmission. In particular, understanding how parasites reach the skin and the level of bottleneck they undergo to do so is vital for understanding how parasite populations may change over multiple cycles of transmission between different hosts. Understanding these dynamics could establish how founder effects might alter parasite evolutionary trajectories, with implications for predicting disease

outbreaks (Weaver et al., 2021). This transmission bottleneck assessment would allow the dissemination of parasites to be assessed following this infection method, to compare to the data described here, which would also provide key insights into where the sand fly infection model might be more applicable to testing specific hypotheses, and where intravenous injection may be sufficient, depending on the specifics of the question being addresses.

Considering the host as an ecological community of pathogen populations has here allowed the application of powerful methods to decipher how *L. donovani* parasites interact with different tissues in a mouse model. The use of STAMP to understand bottleneck events has previously been established in various bacterial infection models (Abel et al., 2015a; Abel et al., 2015b; Bachta et al., 2020; Campbell et al., 2023; Hotinger et al., 2024; Hullahalli and Waldor, 2021; Jiang et al., 2024; Louie et al., 2019; Zhang et al., 2017), and to date in a single protozoan parasite model (Wincott et al, 2022), but bottlenecks in *in vivo Leishmania* dissemination have not been assessed. To our knowledge, community ecology-based analysis has not previously been applied in an *in vivo* infection in this manner, although one study has previously used an artificial barcode approach to study the diversity and assembly processes of *Ralstonia solacearum* in tomato plants (Jiang et al., 2024) using similar methods. Thus, this work provides a framework for integrating ecological methods with experimental infection models to understand pathogen population changes within the 'ecosystem' of the host, allowing a deeper understanding of host-pathogen interactions.

Our findings describe tissue colonisation and dissemination patterns which were previously unresolved, indicating differences between different life cycle stages on infection. This has led to further hypotheses to test, regarding key aspects of parasite-host interactions, cell trafficking and parasite quiescence, as well as a nuanced understanding of key considerations for experimental design. Assessing the underlying immune mechanisms causing both early and ongoing bottlenecks and dissemination patterns could help develop interventions to prevent disease progression, whilst understanding changes in parasite diversity could also provide vital insights into the evolution of *Leishmania* parasites *in vivo*.

Chapter 5: The Role of Complement in Early Bottlenecks and Dissemination Patterns in the C57BL/6J model of VL

5.1 Introduction:

The early response to infection is mediated via the innate immune system, which acts as a key initial barrier to the establishment of pathogens within a host. Early events are important in determining the course of disease and host pathology. Many elements of this process are evolutionarily conserved across vertebrate and even invertebrate species (Al-Sharif et al., 1998; Azumi et al., 2003; Jensen et al., 1981; Miller et al., 2007; Suzuki et al., 2002; Zhu et al., 2005); some pathogens have co-evolved mechanisms to subvert recognition and/or killing via these pathways, including protozoan parasites such as *Leishmania* (Kennedy et al., 2016; Sacks and Sher, 2002; Sikorski et al., 2020).

One key element of the innate immune response is the complement system. This set of protein cascades has three activation pathways. The classical pathway is initiated by antibody binding to antigens on pathogen surfaces, which can then bind to cascade-activating protease complexes. The lectin pathway relies on mannose-binding lectin (MBL) and pattern-recognition receptors (PRRs), whilst the alternative pathway is activated by C3 hydrolysis and formation of a C3 convertase complex as part of a feedback loop (reviewed in Dunkelberger and Song, 2009; Figure 5.1).

The cascades all converge at the level of C3 activation, where they initiate responses which fall broadly into three categories: i) lysis of pathogens is induced by the formation of the Membrane Attack Complex (MAC), which induces pore formation on pathogen surface membranes; ii) opsonisation, involving C3 cleavage products such as C3b and C3bi; and iii) the release of the anaphylatoxins C3a and C5a that promote chemotaxis of neutrophils, monocytes and macrophages to the site of infection (Dunkelberger and Song, 2009).

Whilst protective against invading pathogens, complement activation can also cause significant tissue damage and as such, a range of regulatory mechanisms exist to limit activation of the complement pathways. For example, inactivation of C3 is

controlled by the serine protease Factor I, in a manner dependent on cofactors, including Factor H, which can also hasten the decay of convertases. Other cofactors and direct inhibitors can block MAC assembly or inactivate various complement components (reviewed in Dunkelberger and Song, 2009; Figure 5.1).

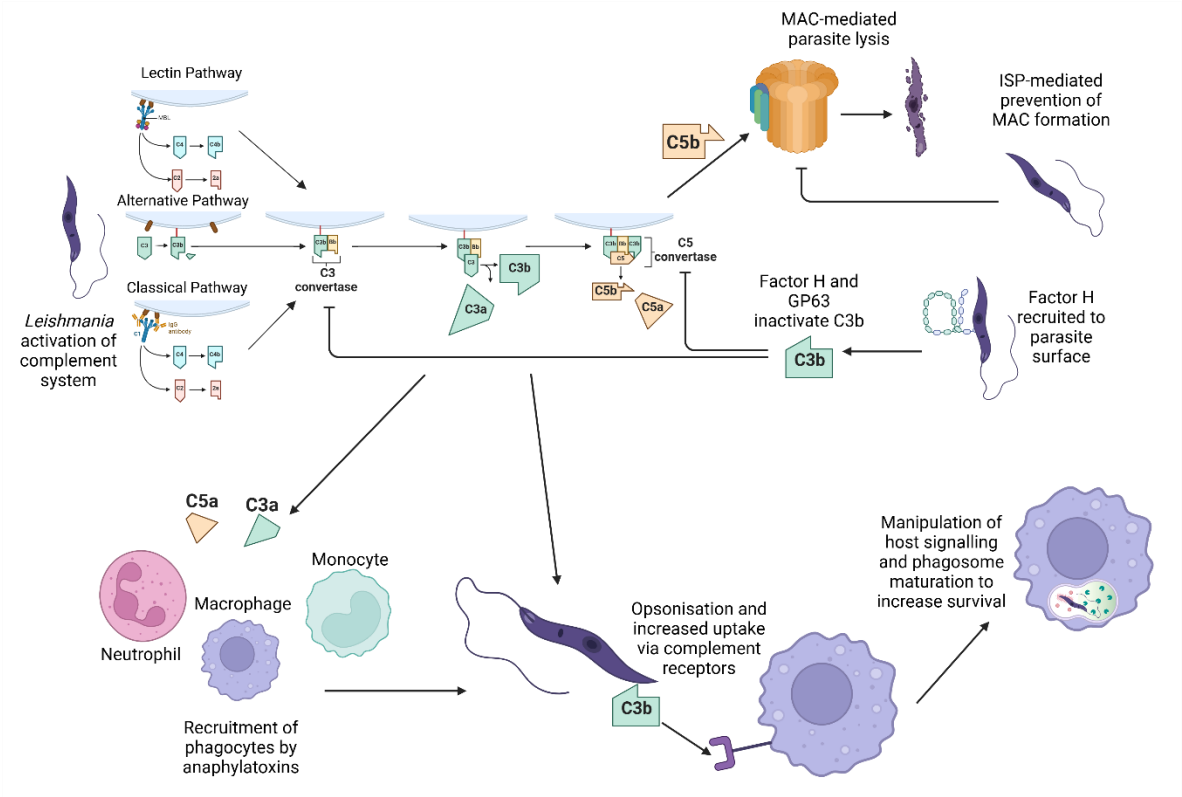


Figure 5.1: The complement system and its manipulation by *Leishmania* parasites. For details, see main text. Created in Biorender.

Leishmania can activate the complement cascade, with evidence supporting roles for all three pathways in this process (Ambrosio et al., 2021; Mosser et al., 1986; Mosser and Edelson, 1984; Tirado et al., 2021). Whilst complement can induce MAC-mediated lysis of *Leishmania* parasites (Mosser and Edelson, 1984; Hoover et al., 1984; Tirado et al., 2021), the lytic power of complement also appears to be highly dependent on both parasite and host species (Mosser and Edelson, 1984), as well as parasite life stage (Tirado et al., 2021); while procyclic promastigotes are relatively susceptible to lysis via MAC, amastigote and metacyclic parasites are more resistant to this process (Franke et al., 1985; Mosser and Edelson, 1984; Tirado et al., 2021). Human serum is noted to be powerfully lytic to *L. major* and *L. amazonensis*, whilst inducing lysis less efficiently for *L. donovani* (Franke et al.,

1985). Differences between sera from mouse, human, cat and dog lysis have been reported for *L. major* (Späth et al., 2003), *L. infantum* (Tirado et al., 2021), *L. enreittii* and *L. tropica* (Mosser and Edelson, 1984). *Leishmania* can also use multiple mechanisms to evade complement killing processes and to improve their own survival by taking advantage of complement-mediated effects (Figure 5.1). Verma et al. (2018) found that *L. donovani* parasites can inhibit the formation of MAC to prevent lysis, via C3-inhibition by inhibitor of serine peptidase 2 (ISP2). Early studies using promastigote *L. major* parasites, meanwhile, found that they can prevent or reverse insertion into the membrane, blocking lytic activity (Puentes et al., 1990). Complement-mediated opsonisation can aid parasite uptake by macrophages (Mosser and Brittingham, 1997; Brittingham et al., 1995; Jacobs et al., 2005; Laufs et al., 2002; Matte et al., 2021; Mosser and Edelson., 1984) and can increase parasite survival by delaying maturation of the lysosome (Polando et al., 2013; Matte et al., 2021). LPG on the parasite surface has been shown to aid in this process in *L. donovani* (Desjardins and Descoteaux, 1997; Quintela-Carvalho et al., 2022); LPG can also assist survival by reducing susceptibility to oxidative burst after neutrophil uptake of *L. infantum* parasites (Quintela-Carvalho et al., 2022). The expression and structure of LPG is known to differ between promastigote and amastigote life stages (McConville and Blackwell, 1991; Moody et al., 1993; Bahr et al., 1993). The GP63 surface protease can inactivate C3b (Brittingham et al., 1995; Brittingham and Mosser, 1996) and can also prevent parasite killing via oxidative burst (Matte et al., 2016), whilst altering the host cell inflammatory response to create a more hospitable environment for parasite replication and survival (Chan et al., 2021). *L. donovani* have been demonstrated to use ISPs to upregulate C5aR signalling on host cells; this leads to a protective effect, whereby parasites show increased intracellular survival *in vitro*, mediated by altering the balance of the cytokines IL-12 and IL-10 via the PI3K/AKT pathway (Verma et al., 2018). Parasites can also recruit host Factor H to assist in C3b inactivation to modulate the complement cascade (Filho et al., 2021).

Another key role of complement is in the recruitment of neutrophils via C3a and C5amediated chemotaxis (Jacobs et al., 2005; Metzemaekers et al., 2020). Neutrophils are a key early responder in *Leishmania* infections, but their role is complex and can either enhance or decrease disease progression in ways which

appear to depend on the timing and tissue-specific locale in which they act (Carlsen et al., 2015; Laufs et al., 2002; Ribeiro-Gomes and Sacks, 2012). Neutrophils can act as a host cell for many species of *Leishmania* parasites (Carlsen et al., 2015; Chang, 1981; Pearson and Steigbigel, 1981; Ribeiro-Gomes and Sacks, 2012; Romano et al., 2017), with some evidence indicating they play a role in increased parasite survival (Carlsen et al., 2015; Ribeiro-Gomes and Sacks, 2012; Romano et al, 2017) and other studies showing potent parasite killing (Chang, 1981; Pearson and Steigbigel, 1981). Hurrell et al (2015) showed that neutrophils can act as a protective host cell for *L. mexicana* early in infection, and that neutrophils can impair lesion healing processes in CL mouse models. Rapid recruitment of neutrophils to the draining lymph nodes in *L. mexicana* infection has been shown to occur in a C5a-dependent manner, with the incumbent neutrophils originating from blood circulation and not the initial infection site (Prat-Luri et al., 2022). Interestingly, lyphatic endothelial cells can produce C5 locally, but the importance of this local, as opposed to systemic, production in neutrophil chemotaxis is not yet established (Prat-Luri et al., 2022). Neutrophil NET-osis may be important in increasing parasite phagocytosis by other cells for *L. donovani* parasites specifically (de Menezes et al., 2016; Gabriel et al., 2010); the induction and impact of NETs on *Leishmania* parasites appears to vary depending on species and life stage (de Menezes et al., 2016; Gabriel et al., 2010; Guimarães-Costa et al., 2009; Oliveira et al., 2023; Oualha et al., 2019). Neutrophils have also been shown to promote visceral disease in an IFNβ-dependent manner (Dias et al., 2019). Dias et al (2019) described a role for Neutrophil Elastase expression in macrophages leading to induction of IFNβ after *L. donovani* infection, which increased parasite survival, whilst Xin et al (2010) showed that Type 1 IFN was important for clearance of *L. amazonensis in vivo*.

There are many known sex differences in immune responses, including those of the innate immune system. Previous studies have established that there are differences in complement activity between male and female mice, with studies findings indicating higher pathway activity levels (Holt et al., 2001; Kotimaa et al., 2016; Tanaka et al., 1986) and pathogenic outcomes of complement-deficiency (Bloom et al., 2016) in males. In a similar vein, differences between neutrophil maturity, phenotypic profile, metabolism and functionality have also been reported in multiple studies (Blazkova et al., 2017; Kim et al., 2022; Lu et al., 2021; McGill et al.,

2023; Molloy et al., 2003; Pokhrel et al., 2020; Xie et al., 2020). Kotimaa et al. (2016) found that activity differences occur post-C3 activation, and specifically limit lytic function in female C57BL/6J and BALB/c mice via reduced C6 and C9 activation. Conversely, other studies have indicated that differences may also occur at the C3 level, with Pokhrel et al (2020) reporting that female mice have increased C3 serum levels compared to males, and that female bone marrow neutrophils also have higher expression of complement receptor 3 (CR3) than neutrophils from male mice. Additionally, evidence shows that female neutrophils are more able to kill some pathogens, including *Staphylococcus aureus* (Pokhrel et al., 2020). In the case of *S. aureus*, blocking CR3 induced a greater decrease in killing power in female neutrophils, compared to males, with CR3-induced ROS induction being increased in female bone marrow neutrophils immediately after infection with opsonised bacteria (Pokhrel et al., 2020). Castleman et al (2018) similarly found that female neutrophils were more efficient at bacterial killing.

Based on this body of data and our findings of a restrictive bottleneck to tissue colonisation in C57BL/6J mice at 2 days p.i (Chapter 4), here we sought to determine whether complement played a role in this process and whether there were any associated sex differences.

5.2 Results:

5.2.1 Complement plays a role in imposing bottlenecks to colonisation in the early stages of *L. donovani* infection of C57BL/6J mice

To assess how complement impacts early tissue bottlenecks, male and female C57BL/6J (WT) and mice and B6.129S4-C3tm1Crr/J (C3KO) (Wessels et al., 1995) mice were infected with 1x10^8 promastigotes, consisting of the full barcoded library of LV9 T7 cas9 *L. donovani*. The experimental set up is shown in Figure 5.2. The mice were killed two days p.i. and the spleen, liver, one lymph node and two sets of tibia plus femur were taken. All tissues except bone marrow were flash frozen on dry ice and stored at -80 °C until use for DNA extraction. Bone marrow was flushed, pelleted and stored at -20 °C until DNA extraction was performed. The

DNA samples were amplified using the Barseq-PCR protocol previously described, then sent for Illumina sequencing and processed for Barseq analysis.

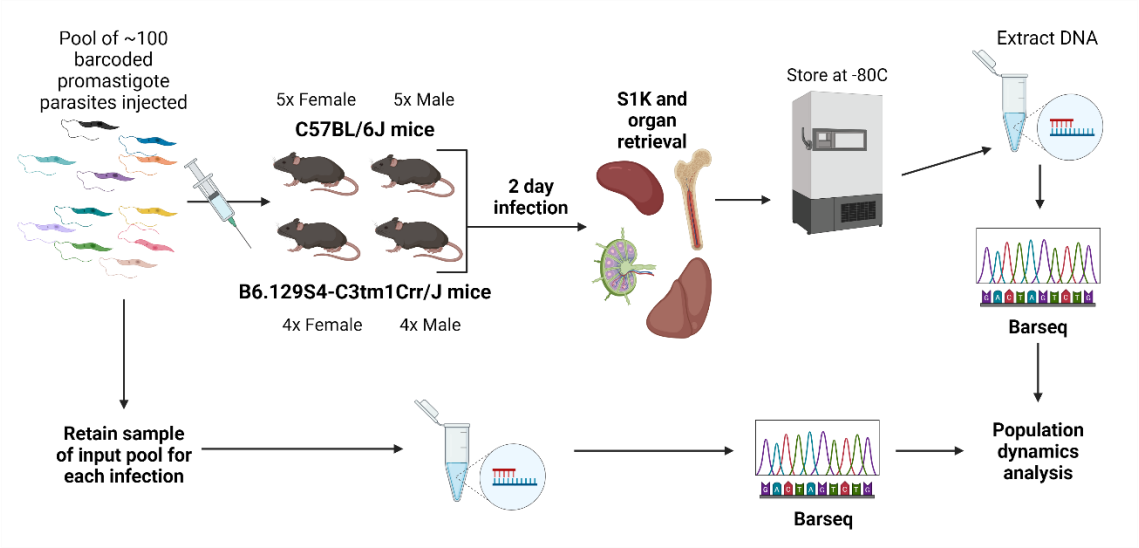


Figure 5.2: Schematic of experimental plan to investigate the role of complement in early infection population dynamics of *L. donovani* in C57BL/6J mice. A). Schematic illustration of the experimental protocol for the infection set-up. Five male and five female C57BL/6J and four male and four female B6.129S4-C3tm1Crr/J mice were infected with 1x10^8 barcoded LV9 T7 cas9 *L. donovani* stationary phase promastigotes, consisting of a library of 102 isogenic 20 bp barcodes, per mouse, via tail intravenous injection Tissues were taken from each mouse and flash frozen on dry ice, then stored at -80 °C before subsequent DNA extraction, Illumina sequencing, Barseq analysis and population dynamics studies, alongside a sample of the input parasite pool from each infection. Tissues taken consisted of liver, spleen, one lymph node and two each of tibia plus femur. Bone marrow was extracted by flushing and stored at -20 °C for DNA extraction. Figure created in Biorender.

FP was calculated as previously described (Chapter 3; Abel et al., 2015). The liver exhibited the highest FP of all four tissues in both WT and C3KO mice, but there was no significant difference between the two mouse lines in this tissue when all mice were compared without regard for sex (Figure 5.3A). However, the heterogeneity seen in the WT mice was far greater than that in the C3KO line for this tissue. A significant increase in FP was observed in the bone marrow (MW: $p<0.0001$) and spleen (MW: $p=0.0031$) of C3KO mice compared to wild-type mice at 2 days p.i. In contrast to the trend seen in the liver, there was wider heterogeneity in the C3KO

mice in the bone marrow and spleen compared to the WT mice (Figure 5.3A). No differences were seen between male and female mice of the same strain when FP for each tissue was directly compared, but an effect of C3-deficiency was found in specific tissues. Whilst both sexes showed a significant increase in bone marrow FP in the C3KO mice (MW_{MALES}: p=0.0005; $MW_{FEMALES}$: p=0.0005), only males showed an increase in the liver (MW: p=0.0159) and spleen (MW: p=0.0159) FP when compared to WT (Figure 5.3B).

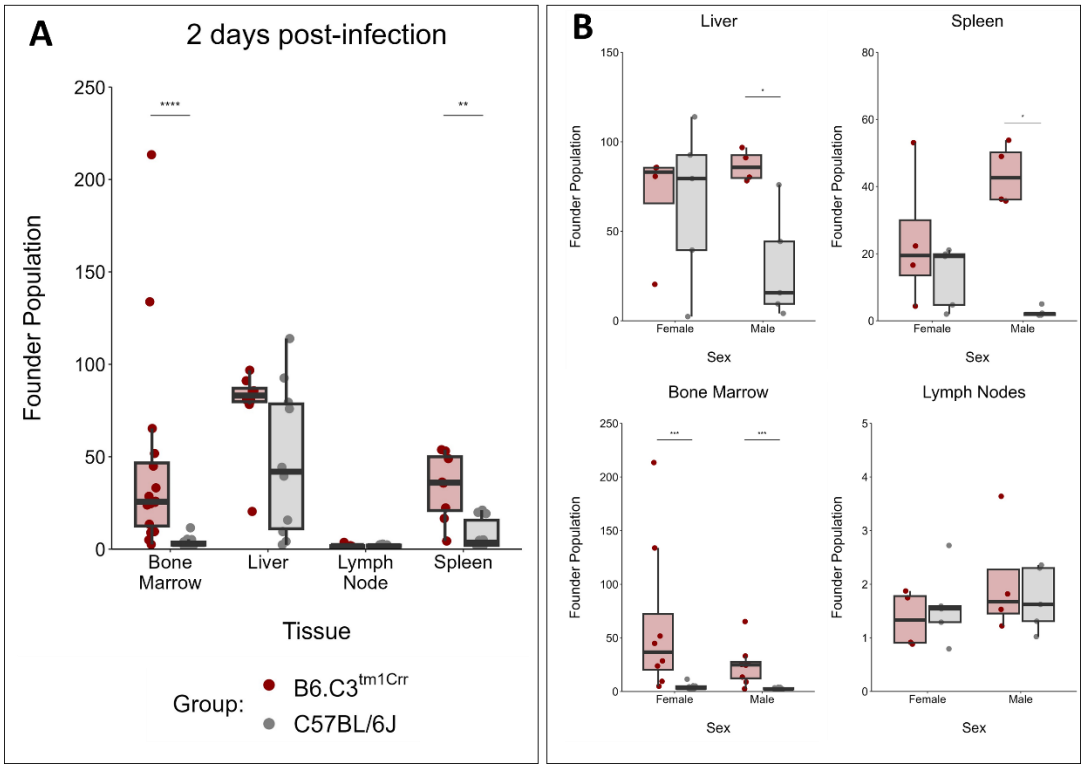


Figure 5.3: Founder populations in multiple tissues in complement-deficient and wild-type C57BL/6J mice infected with promastigote *L. donovani* parasites at 2 days post-infection. A). Founder populations for WT and C3KO mice at 2 days post-infection with 1x10^8 *L. donovani* stationary phase promastigote parasites via intravenous tail injection. (WT, n=5 males and 5 females; C3KO, n=4 males and 4 females). For bone marrow, two were taken per mouse. B). Sex split comparisons of founder populations in each tissue. Each point is an individual mouse tissue. Colour indicates mouse genetic background as shown in the key and is consistent across all panels. Mann-Whitney-Wilcoxon test. Significance of p-value: * p<0.05; **p<0.01; *** p<0.001: **** p<0.0001. Boxplots show the median, 25th and 75th quartiles, with whiskers representing all data-point within 1.5x Interquartile Range of the median; points above or below the whiskers represent outlying data values.

In summary, these data provide evidence that C3 can contribute to early bottlenecks to colonisation of specific tissues in mice infected with *L. donovani*. The importance of this complement-mediated bottleneck appears to differ between male and female mice, with males relying more heavily on complement-mediated processes to prevent parasite establishment in the liver, bone marrow and spleen.

5.2.2 Impact of C3-deficiency on parasite relatedness between tissues

GD was calculated for pairwise comparisons of parasite populations in each tissue, to understand how C3-deficiency impacts population relatedness and, by extension, parasite dissemination patterns (Figure 5.4). Overall, relatedness was greater in WT mice compared to C3KO mice for comparisons between liver and lymph node (MW: $p=0.0044$) (Figure 5.4A), spleen and lymph node (MW: $p=0.0021$) (Figure 5.4B) and bone marrow and lymph node (MW: $p<0.0001$) (Figure 5.4C). Relatedness was greater in the C3KO mice compared to WT mice for liver to bone marrow (MW: $p=0.0035$) and liver to spleen (MW: $p=0.0005$) population comparisons for amalgamated mice (Figure 5.4A).

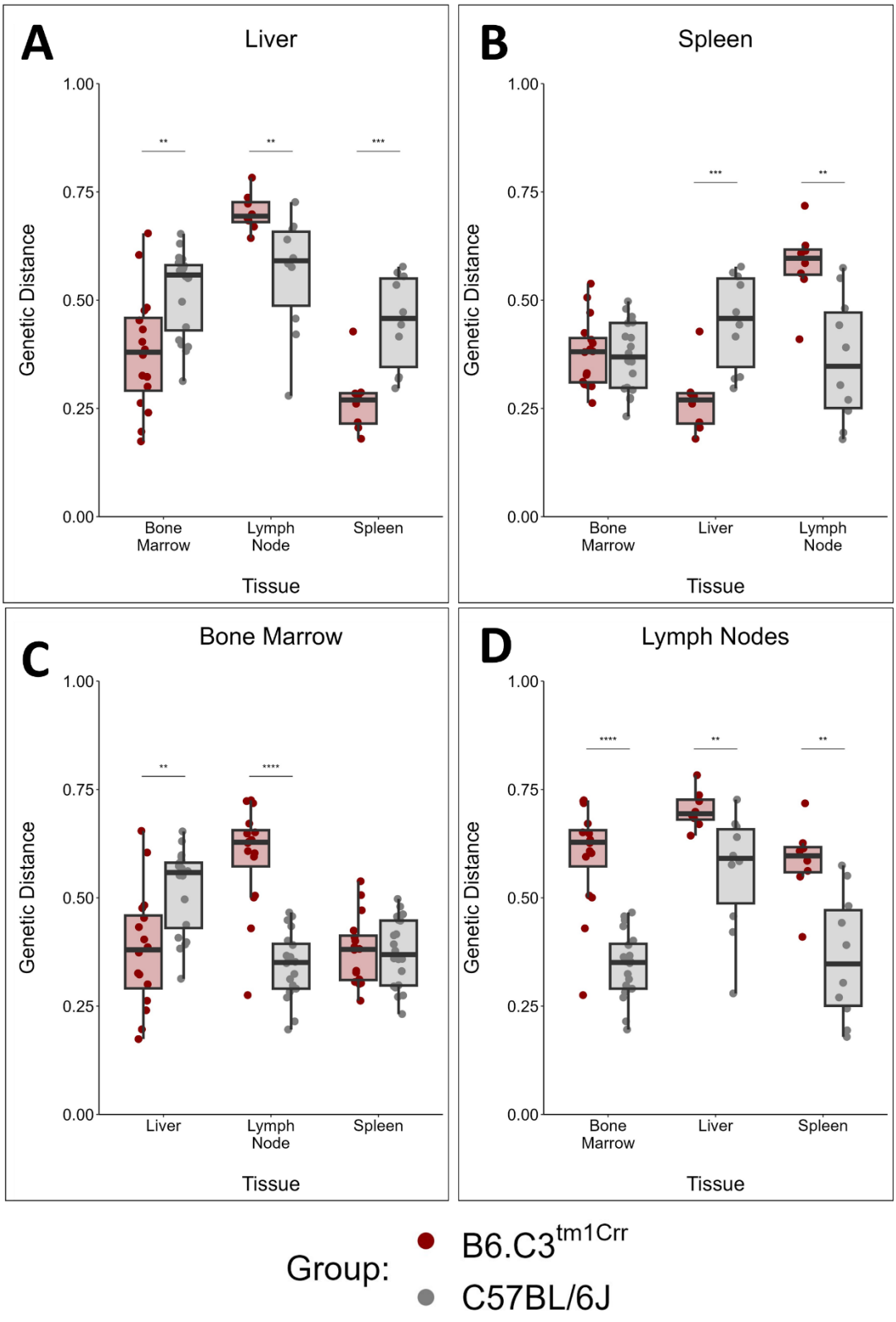


Figure 5.4: Genetic distance analysis for comparisons between multiple tissues in complement-deficient and wild-type C57BL/6J mice infected with promastigote *L. donovani* parasites at 2 days post-infection. A-D). Genetic distance for pairwise comparisons of individual tissues in WT and C3KO mice at 2 days post-infection with 1x10^8 *L. donovani* stationary phase promastigote parasites (WT, n=5 males and 5 females; C3KO, n=4 males and 4 females. For bone marrow, two were taken per mouse. Each data point

represents a pairwise comparison between two tissues. Colour indicates mouse genetic background as shown in the key and is consistent across all panels. Mann-Whitney-Wilcoxon test. Significance of p-value: * p<0.05; **p<0.01; *** p<0.001: **** p<0.0001. Boxplots show the median, 25th and 75th quartiles, with whiskers representing all data-point within 1.5x Interquartile Range of the median; points above or below the whiskers represent outlying data values.

When mice were split by sex (Figure 5.5), males again exhibited a greater impact when C3 was absent. In male C3KO mice, relatedness increased between liver and spleen (MW: p=0.0159). Relatedness was greater in WT mice for liver and lymph node (MW: p=0.0159), spleen and lymph node (MW: p=0.0159), spleen and bone marrow (MW: p=0.0077), and bone marrow and lymph node (MW: p=0.0021) (Figure 5.5A). In females (Figure 5.5B), only bone marrow and lymph nodes showed an effect of complement-deficiency, with WT mice showing more relatedness than C3KO mice between these two tissues (MW: p=0.0005). When male and female mice were directly compared, it was found that male WT mice showed a greater level of relatedness between bone marrow and spleen (MW: p=0.000487) and bone marrow and lymph nodes (MW: p=0.000725) than females, but no differences between males and females were observed in the C3KO mice.

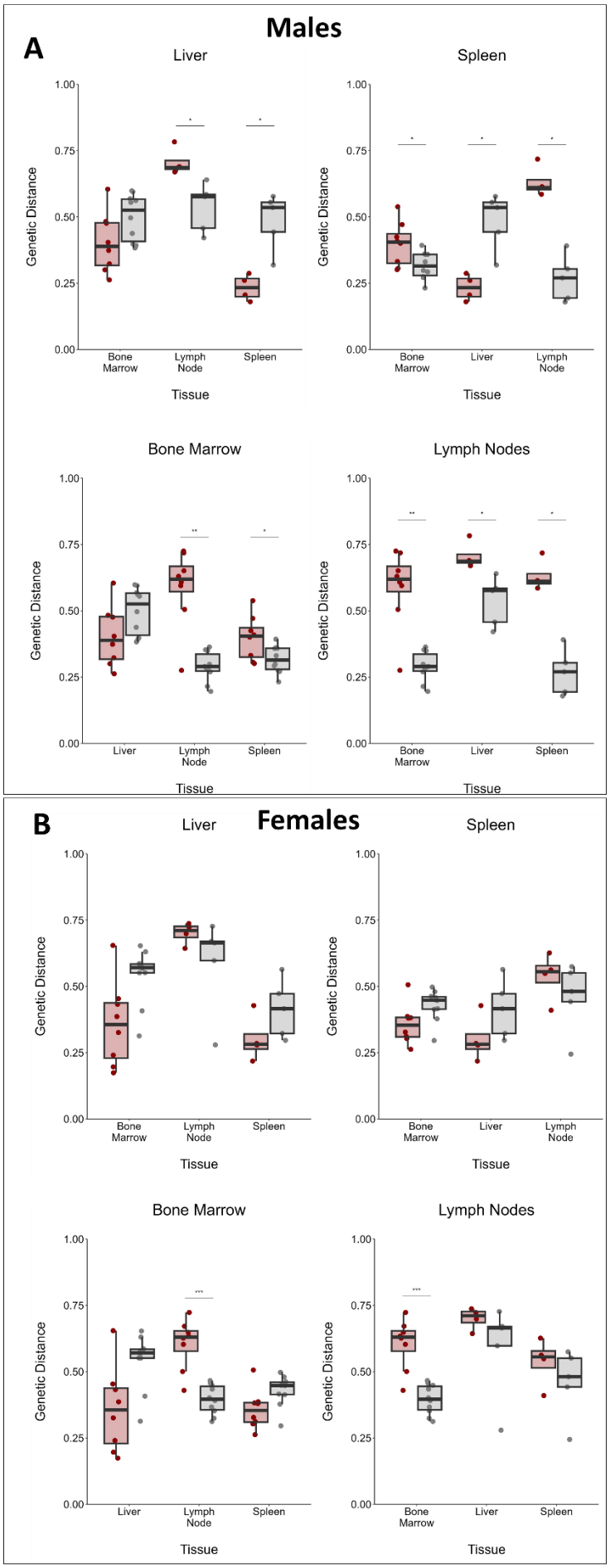

Males
A
Liver
Spleen
Bone Marrow
Lymph Nodes
Genetic Distance
1.00
0.75
0.50
0.25
0.00
Bone Marrow
Lymph Node
Liver
Spleen
Tissue
Females
B
Group:
B6.C3tm1Crr
C57BL/6J

Figure 5.5: Genetic distance analysis for comparisons between multiple tissues in complement-deficient and wild-type C57BL/6J mice infected with promastigote *L. donovani* parasites at 2 days post-infection. A). B). Genetic distance for pairwise comparisons of each tissue in male (A) and female (B) WT and C3KO mice at 2 days post-infection with 1x10^8 *L. donovani* stationary phase promastigote parasites (WT, n=5 males and 5 females; C3KO, n=4 males and 4 females). For bone marrow, two were taken per mouse. Each data point represents a pairwise comparison between two tissues. Colour indicates mouse genetic background as shown in the key. Mann-Whitney-Wilcoxon test. Significance of p-value: * p<0.05; **p<0.01; *** p<0.001: **** p<0.0001. Boxplots show the median, 25th and 75th quartiles, with whiskers representing all data-point within 1.5x Interquartile Range of the median; points above or below the whiskers represent outlying data values.

Overall, relatedness between liver and spleen increased in the absence of complement, whilst relatedness between lymph nodes and all other tissues decreased when C3 was absent. Together, these data indicate that the presence of C3 is important in regulating tissue-specific parasite dissemination dynamics, and plays a more vital role in male mice than in female mice.

5.2.3 Mouse serum does not kill promastigote parasites *in vitro*

A possible explanation for the above finding is that C3-deficiency results in changes to initial parasite load due to complement-dependent killing of parasites early after infection. To assess whether mouse serum was lytic for promastigotes, an *in vitro* killing assay using sera from male and female WT and C3KO mice was performed. The experimental set up is illustrated in Figure 5.6. Serum from a minimum of two mice in each category was pooled, with half of each heat-inactivated to denature complement. Sera were then serially diluted in triplicate, and 1x10^6 parasites were added to each well. Controls included normal human serum and media only. Stationary phase parasites, at a density of 1x10^6 per well, were added and incubated at 37 °C for approximately 1 hour, and were then visually scored. Scoring was done on a scale of 1-5, with 1 being almost all parasites dead and 5 being almost all alive (see Methods for details).

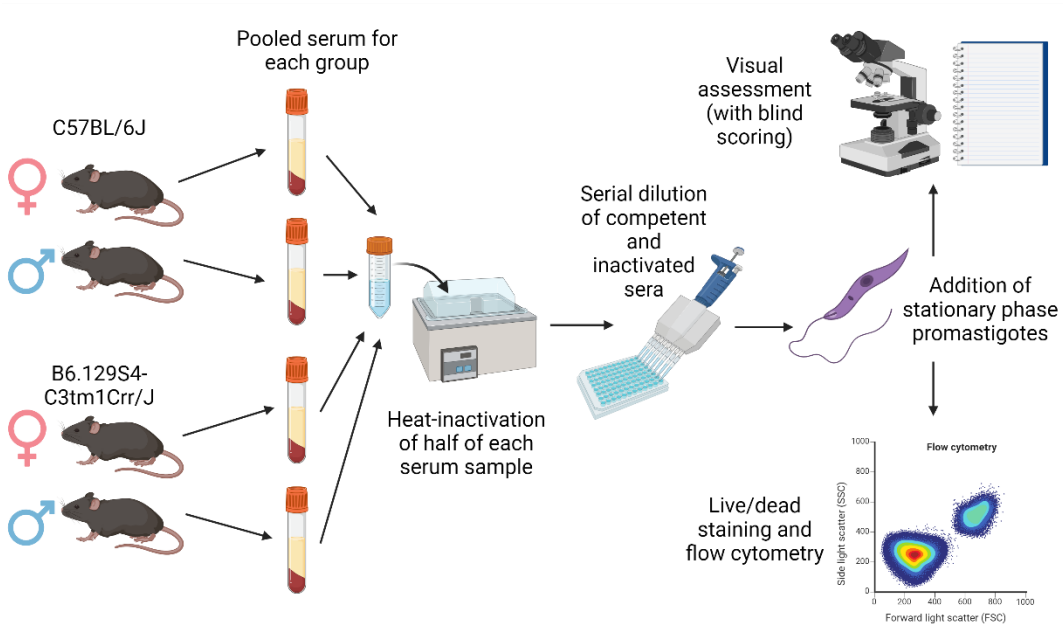


Figure 5.6: Experimental protocol to assess *L. donovani* parasite lysis by murine serum complement *in vitro*. Blood serum was taken via cardiac bleeds from male and female WT and C3KO mice and pooled. Half of each batch was heat-inactivated. Parasite death in serially diluted samples was assessed by eye, on a 1-5 scale as detailed in Methods. Where indicated, parasites were pooled and stained with LDAqua to assess live versus dead proportion via flow cytometry. Figure created in Biorender.

Parasite death at high concentrations was clearly visible in the human serum control, but estimation of viability in mouse sera-treated samples proved unreliable. Therefore, a flow cytometry-based assay was adopted instead. Triplicate wells of parasites were pooled, then stained with LDAqua to allow discrimination of live versus dead parasites, using the gating strategy shown in Figure 5.7A. The data shows the gating applied to both unstained (Figure 5.7A) and stained (Figure 5.7B) cells. The human serum control (Figure 5.7C) again showed a clear trend of increased parasite death at higher concentrations, with almost 95% of parasites killed in 50% normal human serum.

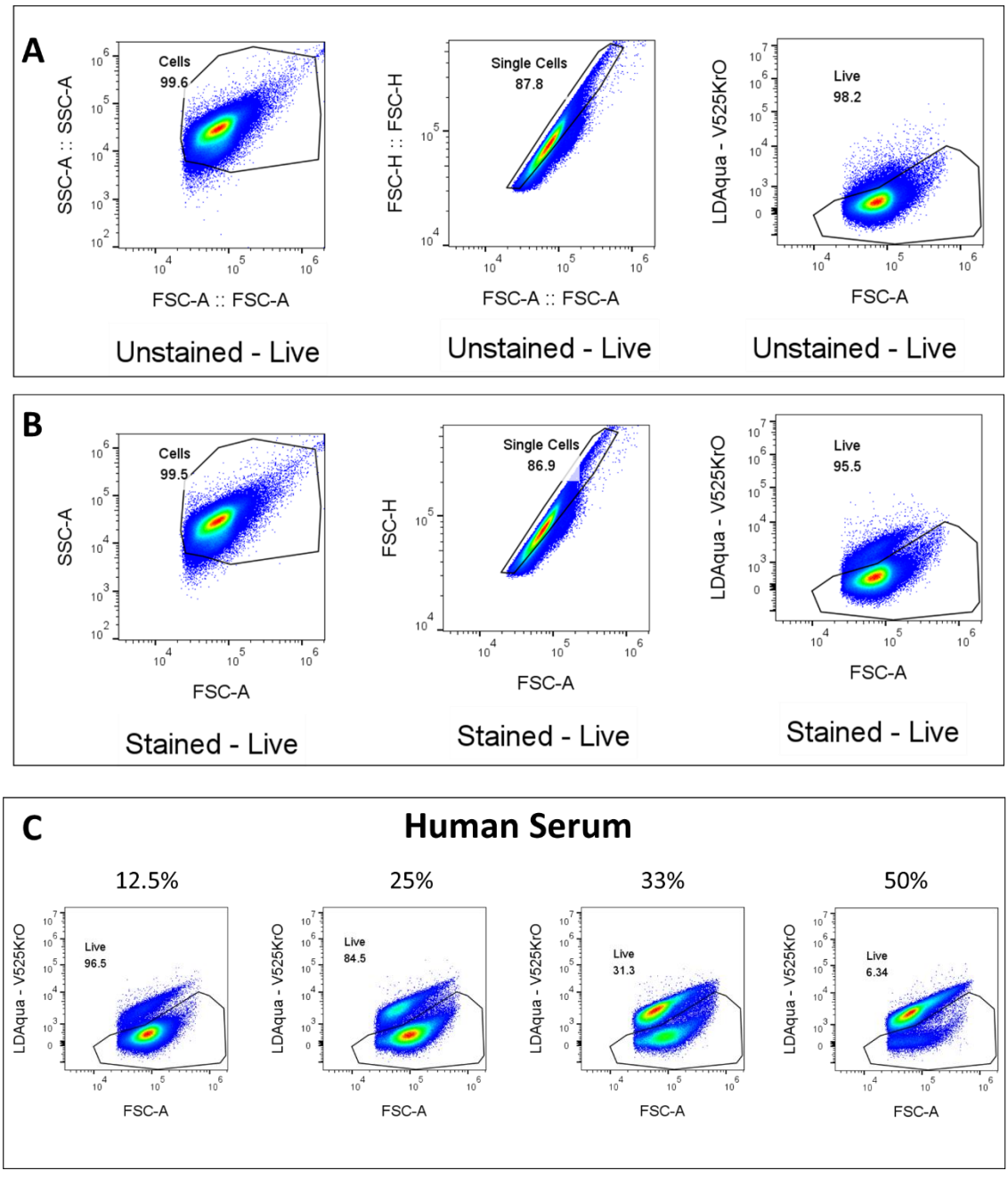


Figure 5.7: Flow gating strategy for assessment of live versus dead *L. donovani* promastigotes. Stationary phase promastigotes were stained with LDAqua, then assessed for staining using flow cytometry, using FSC and SSC to identify single cells. A). Unstained parasites in media. B). Stained parasites in media. C). LDAqu stained parasites following incubation in diluted normal human serum. Plots were generated using FlowJo software.

In contrast to normal human serum, neither WT nor C3KO serum showed any significant killing activity over the concentration range tested, irrespective of heat inactivation.

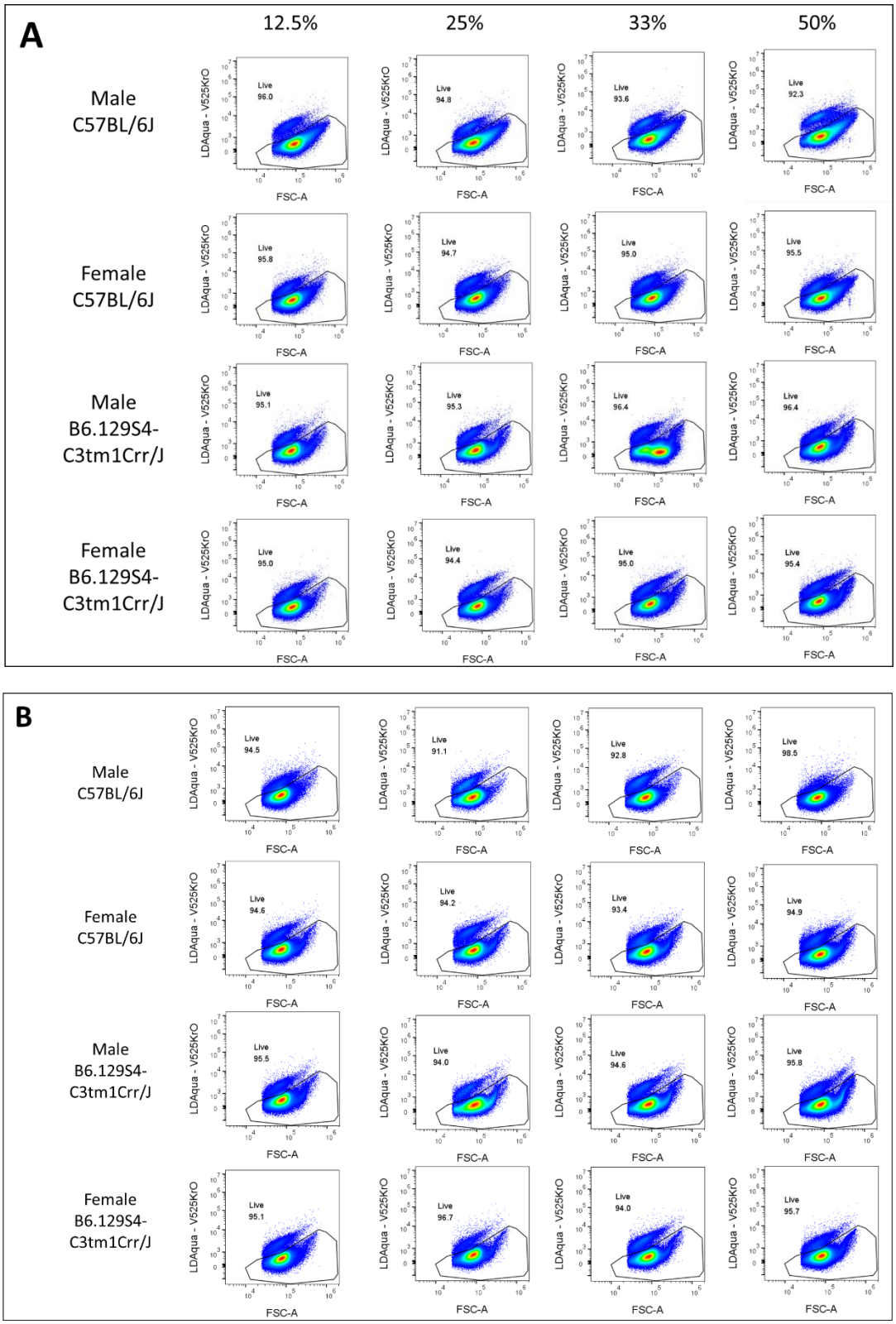


Figure 5.8: Murine serum fails to kill *L. donovani* promastigotes. Stationary phase promastigotes were incubated at a density of 1x10^6 per well with untreated or heat-inactivated normal mouse serum from male or female WT or C3KO mice. After 1 hour incubation at 37 °C parasites were stained with LDAqua and assessed for viability by flow cytometry. A). Serum without heat-inactivation. B). Serum with prior heat-inactivation.

Percentages indicate the serum concentration the parasites were incubated with and apply to both panels. Plots were generated using FlowJo software.

Figure 5.9 shows data from two independent experiments. In the first experiment, the results were slightly more scattered, but the comparison to untreated live cells (Figure 5.9A) indicates that the mouse serum-treated parasites were not dying at an increased rate. Human serum clearly results in parasite death at higher concentrations (indicated by the black circles in Figures 5.9B and C), but increasing serum concertation in the mouse-sera samples does not lead to parasite death (Figure 5.9B and C), as live percentages are similar to those in the untreated samples in Figure 5.9A in both experiments.

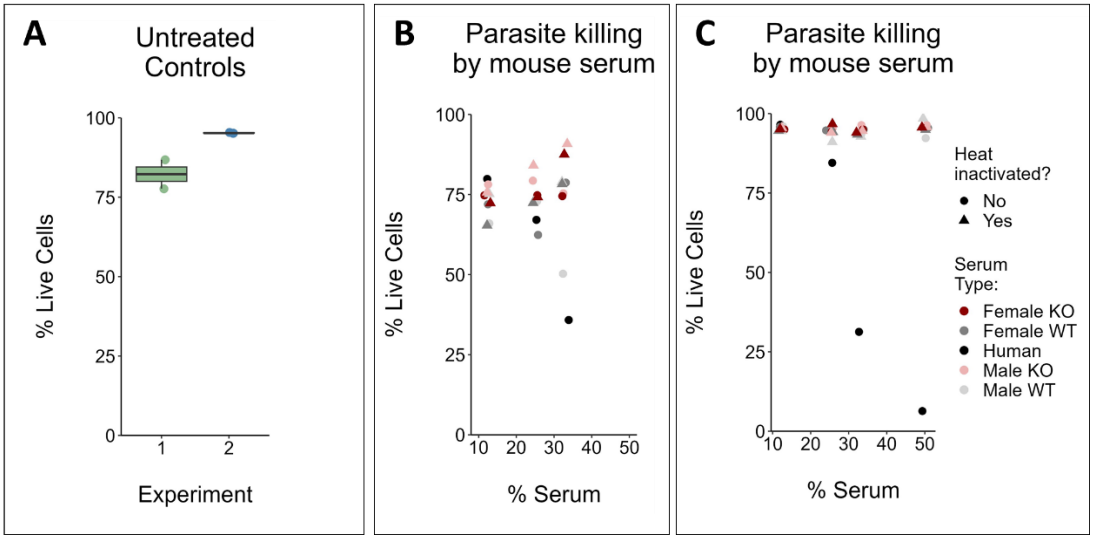


Figure 5.9: Live cell percentage using flow cytometry analysis of promastigote *L. donovani* parasite death after incubation with male and female C57BL/6J (WT) or B6.129S4-C3tm1Crr/J (C3KO) serum. Stationary phase promastigotes were incubated with murine serum from WT or KO male or female mice, at a density of 1x10^6 per well, with or without prior heat-inactivation, for 1 hour at 37 °C. Triplicate wells of parasites were pooled and stained with LDAqua, then assessed for staining using flow cytometry. A). Live cells from the two experiments with no serum treatment applied, to define the control live percentage expected. Green indicates experiment 1 (n=2) and blue indicates experiment 2 (n=3). B and C). Live parasite percentage after treatment with each serum type for the experiments termed 1 and 2 respectively. Colour scheme for plots B and C is shown on the right, and

indicates serum type. Shape indicates whether serum was heat-inactivated prior to incubation.

Collectively, these data indicate that mouse serum lacks lytic activity against *L. donovani* promastigotes and thus that the mechanism by which C3 contributes to early bottlenecks *in vivo* is unlikely to be due to differences in early parasite lysis in the blood post-infection. These data suggest that a more complex C3-dependent mechanism influences parasite colonisation of tissues.

5.2.4 The role of complement in early infection bottlenecks is both tissue and sex dependent

To further understand the role of complement in early colonisation bottlenecks, a repeat *in vivo* infection using the two mouse lines was performed, but with the additional inclusion of a serum 'addback' group; this encompassed an intraperitoneal injection of sex-matched WT serum into C3KO mice prior to infection with the barcoded *L. donovani* library. These mice are referred to as 'serum addback' mice. The experimental set up is shown in Figure 5.10, with a dose of 5x10^7 promastigotes per mouse, due to the previously mentioned choice to reduce parasite dose for welfare reasons. All tissues except bone marrow were flash frozen on dry ice and stored at -80 °C, whilst bone marrow was flushed and stored at -20 °C, for DNA extraction. Tissue sections were also obtained for liver, spleen, gut and lung from each mouse, along with serum samples from cardiac bleeds. An additional two male and female mice of each line were not infected and were used for control sera and tissue sections.

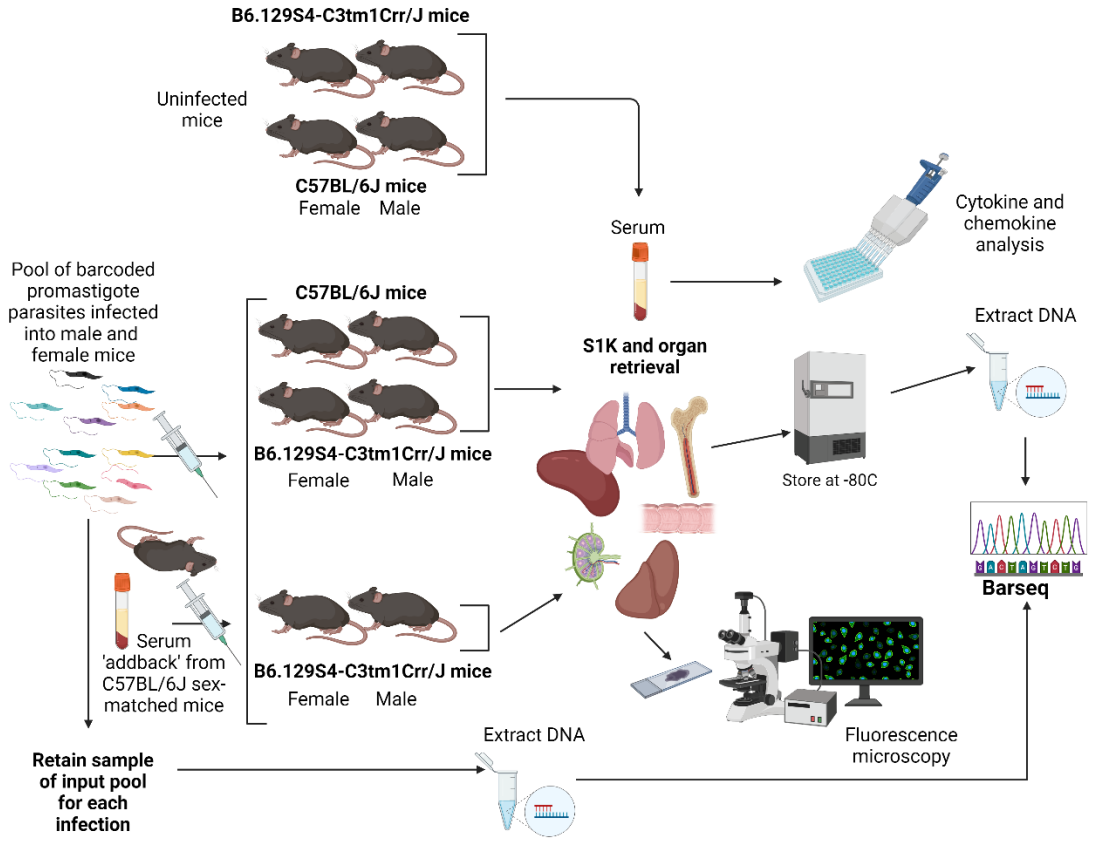


Figure 5.10: Schematic of experimental plan to investigate the role of complement in early infection population dynamics of *L. donovani* in C57BL/6J mice. A). Schematic illustration of the experimental protocol for the infection set-up. WT (n=5 males and 5 females), C3KO (n=5 males and 4 females) and "serum addback" (n=5 males and 4 females) mice were infected with 5x10^7 barcoded LV9 T7 cas9 *L. donovani* stationary phase promastigote parasites, consisting of a library of 102 isogenic 20 bp barcodes, per mouse, via tail intravenous injection. For serum 'addback' mice, C3KO mice were given an intraperitoneal injection of pooled sex-matched WT serum prior to infection. Tissues were taken from each mouse and flash frozen on dry ice, then stored at -80 °C before subsequent DNA extraction, Illumina sequencing, Barseq analysis and population dynamics studies, alongside a sample of the input parasite pool from each infection. Tissue sections of liver, spleen, gut and lung for immunofluorescence analysis from all mice, plus two male and two female naïve mice of each strain, were also stored in OCT at -80 °C. Serum samples were taken via cardiac bleeds for all mice. Tissues taken consisted of liver, spleen, gut, lung, two lymph nodes and two each of tibia plus femur. Bone marrow was extracted by flushing and stored at -20 °C for DNA extraction. Figure created in Biorender.

Bottlenecks in each tissue were assessed using the STAMP protocol, as previously described, to calculate FP for each tissue (Figure 5.11). Males again showed a greater impact for loss of complement than female mice, but the effect seen in the previous experiment was not recapitulated; FP in bone marrow, liver and spleen was not found to differ significantly between WT and C3KO mice for either sex. Of note, the liver FP was approximately ten times greater for both C3KO and WT mice in this experiment than in the previous study (Figure 5.3). However, in the lung, gut and lymph nodes, loss of C3 resulted in reduced FP values in male mice, but only affected FP in the lymph nodes of female mice ($MW_{male\ lung}$: p= 0.024; $MW_{male\ gut}$: p= 0.024; $MW_{male\ LN}$: p= 0.034; $MW_{female\ LN}$: p=0.044). Addition of serum to C3KO mice resulted in a reduced FP in the spleen of male mice (MW: p=0.048) and in the lymph nodes of female mice (MW: p=0.044). Serum addback also led to lower FP values than the WT mice (but not the C3KO mice) in the gut (MW: p=0.024) and lymph nodes (MW: p=0.021) of male mice.

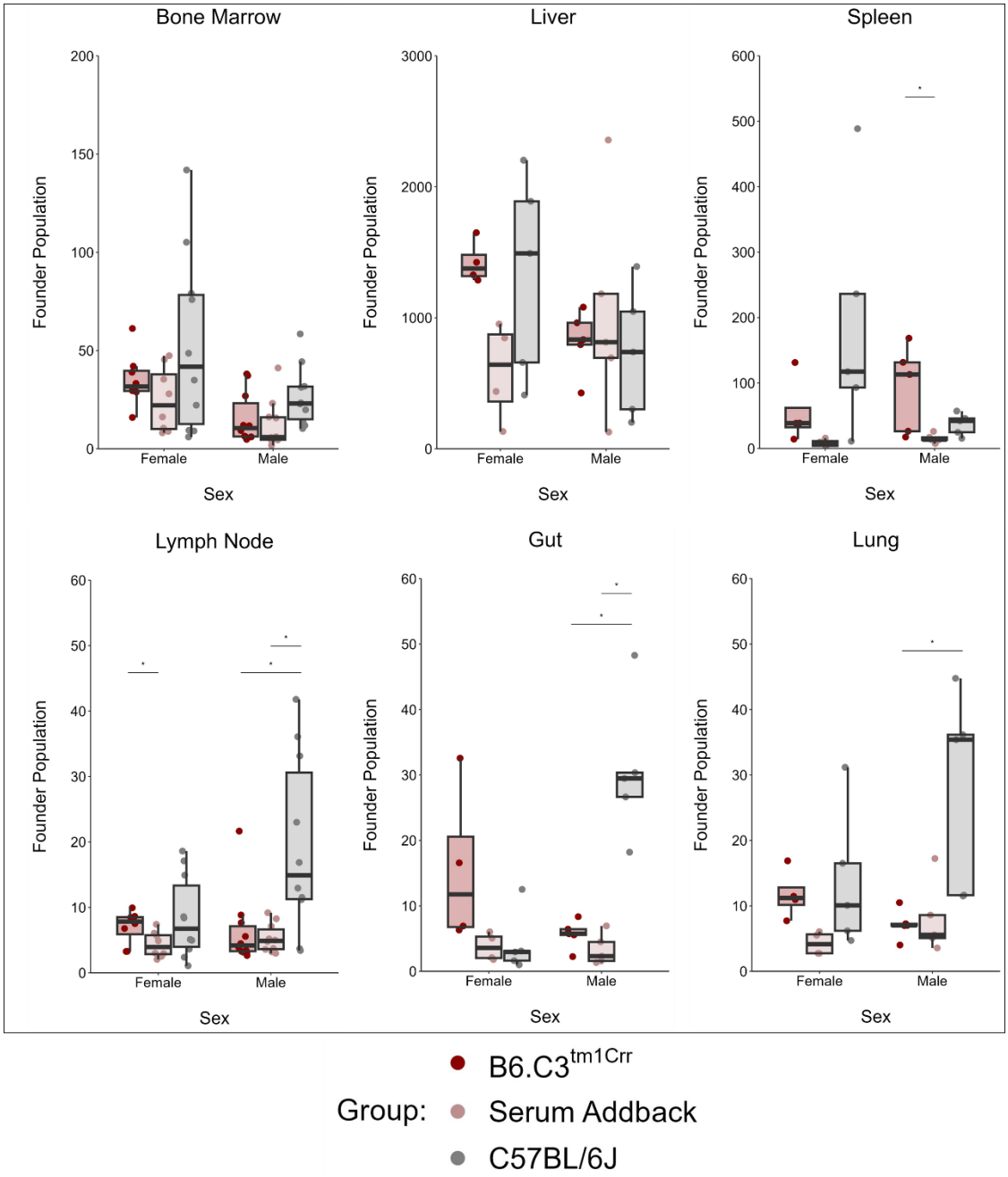


Figure 5.11: Founder populations in multiple tissues in complement-deficient (with or without serum 'addback') and wild-type C57BL/6J mice infected with promastigote *L. donovani* parasites at 2 days post-infection. Founder populations in WT, C3KO and serum addback mice at 2 days post-infection with 5x10^7 *L. donovani* stationary phase promastigote parasites (WT, n=5 males and 5 females; C3KO, n=5 males and 4 females; serum addback, n=5 males and 4 females). For bone marrow and lymph nodes, two were taken per mouse. Each point is an individual mouse tissue. Colour indicates mouse genetic background as shown in the key and is consistent across all panels. Mann-Whitney-Wilcoxon test with Bonferroni correction. Significance of adjusted p-value: * p<0.05; **p<0.01; *** p<0.001: **** p<0.0001. Boxplots show the median, 25th and 75th quartiles,

with whiskers representing all data-point within 1.5x Interquartile Range of the median; points above or below the whiskers represent outlying data values.

Males and females also showed differences in FP in specific tissues when compared directly. In WT mice, gut FP was higher in male mice (MW: p=0.00794), whilst in C3KO mice, females had a higher FP in bone marrow (MW: p=0.00855), liver (MW: p=0.0159) and lung (MW: p=0.0317) FP. In the serum addback mice, females also had higher bone marrow FP compared to males (MW: p=0.0343) (Figure 5.11).

Whilst these data are not consistent with the previous findings with regard to the liver, bone marrow or spleen, it does indicate that the effect of C3-deficiency is greater in male mice, given the differences in lymph nodes, gut and lung are observed only in males here. Serum reconstitution was not able to recapitulate the WT bottleneck in most tissues, and parasite dose may be an important factor in the impact of C3 on parasite invasion, given differences between these data and the initial experiment. Male mice consistently show a heavier reliance on complement-related processes in the dynamics of tissue colonisation, but the specifics of this effect differed between experiments.

5.2.5 Early *L. donovani* dissemination dynamics are more dependent on complement in male mice

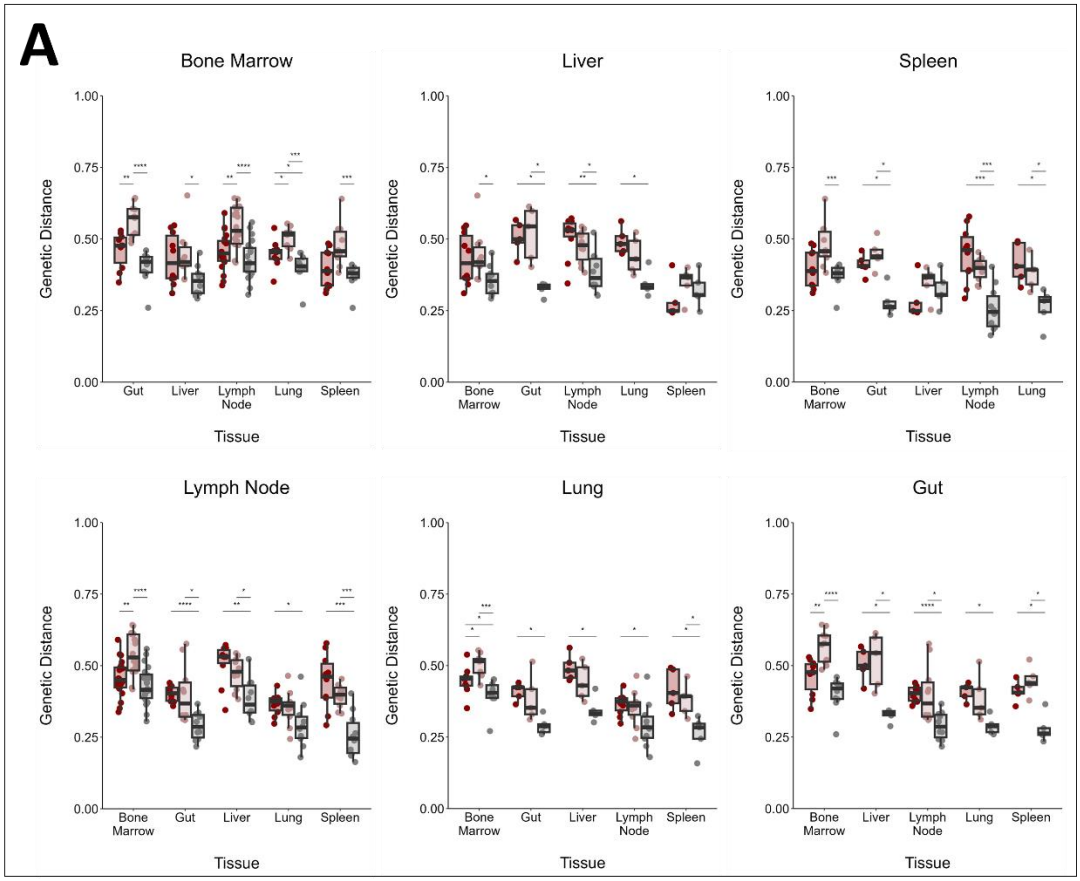
GD was calculated for the parasite populations within each tissue for the WT, C3KO and serum addback mice. Overall, C3KO mice showed few statistical differences to WT mice for amalgamated data, with only the lymph node to gut (MW: p=0.017), lymph node to liver (MW: p=0.04) and liver to lung (MW: p=0.017) showing that removal of complement resulted in reduced relatedness (higher GD).

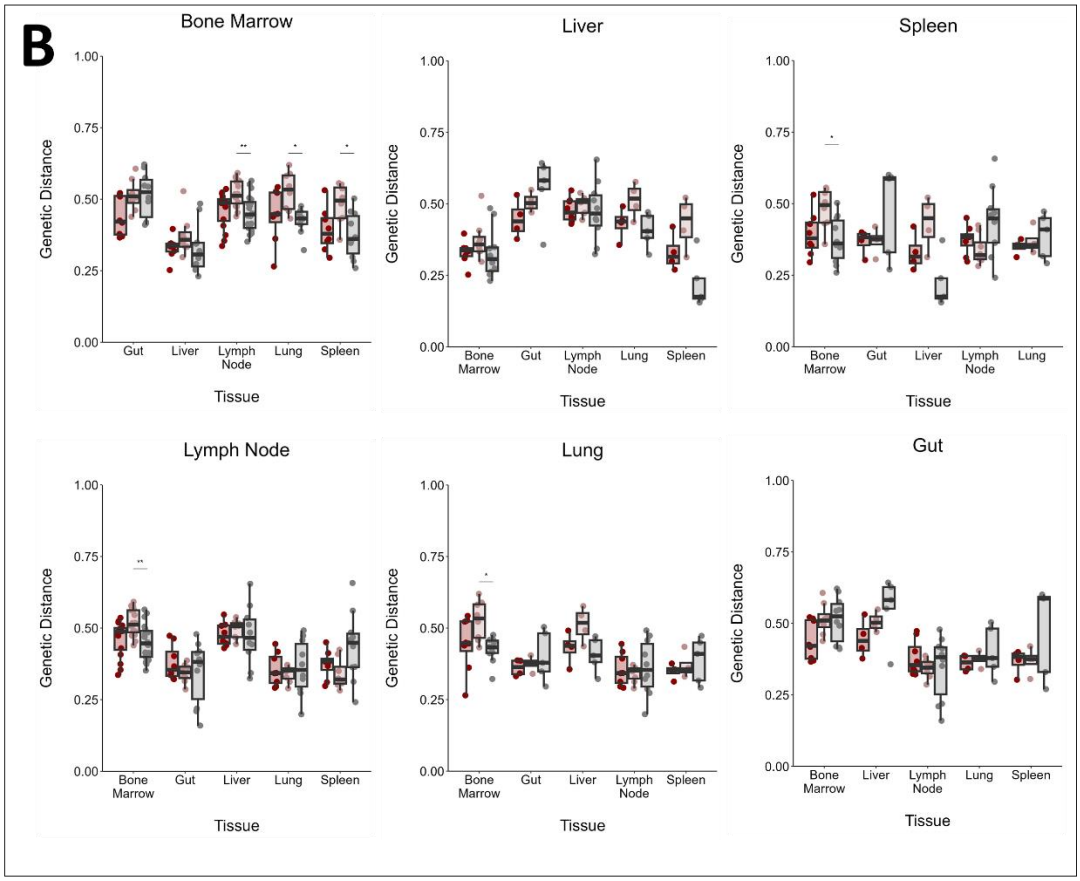
Interestingly, the addition of serum did appear to affect dissemination, with C3KO mice showing more relatedness in relation to the bone marrow and almost all other tissues, compared to the serum addback mice ($MW_{BM\text{-}GUT}$: p=0.003; $MW_{BM\text{-}LN}$: p=0.000177; $MW_{BM\text{-}LUNG}$: p=0.011; $MW_{BM\text{-}SPLEEN}$: p=0.002). In most bone marrow comparisons, along with the liver to lung comparison, a significant difference was also found between the serum addback and WT mice, even when no statistical

difference between C3KO and WT mice was seen. In all of these cases, the serum addback mice had less related tissue populations than the WT mice did.

When mice were split by sex, females showed very little effect of either complement absence or reconstitution with competent serum (Figure 5.12A). However, in male mice, multiple pairs of tissues showed a difference between C3KO and WT mice (Figure 5.12B), with consistently greater relatedness between tissues in WT mice; this is in contrast to the previous study, which noted the liver and spleen were more closely related in the C3KO mice (Figure 5.4A and B). The bone marrow was not as well-represented in these pairs as other tissues. Addition of serum to C3KO mice only caused a significant difference in specific tissues paired with bone marrow in males ($MW_{BM:GUT}$: $p<0.0001$; $MW_{BM:LN}$: $p=0.003$; $MW_{BM:LUNG}$: $p=0.034$), and in all of these cases, the serum addback had less closely related populations in the paired tissues compared to the un-reconstituted C3KO mice (Figure 5.15B).

Directly comparing male and female mice of each group indicated that tissue pairs with gut or lymph node were often more closely related in male WT mice compared to females for the same comparison ($MW_{BM\text{-}GUT}$: $p=0.00893$; $MW_{LIVER\text{-}GUT}$: $p=0.00794$; $MW_{LUNG\text{-}GUT}$: $p=0.0159$; $MW_{LIVER\text{-}LN}$: $p=0.0232$; $MW_{SPLEEN\text{-}LN}$: $p=0.00105$). C3KO mice differed only in comparisons between bone marrow and liver, with male mice showing less relatedness between these two tissues (MW: $p=0.0117$).





Group: B6.C3^{tm1Crr}
Serum Addback
C57BL/6J

Figure 5.12: Genetic distance analysis for comparisons between multiple tissues in complement-deficient mice (with or without serum 'addback') and wild-type C57BL/6J mice infected with promastigote *L. donovani* parasites at 2 days post-infection. A). B). Genetic distance for pairwise comparisons of each tissue in male (A) and female (B) WT, C3KO and serum addback mice at 2 days post-infection with 5x10^7 *L. donovani* stationary phase promastigote parasites (WT, n=5 males and 5 females; C3KO, n=5 males and 4 females; serum addback, n=5 males and 4 females). For bone marrow and lymph nodes, two were taken per mouse. Each data point represents a pairwise comparison between two tissues. As such, the reverse pairwise comparison, in the corresponding switched direction plot (i.e. the x-axis tissue is the title tissue and vice versa), is identical. Colour indicates mouse genetic background as shown in the key. Mann-Whitney-Wilcoxon test with Bonferroni correction. Significance of adjusted p-value: * $p<0.05$; **$p<0.01$; *** $p<0.001$: **** $p<0.0001$. Boxplots show the median, 25^{th} and 75^{th} quartiles, with whiskers representing all data-point within 1.5x Interquartile Range of the median; points above or below the whiskers represent outlying data values.

Collectively, the data from these experiments provides a suggestion that C3-deficiency impacts dissemination patterns more in male than female mice, but the exact nature of this effect remains unclear.

5.2.7 Neutrophil and macrophage numbers in the liver after *L. donovani* infection are not affected by C3-deficiency

Phagocyte recruitment, especially of neutrophils, is thought to be an important early event in *Leishmania* infections, and this process can be influenced by the complement system. The liver is a key early site of parasite colonisation, and therefore it was hypothesized that neutrophils in the liver may be associated with underlying differences in bottlenecks and dissemination dynamics in C3KO and WT mice. To assess how neutrophil and macrophage recruitment may relate to differences in colonisation, liver sections from male and female C3KO and WT mice, at 2 days p.i with *L. donovani* promastigotes, were stained for Ly6G and F4/80 to determine the proportions of neutrophils and macrophages respectively.

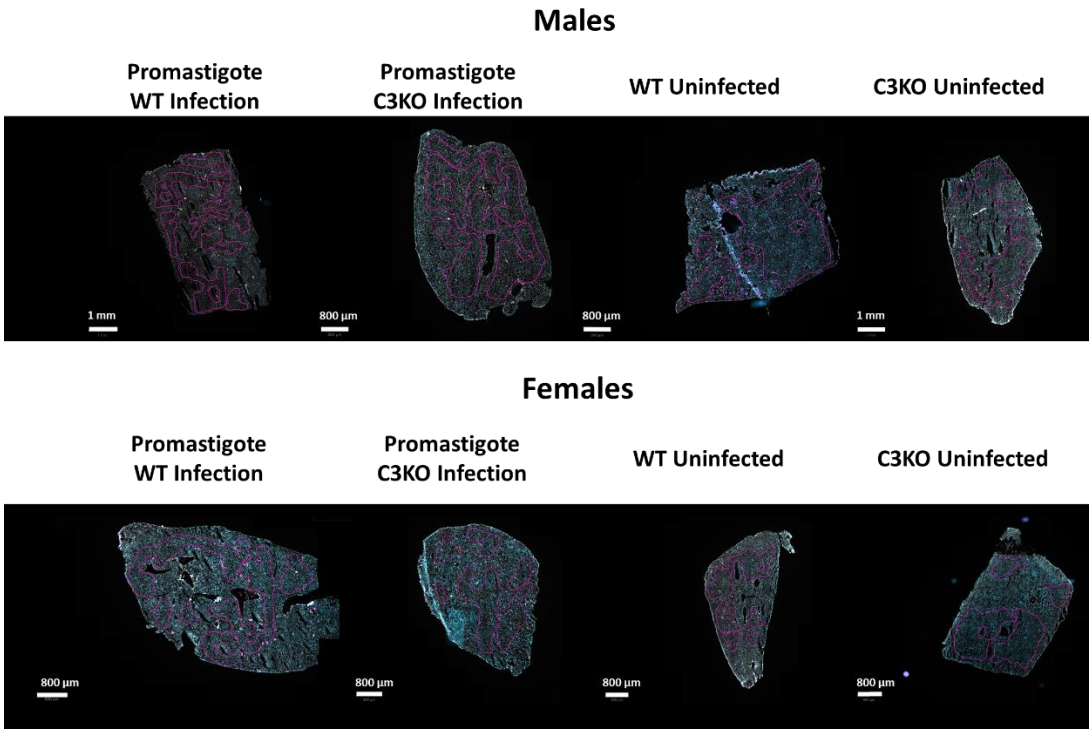
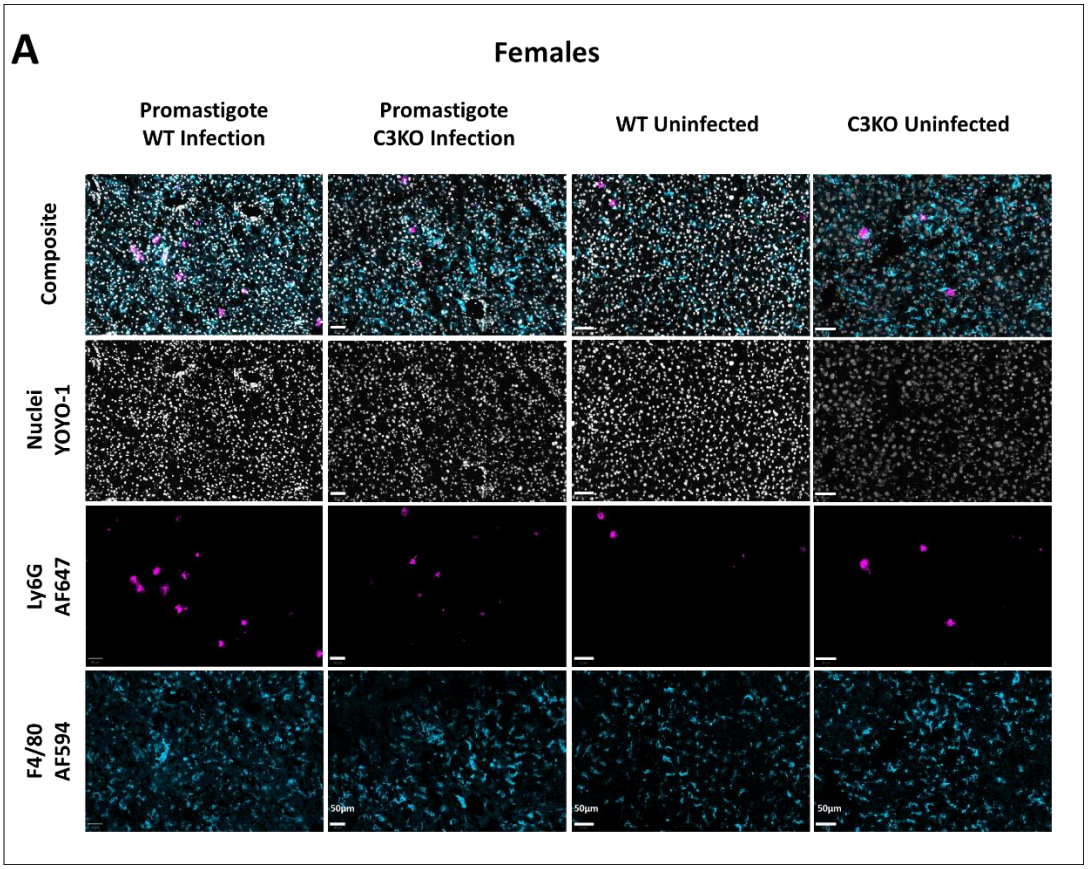


Figure 5.13: Exemplar images of whole liver sections stained for Ly6G and F4/80 to quantify neutrophils and macrophages. Liver sections from WT, C3KO and serum addback mice at 2 days post-infection with 5x10^7 *L. donovani* stationary phase promastigote parasites (WT, n=5 males and 5 females; C3KO, n=5 males and 4 females; serum addback, n=5 males and 4 females). Uninfected controls (WT, n=2 males and 2 females; C3KO, n=2 males and 2 females). Liver cryo-sections were stained using conjugated antibodies for Ly6G (AF647) (magenta) and F4/80 (AF594) (blue) to quantify the proportion of cells expressing each marker. Total cells were quantified using YOYO-1 nuclear staining (white). One representative image from each group is shown. Pink outlines indicate the regions quantified in each tissue using QuPath. Scale bars are shown in white at the bottom left of each image and measure either 800 µm or 1 mm as indicated. Images were produced in QuPath.

Figure 5.13 shows exemplar staining of whole liver sections for one male and one female mouse from each infection group, plus an uninfected control for each mouse line (n=2 males and 2 females for each). The quantified regions for each tissue are marked in pink, to exclude regions with obvious cutting artefacts. Figure 5.14 shows representative individual staining and composite images to identify macrophages and neutrophils.



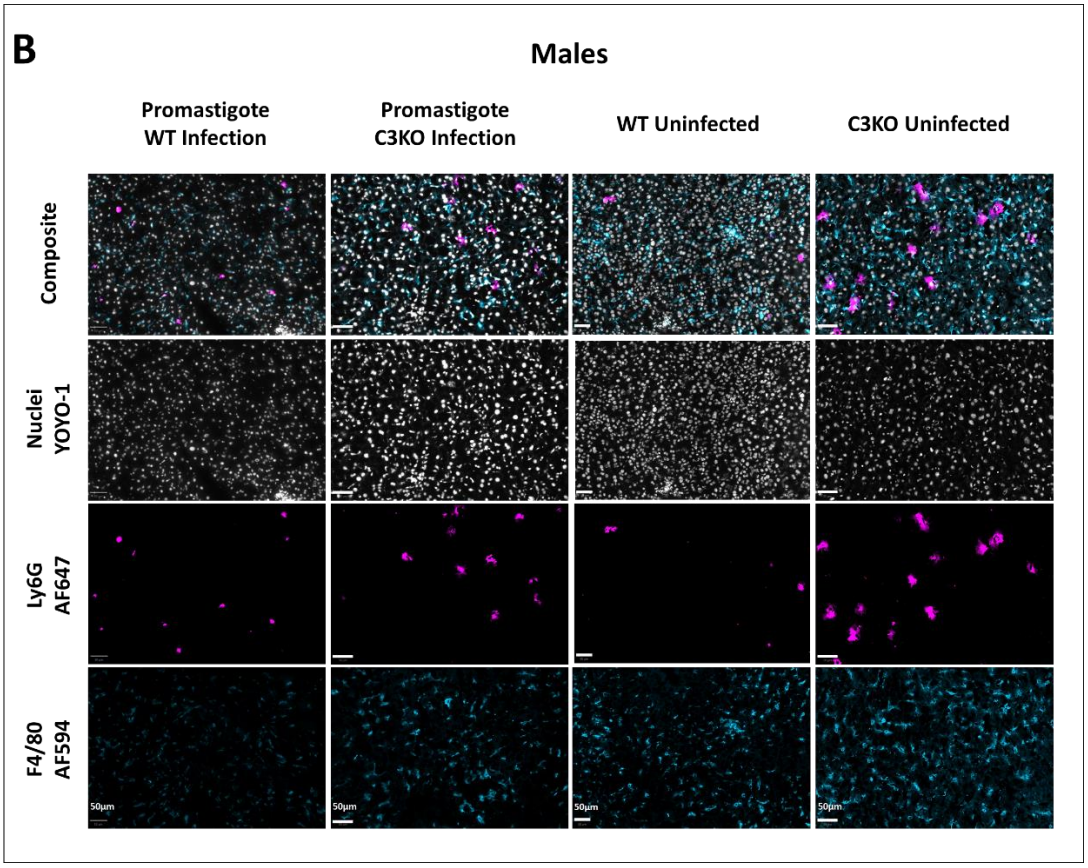


Figure 5.14: Exemplar staining of liver sections for Ly6G and F4/80 to quantify neutrophil and macrophage proportions in male and female C57BL/6J and B6124S-C3tm1Crr/J mice infected with *L. donovani* promastigotes at 2 days p.i. Liver sections from WT, C3KO and

serum addback mice at 2 days post-infection with 5x10^7 *L. donovani* stationary phase promastigote parasites (WT, n=5 males and 5 females; C3KO, n=5 males and 4 females; serum addback, n=5 males and 4 females). Uninfected controls (WT, n=2 males and 2 females; C3KO, n=2 males and 2 females). Liver cryo-sections were stained using conjugated antibodies for Ly6G (AF647) (magenta) and F4/80 (AF594) (blue) to quantify the proportion of cells expressing each marker. Total cells were quantified using YOYO-1 nuclear staining (white). A). Female mouse representative images of (from top) composite, nuclei, Ly6G and F4/80 staining. B). Male mouse representative images of (from top) composite, nuclei, Ly6G and F4/80 staining. One representative image from each group is shown. Scale bars in white at the bottom left of each image all equal 50 µm in length. Images were produced in QuPath.

To analyse the total cell number, individual nuclei within the selected region were quantified using automated QuPath cell detection software. The classification settings to detect cells showing positive staining for AF647 (Ly6G) and AF594 (F4/80) were set by comparing stained sections to isotype controls, as shown in Figure 5.15, to find the intensity settings to apply (further described in Methods). Negative cells are outlined in black. Two intensity levels were applied for each marker: intermediate staining and high staining, which are indicated in the classification images (Figure 5.15) as orange and red respectively. The mask shows the intensity of staining on a red scale, where a higher red intensity indicates greater staining positivity for the respective marker.

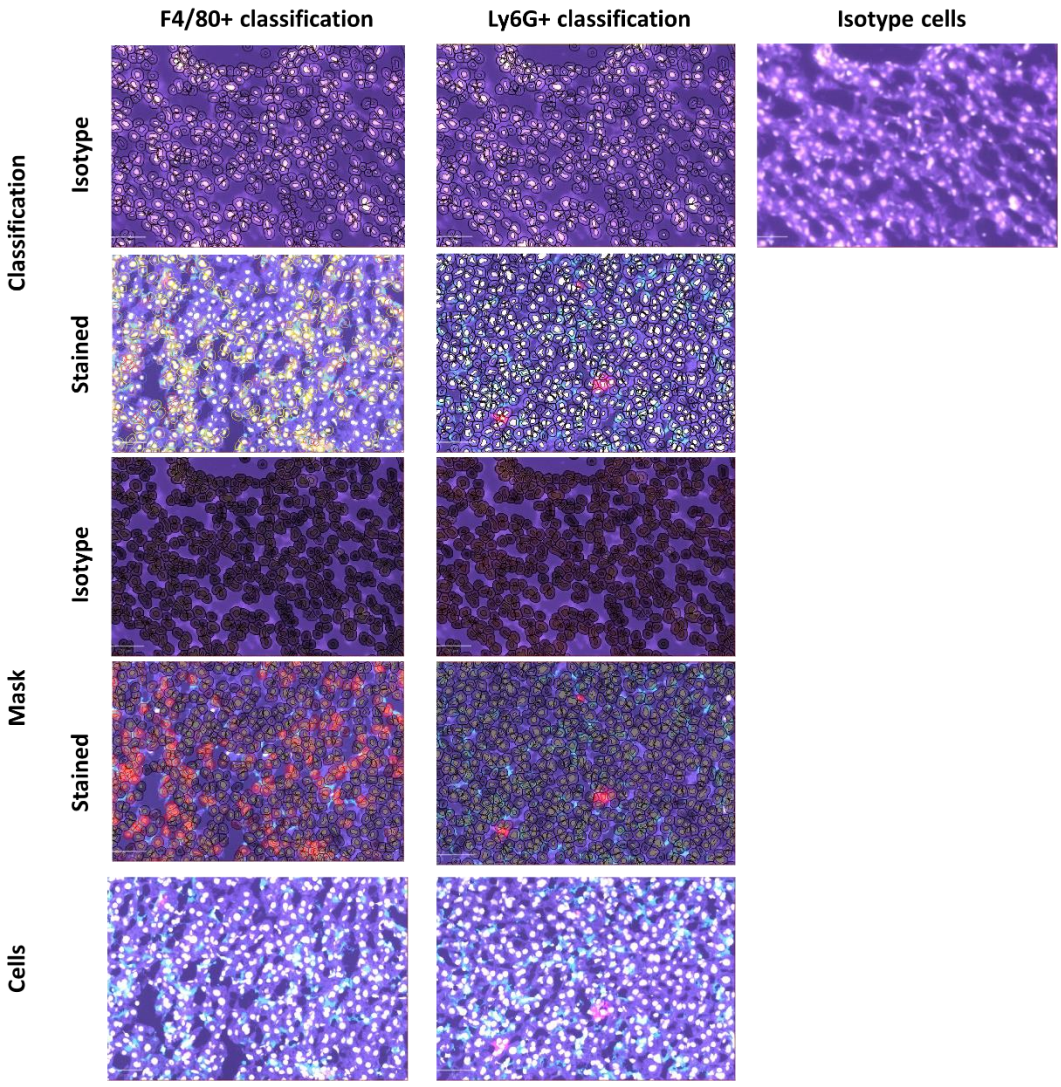


Figure 5.15: QuPath analysis protocol for classification of Ly6G and F4/80 positive cells in *L. donovani* infected mouse livers. Cell classification was performed in QuPath to define cells as positive for F4/80 (AF594) or Ly6G (AF647) or negative for both markers. The top image shows the isotype staining, which was used to define the classification settings. Cells were classified as positive if intensity of staining was above 1000 and as high-expressing if above 2000, for both fluorophores, with measurement based on Cytoplasm Mean for AF647 and Cytoplasm Max for AF594. Isotype livers showed negligible positive cells under this classification. In classification, negative cells are outlined in black, intermediate positive cells are in orange and highly positive cells are in red. Mask shows the intensity scale with red intensity indicating the level of expression of the relevant marker; black indicates no expression. Images generated in QuPath.

Using this classification method, the proportion of total cells which were classed as neutrophils was calculated for each liver section, with results shown in Figure 5.16.

It was found that there were no statistical differences between infected mice, either when amalgamated (Figure 5.16A) or split by sex (Figure 5.16B and C). However, there was a possible increase in the numbers of neutrophils in uninfected C3KO mice when amalgamated (MW: p=0.0286); this effect was not significant when split by sex (most likely due to the small sample sizes), but appears to be driven by the two male mice, which both had higher than expected neutrophils. Since the sample size was only two for the uninfected groups, this could be random variation in these specific mice and not consistent across a larger sample. Neutrophils in infected female WT mice were also noted to be significantly higher than in infected WT males, when directly compared (MW: p=0.0159).

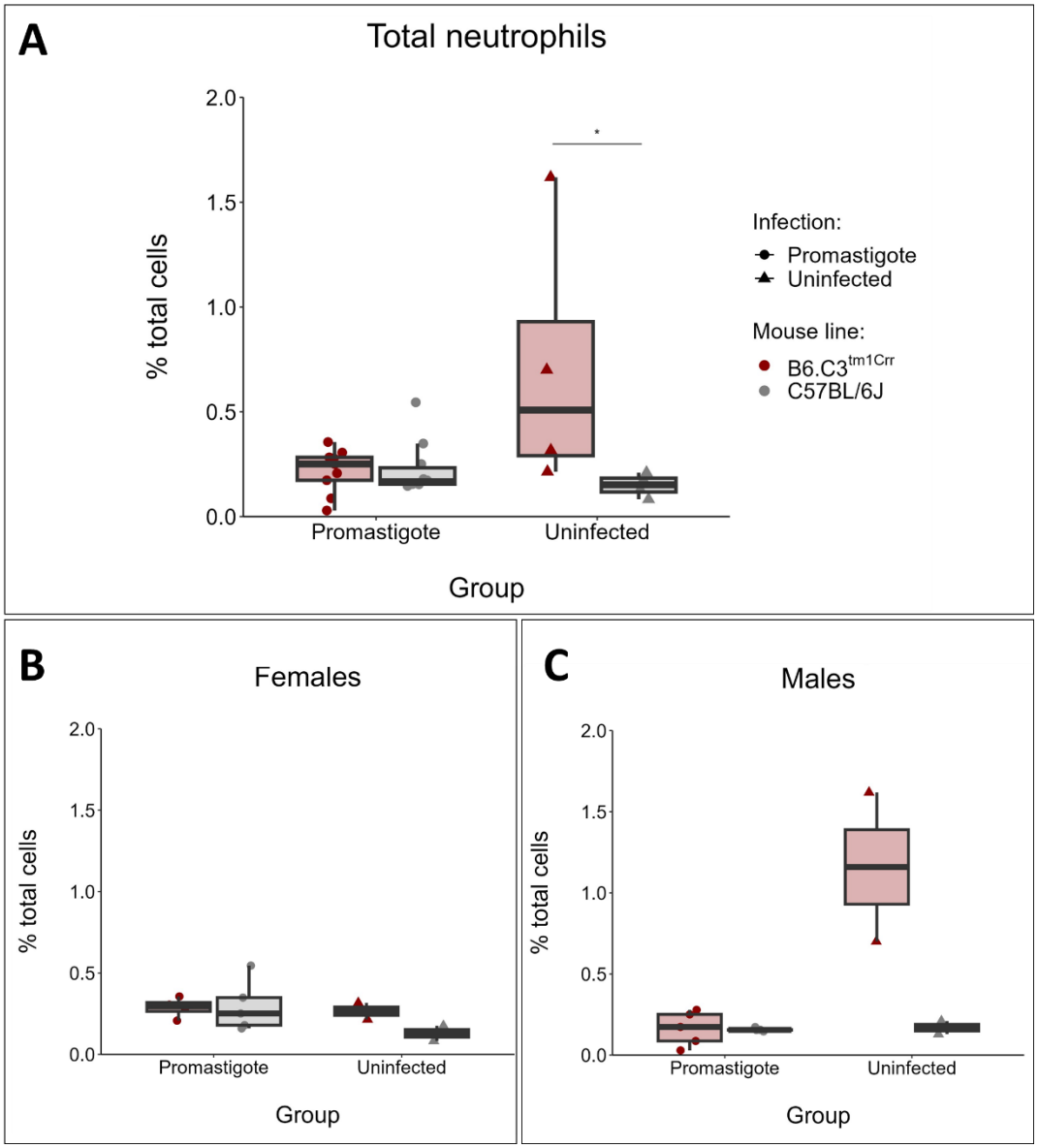


Figure 5.16: Total neutrophils as a percentage of total cells in livers of male and female C57BL/6J and B6124S-C3tm1Crr/J mice infected with *L. donovani* parasites at 2 days p.i. Liver sections from WT and C3KO mice at 2 days post-infection with 5x10^7 *L. donovani* stationary phase promastigote parasites (WT, n=5 males and 5 females; C3KO, n=5 males and 4 females). Uninfected controls (WT, n=2 males and 2 females; C3KO, n=2 males and 2 females). Liver cryo-sections were stained using conjugated antibodies for Ly6G (AF647) (magenta) and F4/80 (AF594) (blue) to quantify the proportion of cells expressing each marker, as a proportion of total nuclei. A). Total neutrophils (Ly6G positive cells) as a percentage of total cell nuclei, for mice of both sexes. B). Total neutrophils as a percentage of total cell nuclei for female mice. C). Total neutrophils as a percentage of total cell nuclei for male mice. Colours and shapes in key apply to all panels. Mann-Whitney-Wilcoxon test within each infection status group. Significance of adjusted p-value: * p<0.05; **p<0.01; *** p<0.001: **** p<0.0001. Boxplots show the median, 25th and 75th quartiles, with

whiskers representing all data-point within 1.5x Interquartile Range of the median; points above or below the whiskers represent outlying data values.

Neutrophils were then segregated into high and intermediate expression of Ly6G (Figure 5.17), which previous data in spleen and bone marrow has indicated corresponds to mature and immature neutrophils respectively (Deniset et al., 2017; Mackey et al., 2021). The difference in uninfected mice appeared to be the result of increased mature neutrophils (MW: p=0.0286), but this difference was again not significant when split by sex.

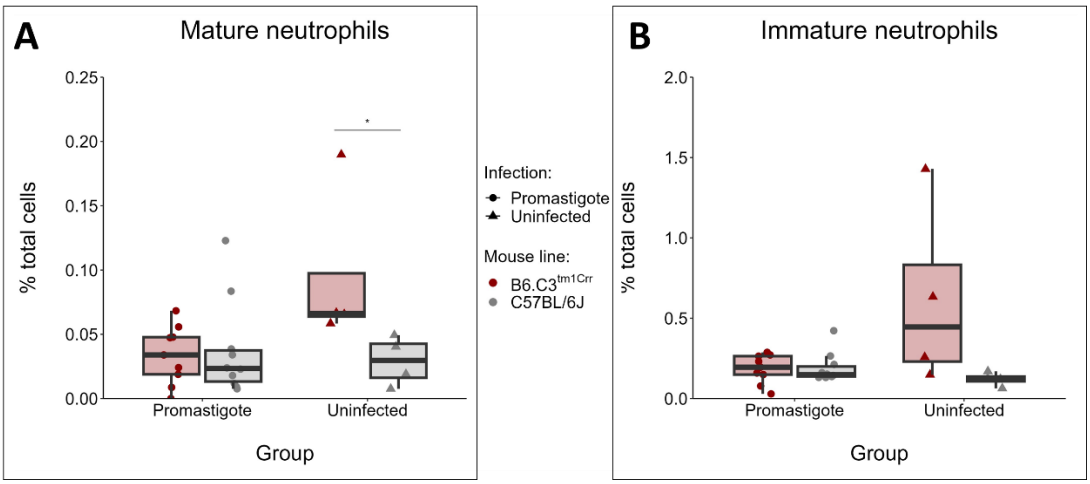


Figure 5.17: Total neutrophils expressing high or low-medium Ly6G as a percentage of total cells in livers of male and female C57BL/6J and B6124S-C3tm1Crr/J mice infected with *L. donovani* parasites at 2 days p.i. Liver sections from WT and C3KO mice at 2 days post-infection with 5x10^7 *L. donovani* stationary phase promastigote parasites (WT, n=5 males and 5 females; C3KO, n=5 males and 4 females). Uninfected controls (WT, n=2 males and 2 females; C3KO, n=2 males and 2 females). Liver cryo-sections were stained using conjugated antibodies for Ly6G (AF647) (magenta) and F4/80 (AF594) (blue) to quantify the proportion of cells expressing each marker, as a proportion of total nuclei. A). Percentage of total cells expressing high levels of Ly6G, indicating mature neutrophils. B). Percentage of total cells expressing intermediate levels of Ly6G, indicating immature neutrophils. Colours and shapes in key apply to all panels. Mann-Whitney-Wilcoxon test within each infection status group. Significance of adjusted p-value: * p<0.05; **p<0.01; *** p<0.001: **** p<0.0001. Boxplots show the median, 25th and 75th quartiles, with whiskers representing all data-point within 1.5x Interquartile Range of the median; points above or below the whiskers represent outlying data values.

The proportion of macrophages as a percentage of total cells is shown in Figure 5.18. Amalgamated data for all mice indicated no differences between C3KO and WT mice, regardless of infection status, and this was the same when mice were split by sex (Figure 5.18A and B). Again, infected WT male and female mice did show a difference, with female mice having higher macrophage numbers (MW: p= 0.0159), but no difference between male and female C3KO mice was found. F4/80^{high} resident and F4/80^{low} monocyte-derived macrophages were also assessed, with the thought that differences in these subtypes could be linked to bottleneck differences, but no difference between C3KO and WT mice was found.

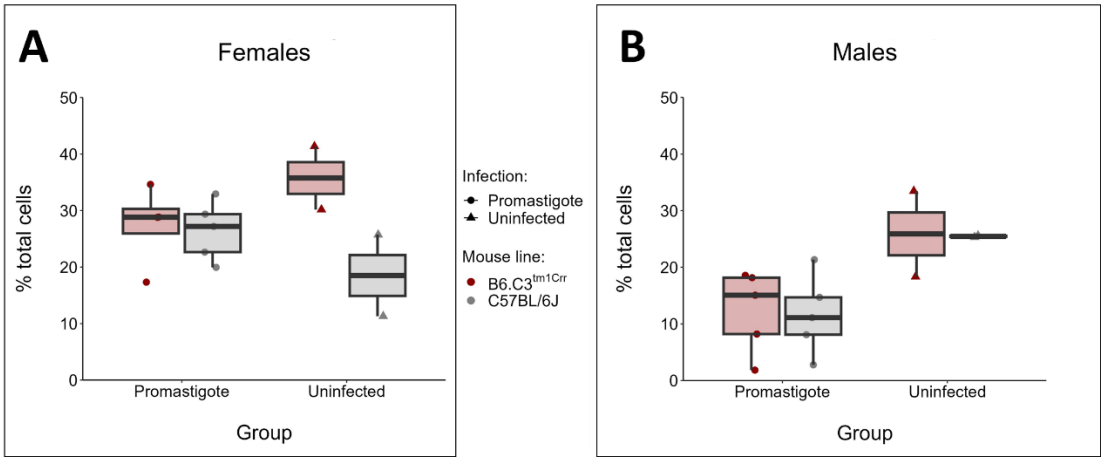


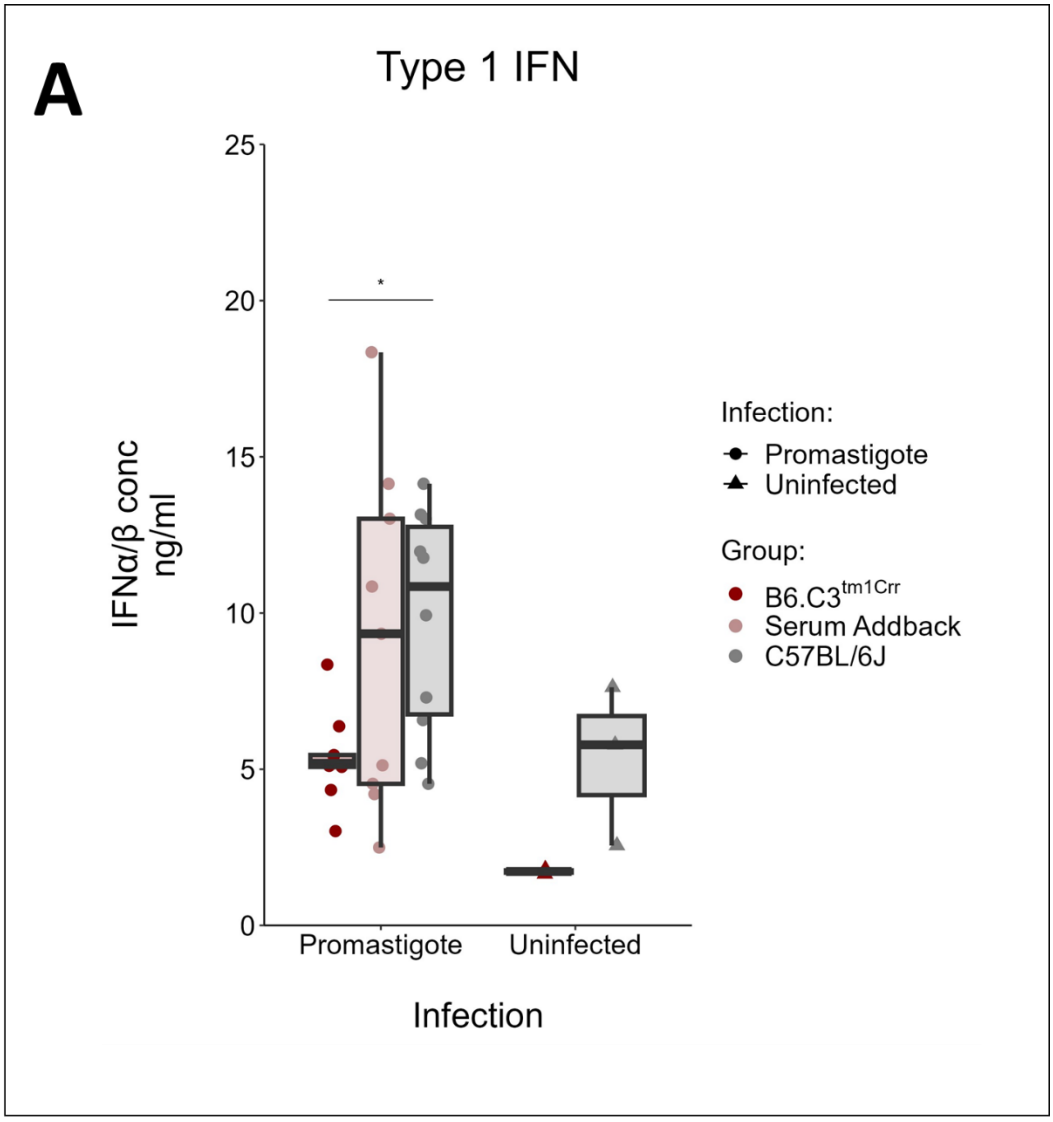
Figure 5.18: Total macrophages as a percentage of total cells in livers of male and female C57BL/6J and B6124S-C3tm1Crr/J mice infected with *L. donovani* parasites at 2 days p.i. Liver sections from WT and C3KO mice at 2 days post-infection with 5x10^7 *L. donovani* stationary phase promastigote parasites (WT, n=5 males and 5 females; C3KO, n=5 males and 4 females). Uninfected controls (WT, n=2 males and 2 females; C3KO, n=2 males and 2 females). Liver cryo-sections were stained using conjugated antibodies for Ly6G (AF647) (magenta) and F4/80 (AF594) (blue) to quantify the proportion of cells expressing each marker, as a proportion of total nuclei. A). Total macrophages as a percentage of total cell nuclei for female mice. B). Total macrophages as a percentage of total cell nuclei for male mice. Each data point indicates one individual mouse. Colours and shapes in key apply to all panels. Mann-Whitney-Wilcoxon test within each infection status group. Significance of adjusted p-value: * p<0.05; **p<0.01; *** p<0.001: **** p<0.0001. Boxplots show the median, 25th and 75th quartiles, with whiskers representing all data-point within 1.5x Interquartile Range of the median; points above or below the whiskers represent outlying data values.

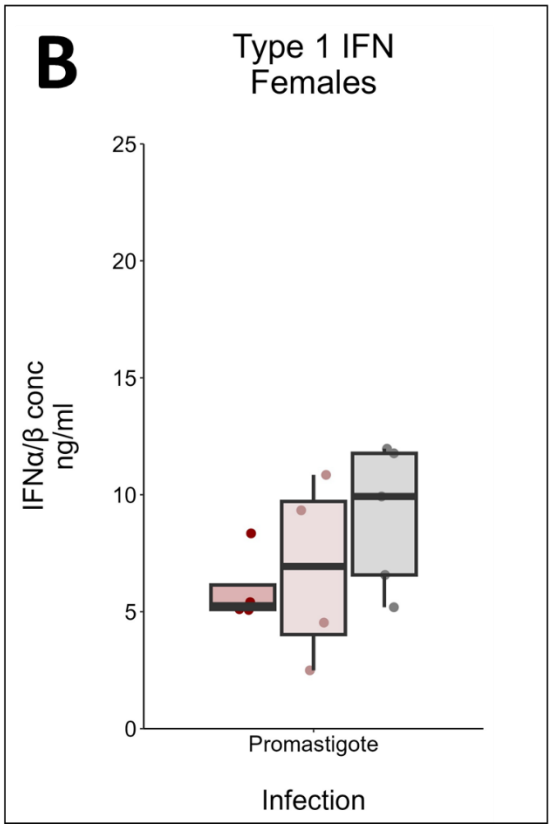
These findings indicate that there may be some differences in liver neutrophils in male mice lacking C3, but the small sample size means this could be a result of random diversity or related to specific conditions in the two mice assessed. No differences after infection with *L. donovani* were noted and no C3-related variation in macrophage numbers was found. Male and female WT mice did show some differences after infection with promastigote parasites, which were not present in C3KO mice, but overall differences in the abundance of either cell type appear unlikely to be the cause of the dissemination patterns observed.

5.2.8 Systemic IFNα/β in response to *L. donovani* infection is complement-dependent

Previous studies have identified links between C3 activation and Type 1 IFNs (Lebel et al., 2017; Posch et al., 2021; Tripathi et al., 2021; Wright et al., 2023) and Type 1 IFN is also noted to affect disease outcomes in leishmaniasis (Deng et al., 2023; Dias et al., 2019; Khouri et al., 2009; Mattner et al., 2000; Mattner et al., 2004). Therefore, we asked whether differences in Type 1 IFN production may be linked to early bottlenecks and dissemination dynamics in VL. To test this, sera from C3KO and WT mice infected with promastigote parasites were assessed at 2 days p.i, to determine serum concentration of IFNα/β.

Figure 5.19 shows the serum Type 1 IFN expression levels, with amalgamated data in Figure 5.19A, and sex split data in Figure 5.19B and C. A significant difference was found between infected C3KO mice and WT mice when amalgamated, with C3KO mice showing reduced IFN expression (MW: p=0.034); this difference was not present when analysing after splitting by sex, possibly due to the reduced sample sizes. The heterogeneity between mice was very different for C3KO mice compared to WT or those with serum reconstitution; C3KO mice had a very narrow range of IFNα/β expression, whilst all other groups showed far more variation within the group. No significant difference was found between uninfected WT and C3KO mice, when amalgamated or split by sex, or between male and female mice for each grouping (C3KO, WT or serum addback) after infection.


A
Type 1 IFN
IFNα/β conc
ng/ml
25
20
15
10
5
0
*
Promastigote
Uninfected
Infection
Infection:
Promastigote
Uninfected
Group:
B6.C3tm1Crr
Serum Addback
C57BL/6J


B
Type 1 IFN
Females
IFNα/β conc
ng/ml
25
20
15
10
5
0
Promastigote
Infection

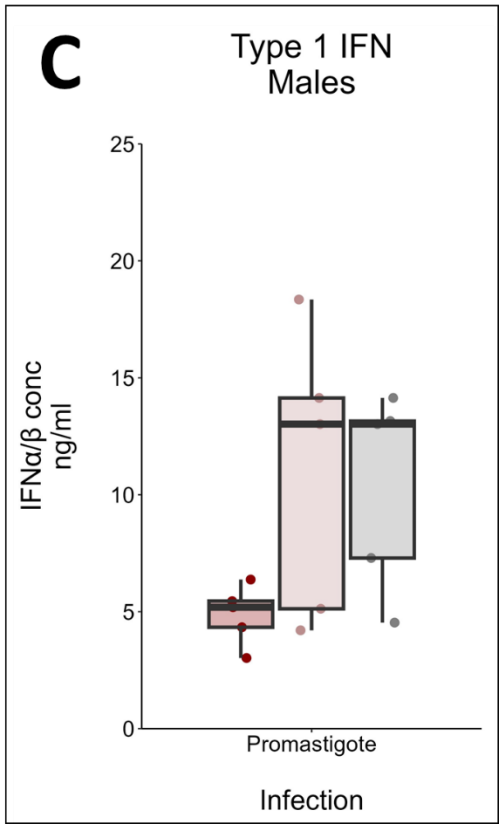

C
Type 1 IFN
Males
IFNα/β conc
ng/ml
25
20
15
10
5
0
Promastigote
Infection

Figure 5.19: Serum interferon-alpha/beta (IFN) expression in male and female C57BL/6J and B6124S-C3tm1Crr/J mice at 2 days post-infection infection with *L. donovani* parasites. Blood serum samples were taken from WT, C3KO and serum addback mice at 2 days post-infection with 5x10^7 *L. donovani* stationary phase promastigote parasites (WT, n=5 males and 5 females; C3KO, n=5 males and 4 females; serum addback, n=5 males and 4 females). Sera was assessed using an ELISA for IFNα/β. A). Overall IFN concentration (ng/ml) in sera for all mice. B). IFN concentration in sera from infected female mice. C). IFN concentration for sera from infected male mice. Colour scheme indicates mouse type and shape indicates infection; both are consistent across all panels. Each data point represents an individual mouse sample. Mann-Whitney-Wilcoxon test with Bonferroni correction within the infection status group. Significance of adjusted p-value: * p<0.05; **p<0.01; *** p<0.001: **** p<0.0001. Boxplots show the median, 25th and 75th quartiles, with whiskers representing all data-point within 1.5x Interquartile Range of the median; points above or below the whiskers represent outlying data values.

Overall, these findings are evidence that systemic IFNα/β signalling in response to *L. donovani* infection might be partially regulated by complement, but this is unlikely to be a key mechanism regulating sex differences in C3-mediated infection bottlenecks, due to the lack of sex-specific differences in Type 1 IFN induction.

5.3 Discussion:

This study sought to establish how the complement system impacts early tissue colonisation bottlenecks and dissemination patterns of *L. donovani* parasites in a mouse VL model. Our approach deployed a C3-deficient mouse; it is possible that different effects might be present when other complement components are absent. Together, the findings shown here indicate that C3 can play a role in colonisation and dissemination patterns early after *L. donovani* infection, but the exact nature of this impact appears to be variable and possibly tissue-specific. The key comonality across experiments was the greater impact of C3-loss in male mice specifically. Because of inconsistency across *in vivo* experiments, further study is required to understand the interaction of C3 with parasite dissemination.

The FP in the spleen, liver and bone marrow was observed to increase in mice infected with the higher parasite dose when complement was absent, and this effect was greater in male mice compared to females. However, in a second experiment, using a lower parasite dose, this was not the case, with no significant differences seen between WT and C3KO mice in terms of FP in these tissues. This could indicate that parasite dose is important in the activation of complement and the subsequent impacts on early colonisation; comparisons between experiments using FP data from 14 days p.i provide tentative support for this possibility. An increased level of complement activation may result in greater parasite death due to differences in numbers or activity of immune cells recruited to early sites of infection. For example, neutrophils may exhibit greater killing capacity under these conditions, or they may be less liable to apoptosis, which has been shown to increase parasite survival (Salei et al., 2017) and to differ between the sexes (Molloy et al., 2003). The doses given were only divergent by a factor of 2, so it would be prudent to assess this effect with a wider dose range in a single experiment, to determine if this effect is consistent or whether other differences between experiments may have played a role in the results seen here. Since FP in the liver specifically was noted to be far higher in the second experiment, it is also possible that the infection in this scenario was able to override the parasite-killing bottleneck imposed by complement in the liver, spleen and bone marrow, as was seen in the increased FP in these tissues in C3KO males in the initial experiment. In our second experiment, this could instead have resulted in a resource-limit acting to prevent further colonisation (e.g. limited host cell availability in these tissues).

An interesting finding was that colonisation of lymph nodes, gut and lung appeared to be greater in WT mice compared to C3KO mice in the second experiment. The FP in each of these tissues was higher in male WT mice, indicating that C3 may be advantageous for the uptake of parasites by host cells, or for parasite survival after phagocytosis, in these tissues specifically. Complement can be locally synthesised in all three tissues (Prat-Luri et al., 2022; Chaudhary et al., 2022; Detsika et al., 2024; West et al., 2023; Wu et al., 2024), with evidence for a role played by neutrophil recruitment to the lymph nodes in *L. mexicana* infection (Prat-Luri et al., 2022).

Neutrophil recruitment is an important early event in leishmaniasis, and these cells are noted to play a complex role in relation to parasite survival and disease outcomes. Locally synthesised C3 may result in increased chemotaxis of infected neutrophils to these tissues specifically; this impact seems to be present only in male mice, which could be due to differences in neutrophil trafficking. Male and female neutrophils may have differential receptor expression, or differ in their leishmanicidal activity. There is support for both tissue and sex divergence in neutrophil function and maturity (Blazkova et al., 2017; Kim et al., 2022; Lu et al., 2021; McGill et al., 2023; Molloy et al., 2003; Pokhrel et al., 2020; Xie et al., 2020), including complement receptor utilisation (Pokhrel et al., 2020) which supports this hypothesis, so further characterisation of these differences is necessary.

Other host cells may also differ in a similar way (Kovats, 2015; Liang Chi et al., 2024; Lakshmikanth et al., 2024; Masuda et al., 2018; Pujantell et al., 2023), whilst parasites may also express different surface proteins or alter their behaviour in other ways in response to differences in the tissue micro-environment. This may influence their uptake and the subsequent host cell responses. C3 opsonisation may be a key factor in host cell reprogramming by parasites in some tissues, but have little impact in others, due to the specific entry receptors used for phagocytosis (Ueno et al., 2012). This could result in the tissue-specific variation in colonisation observed here.

Understanding the baseline differences between WT and C3KO mice would be valuable, using flow cytometry or transcriptomics approaches to decipher how the cellular response differs both prior to and after infection. This could help establish whether specific cell types may be involved in complement-mediated bottlenecks and trafficking. Of particular relevance would be the macrophage subsets present in each tissue, as well as neutrophils, HSCs, pDCs and monocytes, all of which are known to act as permissive host cells for *Leishmania* parasites (Carlsen et al., 2015; Chang, 1981; Dirkx et al., 2022; Dirkx et al., 2024; Heyde et al., 2018; Karagiannis et al., 2023; Pearson and Steigbigel, 1981; Ribeiro-Gomes and Sacks, 2012; Romano et al., 2017; Soong, 2008; Terrazas et al., 2017). Tissue-resident macrophages have been noted to exhibit tissue-specific differences in complement receptor expression; one of note for its links to C3-mediated phagocytosis is Complement

Receptor of the Immunoglobulin superfamily (CRIg), which is highly expressed on Kupffer cells, as well as macrophages in multiple other tissues, but notably not in bone marrow or splenic macrophages. CRIg can be expressed as different splice isoforms which could further affect binding of parasites in a tissue-specific manner (Helmy et al., 2005). Variation in both receptor expression and inflammatory responses have been reported in human, rat and mouse macrophages (Becerra-Díaz et al., 2021; Ćuruvija et al., 2017; Marriott et al., 2006; Scotland et al., 2011), including in *in vivo* infection models (Nakanishi et al., 1989). *Leishmania* parasites can utilise many different routes to enter host cells, depending on various factors including serum composition, parasite surface markers, parasite life stage and parasite species (Ueno and Wilson, 2012). Understanding how receptors related to C3-mediated phagocytosis and also cell trafficking differ between the two mouse strains might enlighten us as to the differences in dissemination patterns observed.

Other differences between males and females may be linked to the male-specific role of C3 we have observed. Basal TLR expression is noted to be higher in female mice than in males, whilst males have been found to experience greater neutrophil influx immediately after bacterial infection (Scotland et al., 2011). Female mice may regulate the innate response to *Leishmania* using TLR pathways preferentially, rather than being reliant on C3-mediated signalling.

In vitro assessment determined that, whilst human serum could effectively kill parasites at higher concentrations, the presence of mouse complement proteins alone does not cause significant lysis of promastigote parasites. This is consistent with literature (Hoover et al., 1984; Mosser and Edelson, 1984; Späth et al., 2003) and indicates that a more complex mechanism than MAC-mediated lysis is involved in the differences between WT and C3KO tissue colonisation.

Our attempt to reconstitute C3 in the C3KO mice using the 'serum addback' approach described was designed with the thought that it may result somewhat in the recapitulation of the parasite dynamics we observed in WT mice. This was not what was observed, further confirming that C3 presence itself is not the direct cause of tissue bottlenecks. FP in the spleen, lymph nodes and bone marrow was decreased in the serum addback mice, compared to WT mice, and was not different to the C3KO mice in most tissues. There was also a marked difference between the

serum addback and both untreated mouse lines in terms of GD comparisons; the relatedness of populations was lower in many cases when serum was re-constituted, compared to both WT and C3KO mice. This effect was particularly striking in the bone marrow, showing that dissemination either to or from this tissue is affected by the addition of C3-containing serum. Potentially, the influx of C3 leads to dysregulation of the CXCR4/SDF-1 axis (Ratajczak et al., 2006), resulting in host cells which could normally traffic parasites from this tissue instead being sequestered in the bone marrow. The fact that the addback was unable to rescue the phenotype may be due to roles for local secretion or the cell-specific complosome in regulating parasite uptake or host cell migration (West and Kemper, 2023). Certain cell types and tissues may secrete complement proteins with different splice isoforms, which could be vital for uptake by macrophages in that tissue, or for tissue-specific trafficking from elsewhere (West and Kemper, 2023). However, it is also important to note that experimentally it is not possible to fully replace all C3KO serum with WT serum, to achieve an effective concentration which would fully mimic the WT system, *in vivo;* as such, it is also possible that the results are due to this experimental limitation on the circulating levels of C3 present in the 'serum addback' group.

Whilst the data indicates that mouse serum cannot kill parasites efficiently *in vitro*, it is also possible that in an *in vivo* setting, this process works more effectively, due to specific environmental conditions, such as pH or other cofactors present in the blood or tissues (West and Kemper, 2023); parasites could therefore be subject to rapid lysis by the high levels of complement proteins present post-injection. Repeating the experiment with additional control groups, which would include a WT mouse injected with WT serum and both C3KO and WT mice injected with C3KO serum, could help to address these questions by assessing how the levels of complement components and putative host cell trafficking differ between groups.

Macrophage and neutrophil numbers in the liver were assessed, due to the established role of these cells as early host cells for *Leishmania in vivo* (Carlsen et al., 2015; Laufs et al., 2002; Ribeiro-Gomes and Sacks, 2012) and the fact that the liver appears to be the key site of early colonisation. There were some differences between male and female WT mice after infection, but no significant differences

between infected C3KO and WT mice were found. Notably, uninfected C3KO mice had significantly greater neutrophils compared to WT mice, and this effect was apparently due to the high levels in male C3KO mice specifically. Neutrophil staining in the spleen was not shown here, but via a visual assessment appeared to show a very characteristic localisation of neutrophils to the marginal zone in uninfected C3KO male mice, but no other mice showed the same pattern. The sample size for this group consisted of only two mice, so it would be prudent to repeat this staining with a larger group of naïve mice, as well as including additional marginal zone markers to establish the exact location of these neutrophils. Immature splenic neutrophils have been shown to act as a mobilisable reservoir capable of trafficking from the spleen in stress conditions, with implicated roles in inflammatory disease and infection (Deniset et al., 2017; Mackey et al., 2019; Mackay et al., 2021). C3 may play a role in this process in male mice, which could underlie the unusual localisation seen here, and this may be relevant for pathological outcomes after *Leishmania* infection. Staining for neutrophils in the other tissues, particularly those which appeared to show a reliance on C3 for colonisation (the lymph nodes, gut and lung), would help in understanding whether neutrophils play a tissue-specific role in enabling or antagonising the colonisation process. Additional staining for other possible host cells, including markers for specific complement receptors and for parasites themselves, could also be valuable in determining the role of tissue-specific differences in parasite uptake. Likewise, measuring neutrophil abundance in the blood, and whether or not these cells are infected, could provide greater insight.

Previous studies have established sex-linked patterns in neutrophil maturity in the spleen, bone marrow and blood (Lu et al., 2021; McGill et al., 2023; Xie et al., 2020), whilst studies of DCs and HSCs have likewise found that sex can affect maturation and differentiation patterns (Kovats, 2015; Lakshmikanth et al., 2024; Liang Chi et al., 2024). Neutrophil clusters have been found to vary by both tissue and sex in single cell RNA-seq studies, with human female neutrophils having upregulation of Type 1 IFN stimulated genes and signalling via NFkB, whilst male neutrophils show evidence of higher mitochondrial metabolism (Gutpa et al., 2020). Xie et al (2020) assessed the sex differences in bone marrow neutrophil maturity compared to the spleen and blood, whilst Lu et al (2021) found female neutrophils have higher ECM

and autophagy-related gene expression, as well as genes related to myeloid differentiation. Splenic neutrophils were noted to mature more quickly than bone marrow neutrophils (Mackey et al., 2021) and in the spleen specifically, mature neutrophils were found to possess greater killing capacity when exposed to *S. pneumonia,* with evidence supporting extramedullary haematopoiesis and maturation of both neutrophils and monocytes within the splenic red pulp (Deniset et al., 2017). Sex differences in NET-osis (Blazkova et al., 2017; Lu et al., 2021) and apoptosis (Molloy et al., 2003; Xie et al., 2020) have also been reported, both of which have been shown to impact parasite survival (de Menezes et al., 2016; Gabriel et al., 2010; Salei et al., 2017), and both of which have been linked to complement activation (Mevorach et al., 1998; Salei et al., 2017). Meanwhile, hormonal changes appear to influence monocyte and DC responses to specific signalling pathways, including after stimulation with *M. tuberculosis* antigens (Lakshmikanth et al., 2024), and these hormones can affect macrophage receptor expression (Marriott et al., 2006; Scotland et al., 2011), along with a range of cytokines and chemokines, including CCL2 (Scotland et al., 2011).

Ideally, a future experiment would assess the phenotype of neutrophils, monocytes and DCs in the liver and other tissues after infection more conclusively, using flow cytometry, transcriptomics and metabolomics to establish whether different tissue-specific subsets are involved. Functional assessment of male versus female and C3KO versus WT host cells *in vitro* would also establish whether their ability to respond to C3-mediated signalling and to kill parasites differs. Determining the timing of neutrophil influx is another key follow-up experiment. These measurements were taken at 2 days p.i, but neutrophils are known to play a role at earlier time-points (Prat-Luri et al., 2022) – it may be that the timing of this staining misses key differences which underlie the colonisation patterns observed. *Ex vivo* assays comparing sex-specific serum effects on WT and C3KO neutrophils, DCs and macrophages from different tissues could assist in determining the C3-mediated mechanics of parasite phagocytosis and killing by these cells, as could profiling the chemokines present in each tissue.

Of note, it has been reported that complement deficiency can affect mouse behaviour, including by increasing aggression (Batista et al., 2014; Gatewood et al.,

2006; Parker et al., 2022; Shartz et al., 2023; Shi et al., 2015; Westacoot et al., 2021); it is possible that differences seen in these experiments may be wholly or partly related to males fighting and causing injuries to each other at a higher rate, resulting in altered baseline inflammation prior to infection. Thus, it would be advisable for future experiments to implement mitigating measures to reduce the risk of this affecting results (Weber et al., 2022).

Type 1 IFN secretion is known to play important roles in control of viral infections, with some studies noting that IFN can both worsen and ameliorate leishmaniasis infection (Deng et al., 2023). The effects appear to be dose dependent in that higher doses promote parasite survival, whilst lower doses are able to induce parasite killing (Dias et al., 2019; Khouri et al., 2009; Mattner et al., 2000; Mattner et al., 2004). We found a reduction in Type 1 IFN levels in the blood of C3KO mice after *L. donovani* infection, indicating that the induction of this pathway may be partially regulated by complement signalling. Whilst Lebel et al. (2017) noted that female pDCs from C3KO mice have higher IFNα expression after virus exposure, our study found no differences between male and female mice in terms of systemic Type 1 IFN expression. It is possible that tissue-specific differences were missed, and this would be valuable to investigate, as bottleneck effects appear to be tissue-specific in terms of the effect of C3-loss. It is also plausible that hormonal interaction with Type 1 IFN signalling affects the response in a manner which could account for the colonisation patterns above (Kovats, 2015; Lakshmikanth et al., 2024). Interferon regulatory factor 7 (IRF7) has been shown to play a key role in parasite killing by macrophages, in a tissue-specific way (Beattie et al., 2011; Phillips et al., 2010), and its expression in human DCs is impacted by testosterone levels (Lakshmikanth et al., 2024), whilst monocyte responses to Type 1 IFNs can also be hormonally regulated (Lakshmikanth et al., 2024). There are indications that in intracellular bacterial infections, high IFNβ may reduce responsiveness to IFNγ (McNab et al., 2014; McNab et al., 2015; Rayamajhi et al., 2010), which is required for parasite killing via oxidative burst, and might regulate host cell trafficking (Antonelli et al., 2010; McNab et al., 2015). pDC secretion of IFNα has been reported to be induced by NET production from neutrophils, inducing a feedback loop which can then promote greater NET-osis (Garcia-Romo et al., 2011). NETs have been shown to be beneficial for *L. donovani* uptake (de Menezes et al., 2016; Gabriel et al., 2010); the higher

IFNα/β in WT mice might therefore increase parasite uptake in this manner in the lymph node, gut or lung. Interferon-stimulated gene 15 (ISG15) has been noted to increase in DCs after infection with *L. braziliensis* (Vargas-Inchaustegui et al., 2009) and to worsen tuberculosis inflammation (Fernandes dos Santos et al., 2018); increased signalling of this gene pathway could be likewise be relevant to *Leishmania* infection in these tissues specifically.

Whilst blood flow is key for early parasite influx to various tissues after IV infection, it alone cannot explain the differences observed here, as we would assume similar flow routes would occur in both mouse models; further, it also cannot account for the GD dissimilarities observed. Therefore, trafficking via host cells is likely to play some role in differences observed between the C3KO and WT mice. One proposal for the results observed here is that male mice require both local C3 secretion and C3-opsonised parasite uptake to trigger colonisation of the gut, lung and lymph nodes. Tissue-resident macrophages in these organs may express specific receptors such as CRIg, which rely on C3 for parasite uptake, and expression levels may differ between males and females. Infection with promastigote parasites may then trigger chemokine signalling and secretion of C3 in these tissues, which induces neutrophils, DCs or other host cells from the spleen and liver of male mice to traffic to these tissues. If only cells which are infected with C3-opsonised parasites are capable of this, it may explain why both C3KO mice and amastigote-infected mice show little colonisation of these tissues (see Chapter 4); whilst not established for *L. donovani*, studies in *L. mexicana* indicate that amastigotes are not opsonised by C3 (Peters et al., 1995). Differences in neutrophil maturity, uptake and killing capacity, as well as receptor expression on tissue-resident macrophages, neutrophils and monocytes could explain why female mice do not show the same effect of C3 loss. The serum addback mice could not recapitulate the WT colonisation of these tissues as they could induce only C3-opsonised parasite uptake and not the local C3 secretion, and C3KO mice may also have downregulated CRIg or other complement receptors. Local complement secretion may rely on different isoforms to trigger directed host cell migration (West and Kemper, 2023). Hormonal influences and interactions with C3, Type 1 IFNs, host cell receptors and other signalling triggered by complement are likely to be involved in the tissue-specific differences we have found. Figure 5.20 summarises this proposed explanation. A key experiment to test

this model requires infection of C3KO mice with either promastigote or amastigote parasites, to compare responses to each, along with a repeat study in WT mice using both parasite life stages; further work in vitro and *in vivo* could assess whether specific parasite surface markers are required in conjunction with C3 opsonisation, and could investigate which host cell receptors may be playing the key roles for parasite uptake in different tissues. Taking tissues at the 2 day time-point and also at an earlier stage after infection could provide more insight into these early colonisation events, whilst gonad-ectomy or cross-hormone treatments could shed further light on sex differences in C3 responses.

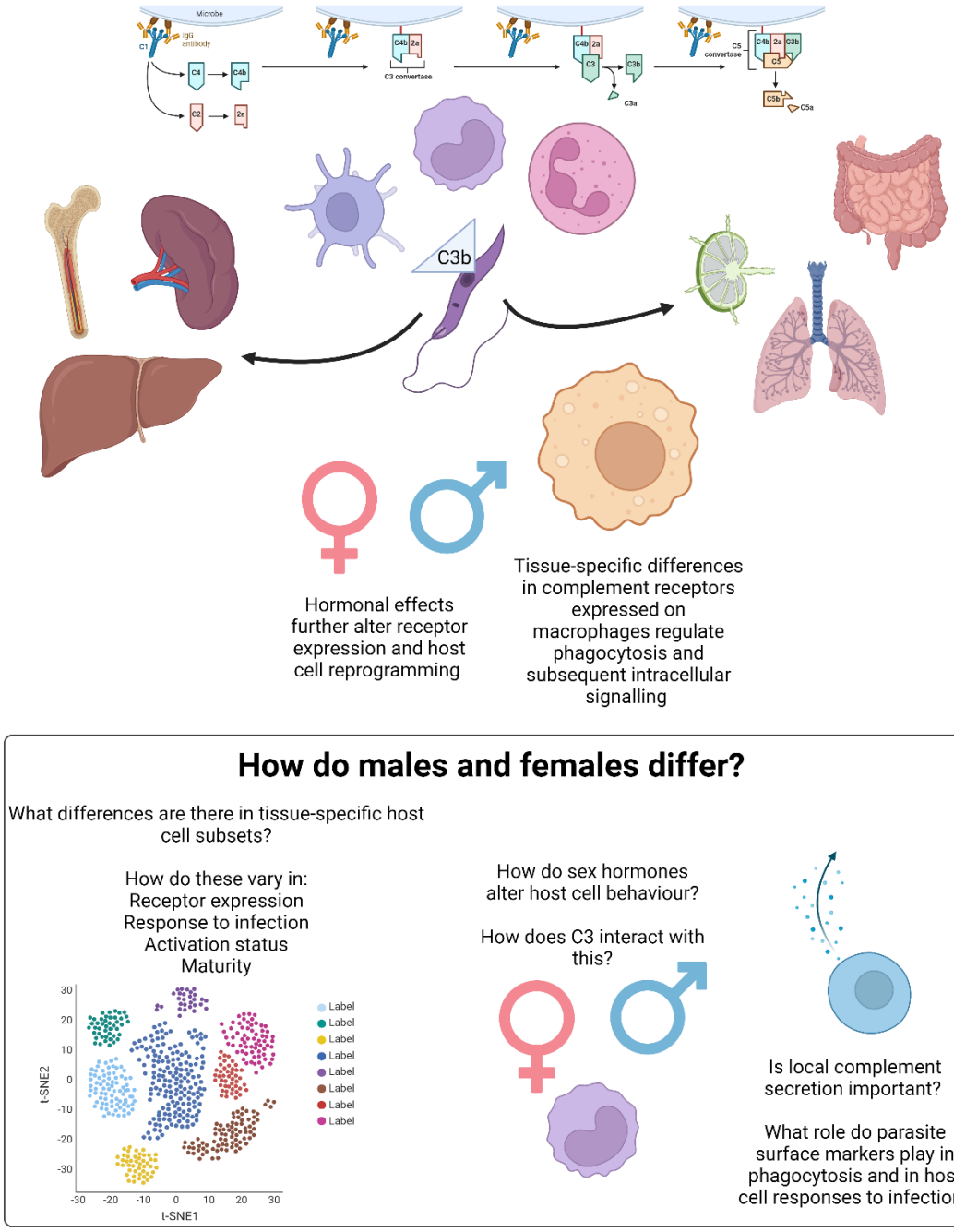


Figure 5.20: Understanding the role of complement in tissue colonisation by *L. donovani.* Schematic illustrates the proposed explanation for tissue specific effects of C3 on early colonisation in C57BL/6J mice infected with *L. donovani* parasites and some key questions to address. The impact of C3-loss appears to be sex linked, with male mice exhibiting a greater reliance on C3-mediated processes to regulate parasite dissemination. This is hypothesised to be a result of differences in tissue-resident macrophage phenotypes and subsets, along with recruitment of host cells relying on local C3 secretion for phagocytosis of parasites and host cell trafficking to the lymph nodes, gut and lung from the spleen and

liver. Hormonal regulation is likely to play a role in these sex-linked host-parasite interaction differences. Figure created in Biorender.

Understanding how the early effects of complement on parasite colonisation might affect the adaptive immune response would be valuable to establish; multiple studies have found that complement plays roles in B cell activation and cytokine secretion (Carroll, 2008; Carroll et al., 2004; Fisher and Hugli, 1997; O'Neil et al., 1988), and in the production of memory T cells (Kim et al., 2004; Fang et al., 2007), which has implications for vaccines and for predicting disease impact in endemic regions. Parsing out how complement-parasite dynamics in early infection versus the systemic results of complement loss could be possible by using Cobra Venom Factor to deplete complement in the short term, or by using antibodies to block key complement receptors. A mouse knock-out model of other complement components, such as C5, could also help to elucidate which aspects of the cascade are involved, which may further enlighten us as to possible mechanisms.

Likewise, investigations into the role complement plays in the skin, both in early response to infection and in skin pathologies such as PKDL, could help understand the mechanics of skin-parasite interactions and inform possible treatment and prevention strategies. Evidence shows parasites can home to the skin following mouse infection with *L. donovani* (Doehl et al., 2017; Doehl et al., 2021) but the timing at which this occurs is not established; it is possible complement could affect trafficking to and colonisation of the skin. Many inflammatory skin conditions show evidence of complement involvement (Giang et al., 2018), whilst sand flies are reported to secrete complement-inactivating compounds on biting (Cavalcante et al., 2003; Ferreira et al., 2016).

In summary, the findings presented here provide some tentative evidence for tissue-specific roles for complement in regulating the initial colonisation of tissue by *L. donovani* in a murine VL model, but further study is necessary given the inconsistency in the data generated. The role of C3 appears to be dependent on sex, with male mice showing a greater reliance on C3-mediated processes to regulate colonisation and dissemination. However, the exact impact of C3-deficiency on dissemination of *L. donovani* remains unclear. Liver neutrophil recruitment was not

associated with the dissemination of parasites, and systemic Type 1 IFN levels were likewise unable to account for the sex-linked differences observed. The interaction of sex hormones with the innate immune system requires further investigation. As yet, the underlying mechanics of early bottlenecks in our VL model are not established, but deciphering these could have important implications for understanding sex differences in innate responses to *Leishmania* infections.

Chapter 6: Tissue-specific population dynamics in a VL re-infection model

6.1 Introduction:

The adaptive immune system has evolved as a protective mechanism to both fight current infections and prevent re-infection with previously encountered pathogens. However, many pathogens have evolved strategies to evade both detection and destruction, including *Leishmania* parasites. Understanding how the adaptive immune system acts to prevent initial disease and when exposed to a secondary infection is key to the development of effective immunotherapies and vaccination strategies.

Early re-infection studies indicated the importance of T cells in both resolution of initial VL infection and in the response to re-infection, with BALB/c mice re-infected with *L. donovani* after 18 weeks showing an uptick in CD8+ T cells in the liver; these were highly concentrated around granulomas and were greater than the numbers seen in the primary infection (McElrich et al., 1988). Carter et al (1989) showed that the liver is able to resist re-infection, with reduced parasite burdens in BALB/c mice previously infected then treated to clear infection, compared to mice experiencing a first infection with *L. donovani*. However, this was only the case for the liver, with spleen and bone marrow burdens being unaffected, indicating the possibility of tissue-specific immunity. The effect seen was also possibly a result of drug persistence, or partially related to the removal of the previous parasite population, raising the question of whether parasite persistence might alter this protective effect. More recently, Ong et al (2020) used bioluminescence to assess parasite burdens after re-infection with *L. donovani*, demonstrating that BALB/c mice were resistant to re-infection to the limit of parasite detection via this system.

Re-infection in cutaneous leishmaniasis is more well-studied, with investigations into *L. major* indicating that skin-resident CD4+ memory T cells, and recruitment of both monocytes and circulating T cells specific to *L. major,* are crucial in preventing establishment of a secondary infection; these T cells activate the monocytes to instigate parasite killing (Glennie et al., 2015; Glennie et al., 2017). Peters et al (2014) noted that recruitment of Ly6C+ effector T cells was necessary for rapid

recruitment of IFNγ secreting CD4+ T cells to the site of re-infection, and that these cells were maintained only in the presence of persistent parasites. Re-infection with *L. major* has been shown to trigger re-activation of initial infection, and this effect appears to rely on an influx of CD4+CD25+ Tregs; chronically-infected mice given Tregs without a second infection experienced relapse of initial disease and induction of IFNγ and IL-10 (Mendez et al., 2004). Inflammatory monocytes have also been indicated to play a role in killing of secondary *L. major* parasites, whilst in primary infection they enabled parasite replication (Romano et al., 2017).

CD8+ T cells are also important in CL immunity. Early studies reported that CD8+ T cells were necessary for IFNγ secretion in the spleen and liver upon re-infection with *L. major* promastigotes, and that this cell population increases in response to re-challenge (Müller et al., 1993; Müller et al., 1994). Recruitment of CD8+ T cells to the site of *L. major* re-infection can promote healing and reduce parasite burdens via IFNγ secretion (Belkaid et al., 2002).

A greater understanding of how chronic *Leishmania* infection mediates the response to secondary infection is thus vital for the development of an effective vaccine against VL, as well as being of interest for investigations into the fundamental immunology of parasite interaction with the adaptive immune system at a tissue level. This work therefore aimed to assess whether colonisation dynamics differ between primary and secondary infections at a tissue level. We then further sought to assess the impact of a secondary infection on a pre-existing chronic infection. Additionally, the ability of visceral parasites to traffic to and colonise the skin was studied using two different *L. donovani* lines, providing a basis for understanding skin homing in a murine PKDL model.

6.2 Results:

6.2.1 Establishing a re-infection VL model

To understand how primary and secondary *L. donovani* infections, and the immune response to each, affect each other, a re-infection model was established using two distinct LV9-derived parasite lines: the previously described barcoded pool (Chapter 3) and the REH-*L. donovani* line, which expresses a red-shifted luciferase gene (see

Methods; Branchini, 2010; Soysa et al., 2015). This allowed discrimination between the prior and subsequent infection, as REH parasites do not contain barcodes, but can be viewed using the IVIS in both live mice and in tissues *ex vivo* by incubating with luciferin. Figure 6.1 summarises the experimental set up used, in order to examine the tissue-specific dynamics of primary and secondary infections.

C57BL/6J mice were infected with 3x10^7 amastigote REH-*L. donovani*; in the initial experiment, both male and female mice were used, whilst in the repeat, only females were included. A subset of mice were re-infected after 9-10 weeks, with barcoded LV9 T7 cas9 *L. donovani*, using 4-5x10^7 stationary phase promastigotes via tail IV. Mice which had previously not been infected with any parasites were also infected at this point, with one group receiving each of the parasite lines. Mice were imaged throughout the experiment using the IVIS to monitor REH-*L. donovani* parasite burdens. Mice were then killed two weeks after the promastigote infections (11-12 weeks post-initial infection), with tissues harvested as previously described.

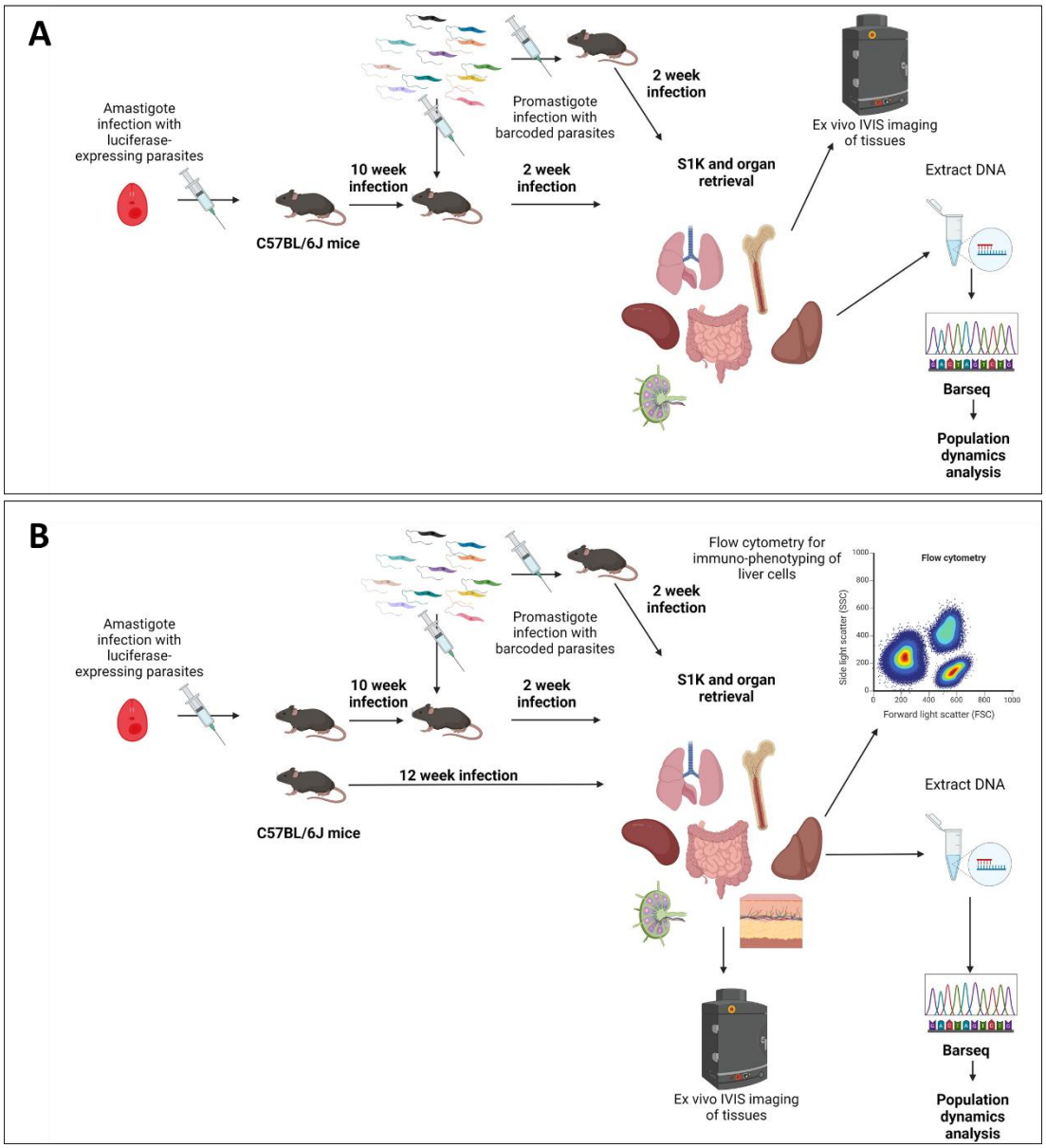


Figure 6.1: Schematic of experimental plan to investigate the population dynamics of *L. donovani* in a re-infection model using C57BL/6J mice. A). Initial experiment set up. n= 5 males and 5 females per group. B). Repeat experiment set up. n= 4 females for 2 week infection and for 10 week followed by 2 week infection, and n=3 females for 12 week single infection. C57BL/6J mice were infected with REH-*L. donovani* amastigotes, expressing a red-shifted luciferase gene. A subset of mice were re-infected after 9-10 weeks, with the barcoded LV9 T7 cas9 *L. donovani* parasite library, using 4-5x10^7 stationary phase promastigotes via tail IV. Mice which had previously not been infected with any parasites were also infected with promastigotes at this point. Mice were imaged throughout the experiment using the IVIS imaging system to monitor REH-*L. donovani* parasite burdens. Mice were then killed two weeks after the promastigote infections (11-12 weeks post-initial infection). Tissues were taken from each mouse, incubated with 30 mg/ml luciferin, then imaged *ex vivo* using the IVIS imaging system. Tissues were subsequently flash frozen on

dry ice, then stored at -80 °C before subsequent DNA extraction, Illumina sequencing, Barseq analysis and population dynamics studies, alongside a sample of the input parasite pool from each infection. Tissues taken consisted of liver, spleen, two lymph nodes and two each of tibia plus femur. Bone marrow was extracted by flushing and stored at -20 °C for DNA extraction. Skin from the rump of the mice (for B only) was removed and imaged *ex vivo*, then five randomly selected patches were biopsied and frozen for DNA extraction. Liver cells from the second experiment only (B) were used for flow cytometry analysis. Figure created in Biorender.

6.2.2 Parasites are not cleared from tissues after 10 to 12 weeks of infection

The trajectory of infection was measured using IVIS imaging of mice infected with REH-*L. donovani,* to check for parasite establishment and to observe how persistent parasites react to re-infection.

In amastigote-infected mice, with and without a subsequent infection, the whole body parasite burden did not differ after re-infection with barcoded promastigote parasites. No significant difference was found between re-infected and single infection mice, indicating that re-infection does not result in a detectable alteration in overall burden of latent parasites. In a separate infection male and female mice were imaged prior to and after re-infection. There were no significant differences in whole body parasite burden between male and female mice as a consequence of re-infection.

To investigate tissue-specific parasite burdens, organs from all REH-infected mice were imaged *ex vivo* after mice were killed. Exemplar images from mice infected for 12 weeks with amastigote REH-*L. donovani* (Figure 6.2A), with and without subsequent infection with barcoded parasites, are shown in Figure 6.2B. The images show clear evidence of colonisation by amastigotes, which can maintain a persistent presence in all tissues up to 12 weeks after initial infection.

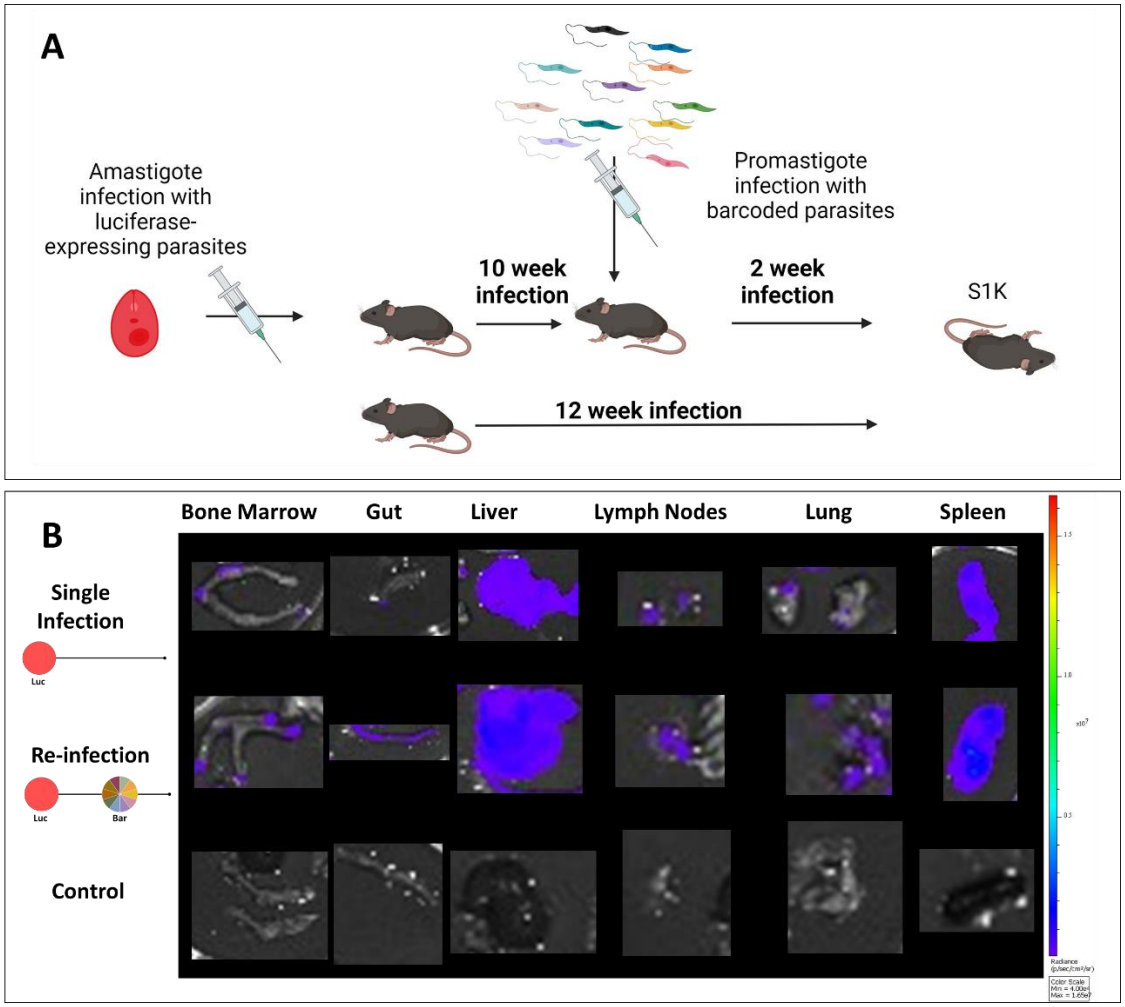


Figure 6.2: Exemplar *ex vivo* imaging of tissue parasite burden of initial amastigote infection in female C57BL/6J mice re-infected with promastigote *L. donovani* parasites after a prior amastigote infection. A). Schematic showing subset of mice analysed. B). Exemplar tissues after *ex vivo* IVIS imaging from mouse from the single infection group versus one mouse from the re-infection group. A control which had no REH-*L. donovani* infection is also shown. For single infection mice, female C57BL/6J mice were infected with 3x10^7 amastigote REH-*L. donovani* for 12 weeks (n=3 females) via intravenous tail injection. For re-infection mice, the same protocol was followed but female mice were re-infected with 5x10^7 stationary phase promastigote barcoded *L. donovani* (n=4 females) on day 70 or day 61 (n=5 females). Scale indicates radiance. Panel A created in Biorender.

When radiance was measured for each tissue of each mouse (Figure 6.3A), there was no detectable difference between re-infected and single infection mice in any tissue, but, when grouped, all infected tissues had significantly increased total flux compared to control mice which were not infected with REH-*L. donovani*. When male and female mice were compared, males were observed to have greater flux

than female mice in the gut (MW: p=0.042) but no other differences in any tissues were found (Figure 6.3B).

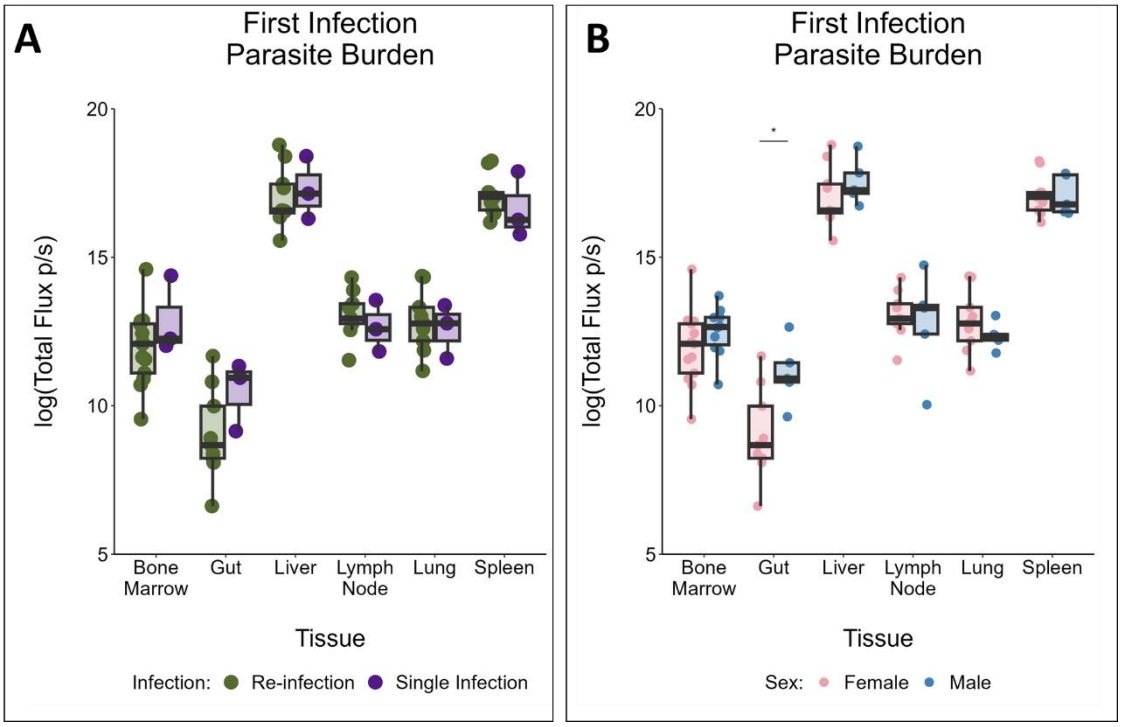


Figure 6.3: Tissue parasite burden, quantified by *ex vivo* imaging, of initial amastigote infection in C57BL/6J mice re-infected with promastigote *L. donovani* parasites after a prior amastigote infection. A). Total flux in tissues from female mouse tissues IVIS imaged *ex vivo* after incubation with luciferin. Colour indicates infection as shown in the key. Single infection n=3; re-infection n=4. B). Total flux in tissues from male and female re-infected mice IVIS imaged *ex vivo* after incubation with luciferin. Each point is an individual mouse tissue. Colour indicates sex as shown in the key. Males n=5; females n=9. All values are with background flux subtracted. One re-infected female mouse bone marrow is missing as flux was negative after background subtraction. Mann-Whitney-Wilcoxon test. Significance of p-value: * p<0.05; **p<0.01; *** p<0.001: **** p<0.0001. Boxplots show the median, 25th and 75th quartiles, with whiskers representing all data-point within 1.5x Interquartile Range of the median; points above or below the whiskers represent outlying data values. Data shown is from two independent experiments.

Together, these findings show that latent parasites remain in multiple tissues of infected C57BL/6J mice, up to 12 weeks post-infection. Parasite burden does not

significantly differ between re-infected mice and those with a single infection, indicating that a secondary challenge does not reactivate a primary infection.

6.2.3 Prior amastigote infection is protective against subsequent colonisation in the bone marrow, liver and spleen

To understand how a prior infection altered tissue colonisation bottlenecks in a secondary infection, Barseq data for multiple tissues in mice which were re-infected with promastigote barcoded parasites, or infected with promastigote barcoded parasites with no prior infection, was analysed to calculate FP (see Chapter 3; Abel et al., 2015 for calculation method). Data from two independent experiments is shown in Figure 6.4. Figure 6.4A summarises the infection groups covered in this analysis.

FP was significantly higher in mice receiving only the single infection in the bone marrow (MW: $p=0.00028$), liver (MW: $p<0.0001$) and spleen (MW: $p=0.00123$) in females (Figure 6.4B), and in the liver in males (MW: $p=0.0317$) (Figure 6.10C). The liver was the most colonised tissue for both sexes in both infection models, but was much closer to other tissue FP after a previous infection had occurred, indicating a severe tightening of the bottleneck in response to prior infection in this tissue. No other tissues showed any significant difference between re-infection and single infection, indicating that prior amastigote infection is not protective against subsequent promastigote colonisation in the lymph nodes, gut or lung.

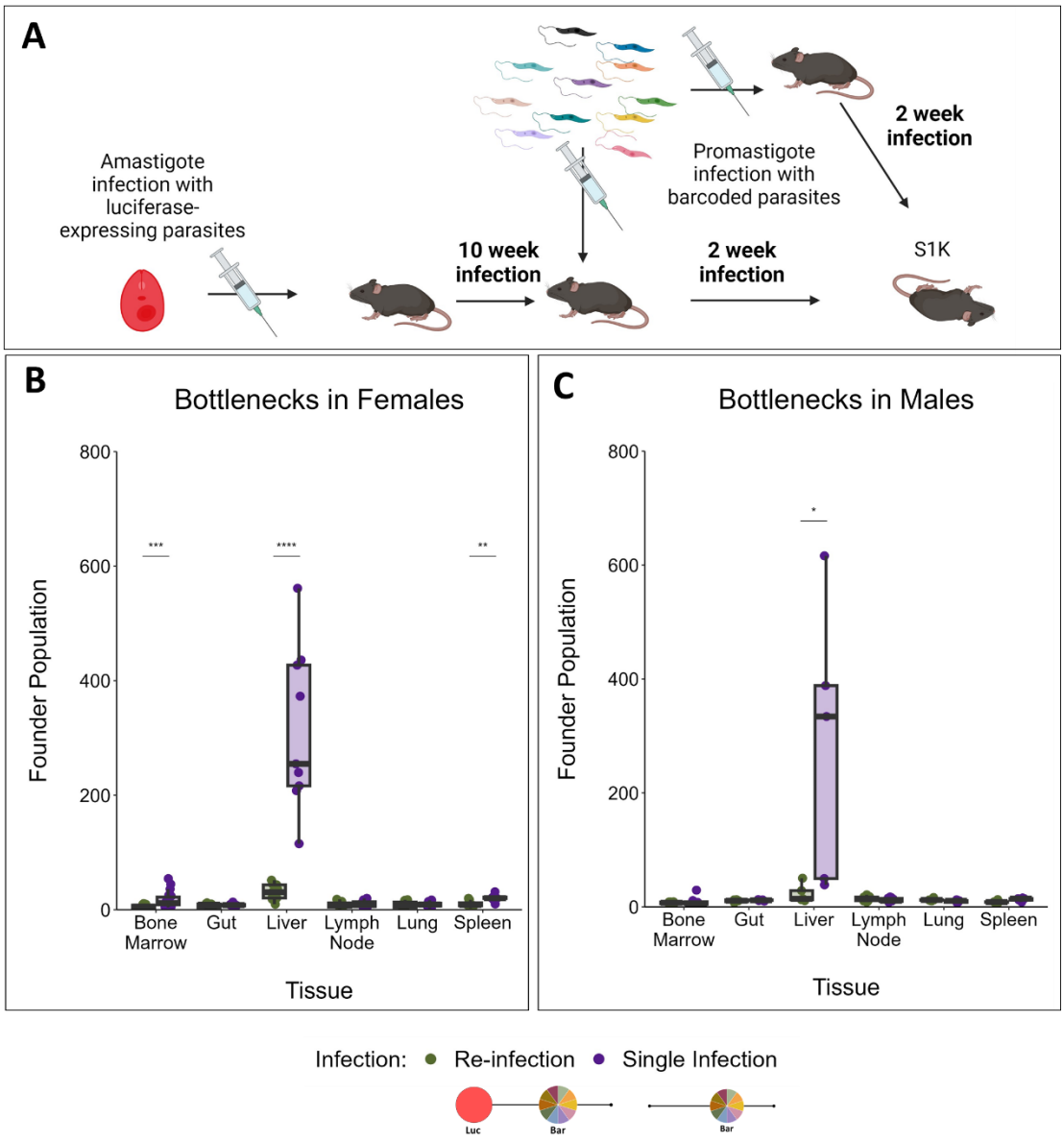


Figure 6.4: Founder populations in secondary infection parasite populations in multiple tissues in female C57BL/6J mice re-infected with promastigote *L. donovani* parasites after a prior amastigote infection. A). Schematic showing the subset of mice analysed. Single infection n=9; re-infection n=9. B). Founder population for the second infection in each tissue in female mice. C). Founder population for the second infection for each tissue in male mice. Single infection n=5; re-infection n=5. Each point is an individual mouse tissue. For one female single infection mouse, no sequence data was returned for the gut and as such, this data point is excluded in the analysis. Colour indicates infection as shown in the key and is consistent across all panels. Mann-Whitney-Wilcoxon test. Significance of p-value: * $p<0.05$; **$p<0.01$; *** $p<0.001$: **** $p<0.0001$. Boxplots show the median, 25th and 75th quartiles, with whiskers representing all data-point within 1.5x Interquartile Range of the median; points above or below the whiskers represent outlying data values. Data comprised of two independent experiments. Panel A created in Biorender.

These findings indicate that a protective effect against recolonisation by promastigote parasites occurs in the bone marrow, spleen and liver when mice have previously been infected with amastigote *L. donovani* parasites, with the liver especially becoming much more difficult to re-colonise due to a more restrictive bottleneck after prior infection. Protection against re-infection with *L. donovani* appears to be tissue-specific.

6.2.4 Prior infection has little impact on parasite dissemination between tissues in subsequent infections

To understand how this protective effect might alter dissemination patterns between tissues, pairwise GD comparisons were run between each tissue for each mouse. The results are shown in Figure 6.5 for females only. The GD did not significantly differ between re-infected and single infected mice for any tissue to tissue comparisons, except the liver to lymph node comparison (MW: p=0.0376) (Figures 6.5A and D), which indicated that populations in these tissues were more closely related in mice which had been previously infected with amastigote parasites.

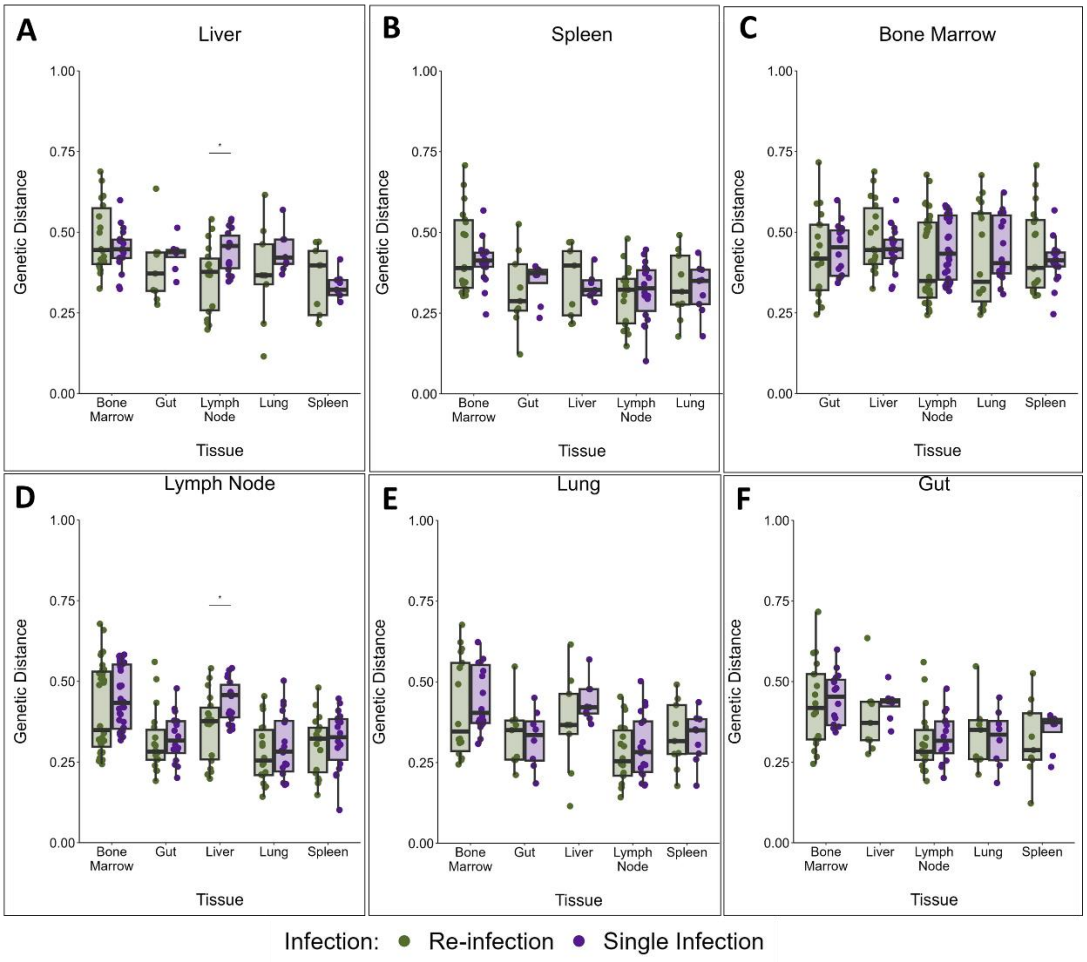


Figure 6.5: Genetic distance analysis for comparisons in secondary infection parasite populations in multiple tissues in female C57BL/6J mice re-infected with promastigote *L. donovani* parasites after a prior amastigote infection. Single infection n=9 females; re-infection n=9 females. A-F). Genetic distance for pairwise tissue comparisons for the second infection parasite population. For bone marrow and lymph nodes, two were taken per mouse. For one female single infection mouse, no sequence data was returned for the gut and as such, these comparisons are missing in the analysis. Each data point represents a pairwise comparison between two tissues. Colour indicates infection as shown in the key and is consistent across all panels. Mann-Whitney-Wilcoxon test. Significance of p-value: * p<0.05; **p<0.01; *** p<0.001: **** p<0.0001. Boxplots show the median, 25th and 75th quartiles, with whiskers representing all data-point within 1.5x Interquartile Range of the median; points above or below the whiskers represent outlying data values. Data comprised of two independent experiments.

These findings indicate that a prior infection has little impact on parasite trafficking between tissues in a subsequent infection; more restrictive bottlenecks in the spleen, liver and bone marrow are unlikely to be due to parasites being unable to

reach these tissues, but likely relate to a bottleneck occurring in the tissues themselves.

6.2.5 Primary infection is not re-activated by a secondary challenge

We then sought to understand how a secondary infection may affect the population dynamics of a chronic primary infection. To assess this, a re-infection model using the reverse order of parasite strains to that described in Figure 6.1 was implemented. The experimental set up is illustrated in Figure 6.6.

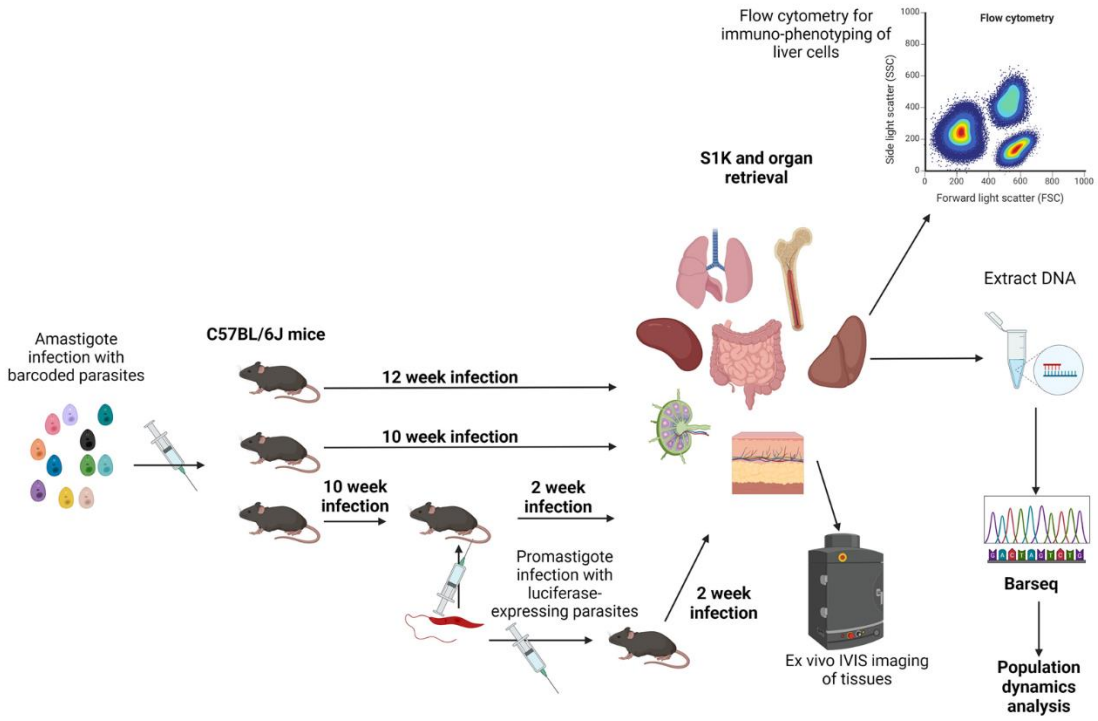


Figure 6.6: Schematic of experimental plan to investigate the population dynamics of the primary infection of *L. donovani* in a re-infection model using C57BL/6J mice. Re-infection n=4 females; 10 week infection n=4 females; 12 week infection n=4 females; 2 week infection n=3 females. C57BL/6J mice were infected with 3x10^7 amastigotes of the barcoded LV9 T7 cas9 *L. donovani* library. One group of these mice were killed at 10 weeks post-infection. A subset of mice were re-infected after 10 weeks, with the REH-*L. donovani* strain, using 5x10^7 stationary phase promastigotes via tail IV. Mice which had previously not been infected with any parasites were also infected with promastigotes at this point. Mice were imaged after the promastigote infection using the IVIS imaging system to monitor REH-*L. donovani* parasite burdens. Mice were then killed two weeks after the promastigote infections (12 weeks post-initial infection). Tissues were taken from each mouse, incubated with 30 mg/ml luciferin, then imaged *ex vivo* using the IVIS imaging system. Tissues were subsequently flash frozen on dry ice, then stored at -80 °C before subsequent DNA extraction, Illumina sequencing, Barseq analysis and population dynamics studies, alongside a sample of the input parasite pool from each infection. Tissues taken consisted of liver, spleen, two lymph nodes and two each of tibia plus femur. Bone marrow was extracted by flushing and stored at -20 °C for DNA extraction. Skin from the rump of the mice was removed and imaged *ex vivo*, then five randomly selected patches were biopsied and frozen for DNA extraction. Liver cells were used for flow cytometry analysis. Figure created in Biorender.

Mice which received REH parasites were subject to whole body imaging after the infection, to assess whether any differences in overall parasite burden were detectable. The trajectory of re-infected and single infected mice was not significantly different after re-infection, indicating that any impact of prior infection on whole body parasite burden in subsequent infection is below detection limits using the IVIS system.

The tissue parasite burden for the second infection, with promastigote REH-*L. donovani* in uninfected or barcoded amastigote-infected mice (Figure 6.7A) was also assessed via *ex vivo* imaging, to understand how a second infection is affected by the occurrence of a prior infection at the tissue level. Exemplar images are shown in Figure 6.7B. Notably, the lung of the single infection mouse shown exhibits luminescence in one lobe only, which implies that the lung lobes may be individually colonised.

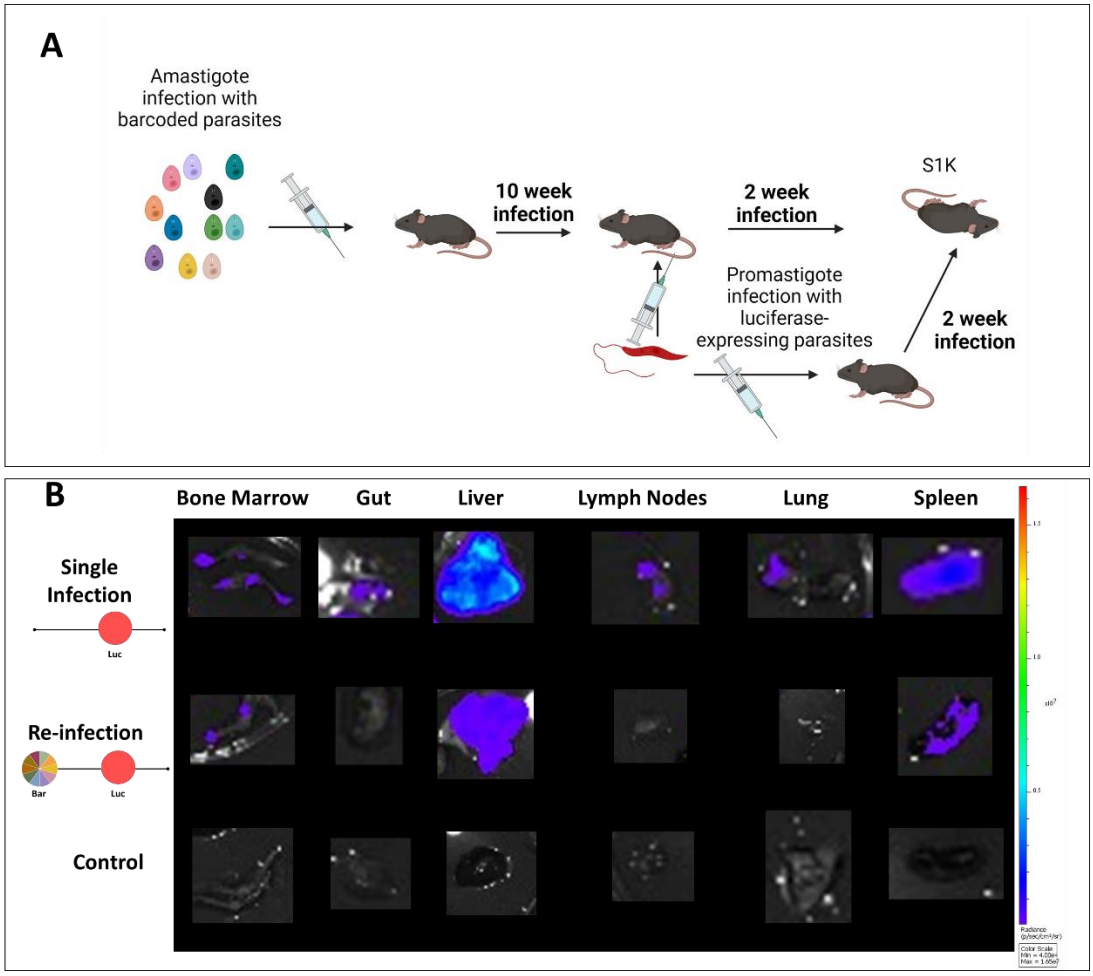


Figure 6.7: Exemplar *ex vivo* imaging of tissue parasite burden of secondary promastigote infection in female C57BL/6J mice re-infected with promastigote *L. donovani* parasites after a prior amastigote infection. A). Schematic showing subset of mice analysed. B). Exemplar tissues after ex vivo IVIS imaging from mouse from the short single infection group versus one mouse from the long infection followed by re-infection group. A control which had no REH-*L. donovani* infection is also shown. Scale indicates radiance. Single infection n=3 females; re-infection n=4 females. Mice were killed two weeks after the promastigote infection. Panel A created in Biorender.

When tissue parasite burdens were quantified for these mice (Figure 6.8), infected mice had significantly greater total flux than mice with no REH infection. No difference was found between re-infected and single infection mice, but this is likely to be a result of the small sample size for each group. There is a clear trend towards decreased radiance in the re-infected mice, but an overall small effect size, indicating that secondary infection parasite burdens are likely lower in most tissues when mice have previously been infected, compared to those which are

experiencing their first infection. Grouped infected tissues had significantly greater total flux than control mice for all tissues, confirming parasite presence.

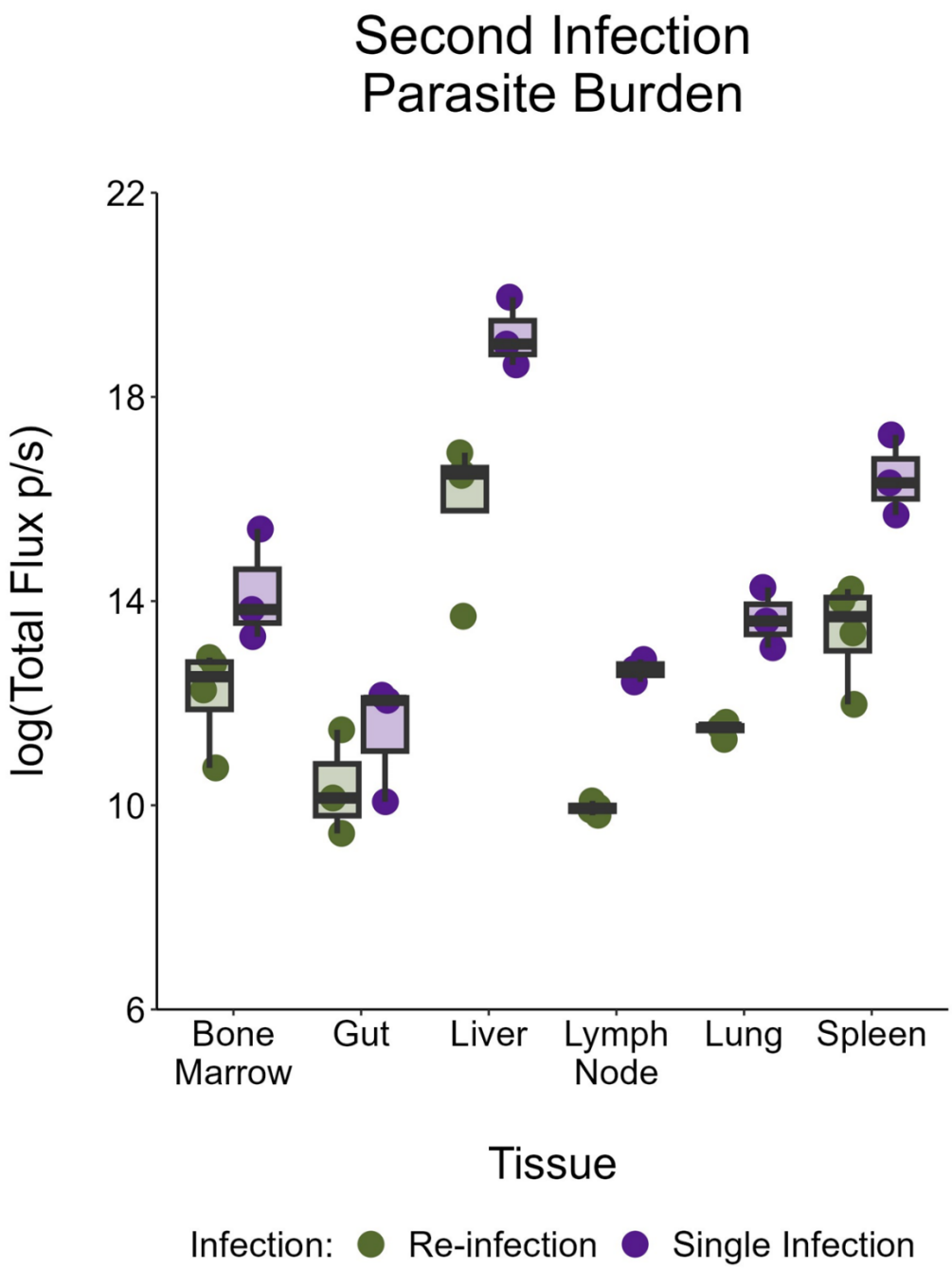


Figure 6.8: Tissue parasite burden, quantified by *ex vivo* imaging, of secondary promastigote infection in female C57BL/6J mice re-infected with promastigote *L. donovani* parasites after a prior amastigote infection. Total flux in tissues from female mouse tissues IVIS imaged *ex vivo* after incubation with luciferin. Single infection n=3 females; re-infection n=4 females. All values are with background flux subtracted. One re-infected mouse gut is missing as flux was negative after background subtraction. Each point is an individual mouse tissue. Colour indicates infection as shown in the key. Mann-Whitney-Wilcoxon test between all infected versus controls. Mann-Whitney-Wilcoxon test. Significance of p-value: * p<0.05; **p<0.01; *** p<0.001: **** p<0.0001. Boxplots show the median, 25th and 75th quartiles, with whiskers representing all data-point

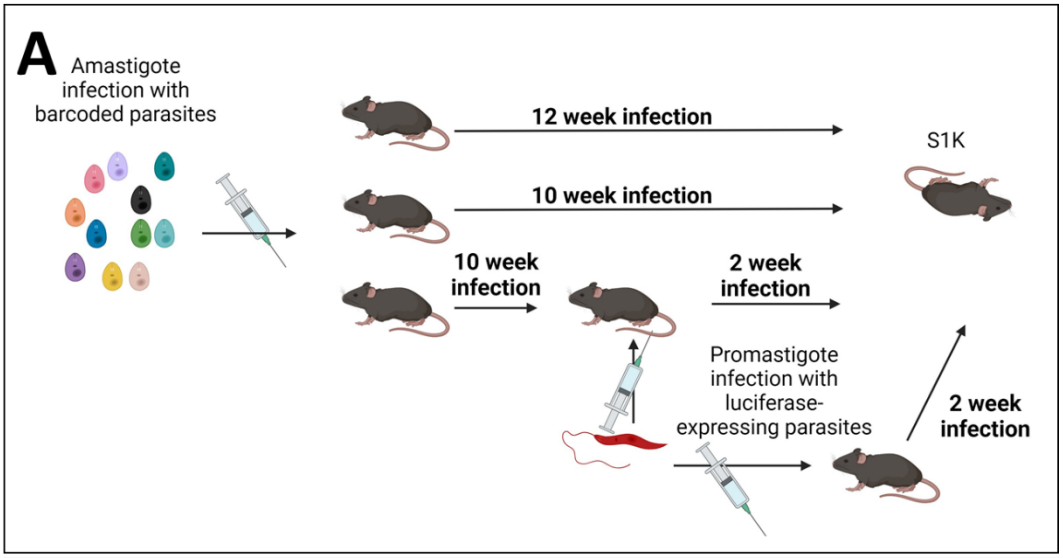
within 1.5x Interquartile Range of the median; points above or below the whiskers represent outlying data values.

These data indicate that a secondary infection is not prevented from establishing in any tissue assessed, despite the continued presence of the prior infection parasite population. However, there are indications that the ability of the new parasites to replicate may be hampered by the occurrence of a previous infection.

6.2.6 Evidence that latent parasites traffic back to the liver from other tissues after a subsequent infection occurs

To understand how the population dynamics of a prior population might be altered by a secondary infection, the FP for the initial amastigote barcoded parasite infection was calculated for each tissue (Figure 6.9A). Little difference due to biological events was expected between mice infected for 10 or 12 weeks with barcoded parasites and as such, these mice were combined as 'single infection' mice for the analysis, on the assumption that heterogeneity was the result of random mouse to mouse differences.

The FP for most tissues did not differ between re-infected and single infection mice (Figure 6.9B). Notably, the spleen FP was the highest in this chronic phase of infection, whilst the liver FP had greatly decreased compared to earlier stages of infection (Figure 6.9B; see chapter 4 for earlier amastigote infection FP data), indicating that a tightening of the bottleneck was occurring over time in this tissue. Interestingly, in the liver the FP significantly increased after re-infection (MW: $p=0.0485$), compared to mice with a single infection. This indicates that latent parasites are able to re-colonise this tissue, presumably via trafficking from other tissues with a chronic parasite reservoir, and that this can be triggered by re-infection. A similar effect was also observed in the lymph nodes (MW: $p=0.0192$), but in no other tissues.


A
Amastigote infection with barcoded parasites
12 week infection
10 week infection
10 week infection
2 week infection
S1K
Promastigote infection with luciferase-expressing parasites
2 week infection

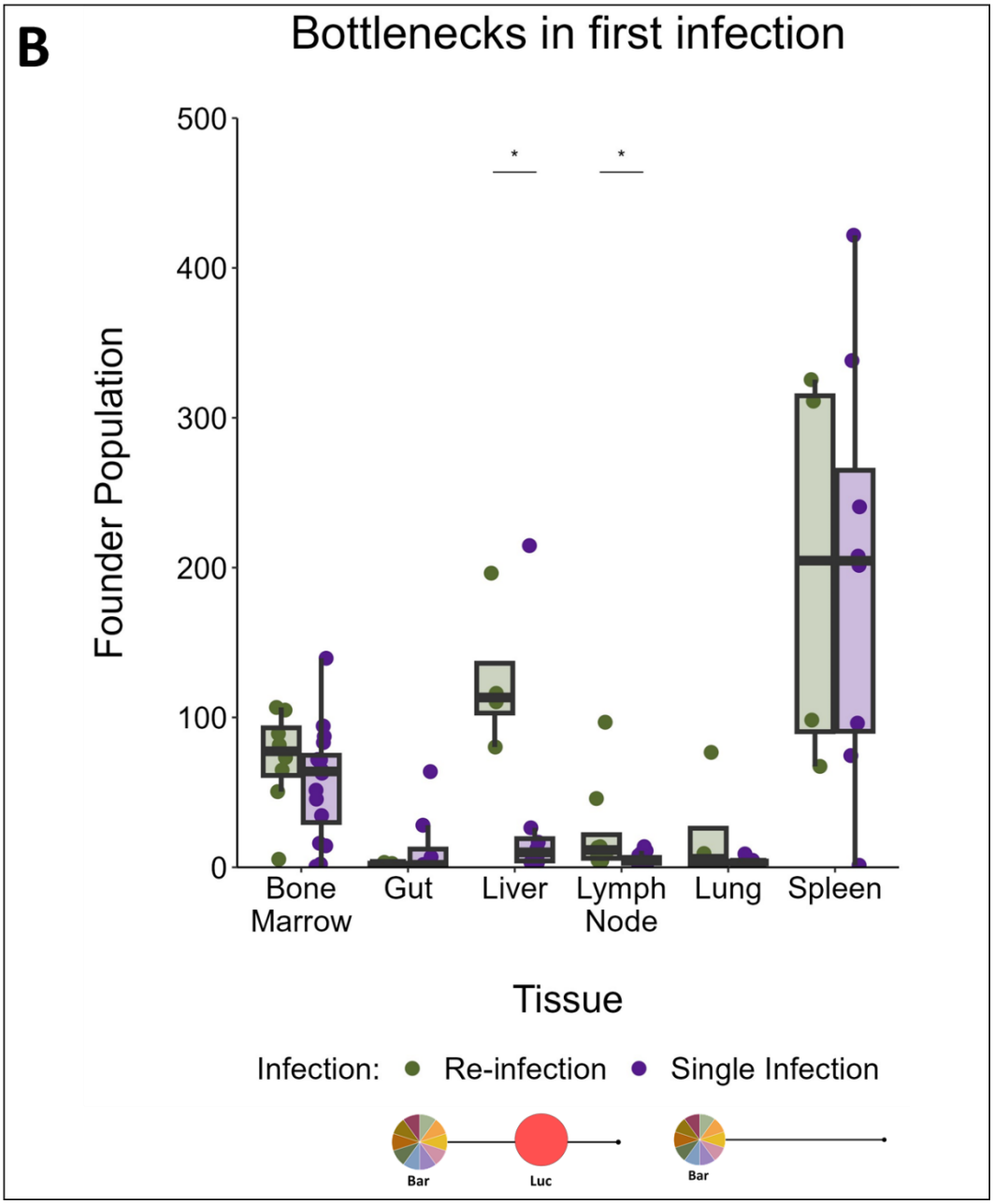

B
Bottlenecks in first infection
Founder Population
500
400
300
200
100
0
*
*
Bone Marrow
Gut
Liver
Lymph Node
Lung
Spleen
Tissue
Infection: Re-infection Single Infection
Bar
Luc
Bar

Figure 6.9: Founder populations in the initial infection in multiple tissues in female C57BL/6J mice re-infected with promastigote *L. donovani* parasites after a prior amastigote infection. A). Schematic showing the subset of mice analysed. Single infection n=8 females; re-infection n=4 females. B). Founder population of the initial infection for each tissue in single or re-infected mice. Each point is an individual mouse tissue. Colour indicates infection as shown in the key and is consistent across all panels. Mann-Whitney-Wilcoxon test. Significance of p-value: * p<0.05; **p<0.01; *** p<0.001: **** p<0.0001. Boxplots show the median, 25th and 75th quartiles, with whiskers representing all data-point within 1.5x Interquartile Range of the median; points above or below the whiskers represent outlying data values. Panel A created in Biorender.

Therefore, GD was assessed to compare the liver parasite population, with and without re-infection, to other tissues in the same mouse, to determine where these parasites may be trafficking from (Figure 6.10). Figure 6.10A shows that the liver parasites from the initial infection in a re-infected mouse were more closely related to those of the bone marrow (MW: p=0.00216), lymph nodes (MW: p=0.0192) and spleen (MW: p=0.0485) than in mice with only a single infection. The gut and lymph node populations were also noted to be more closely related in re-infected mice, although a trend towards greater relatedness to the gut which did not reach the threshold for significance was observed for all tissues (Figure 6.10E).

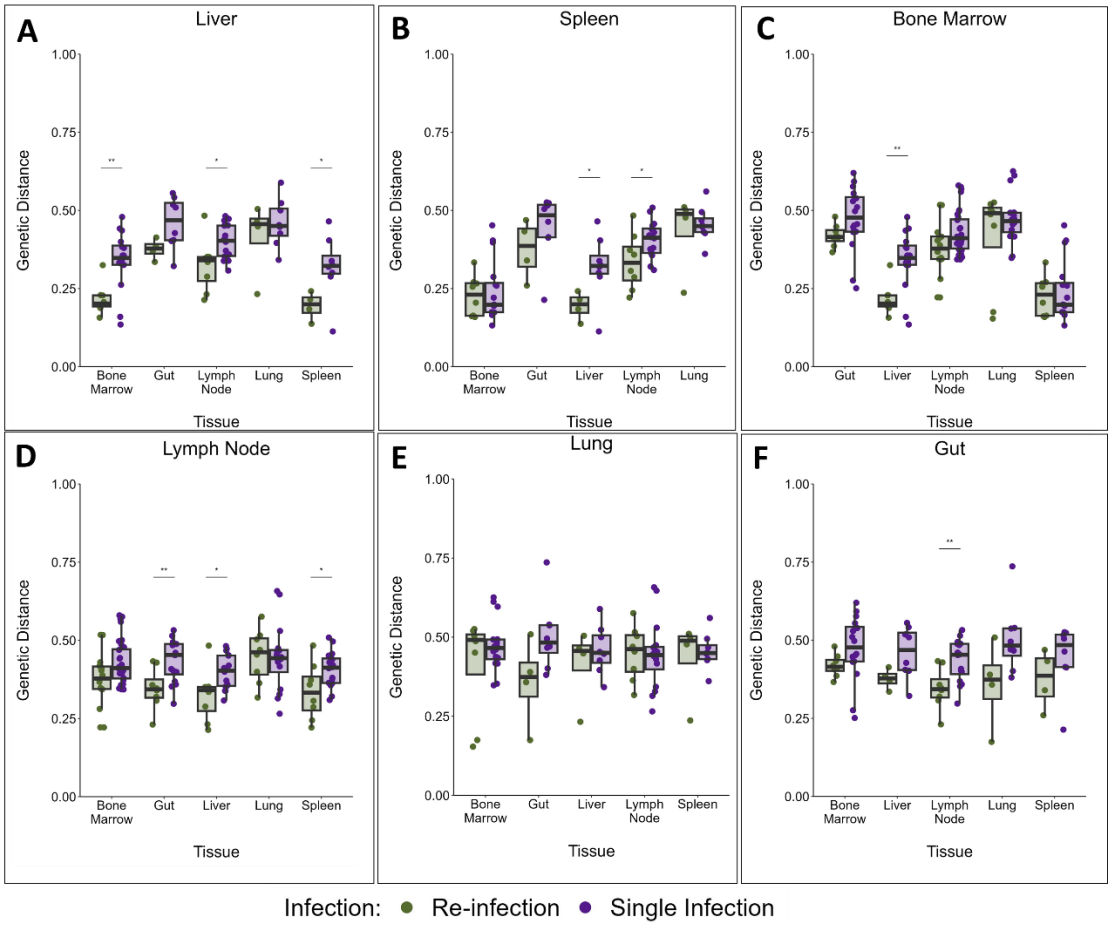


Figure 6.10: Genetic distance analysis for comparisons of initial parasite populations in multiple tissues in female C57BL/6J mice re-infected with promastigote *L. donovani* parasites after a prior amastigote infection. Single infection n=8 females; re-infection n=4 females. A-F). Genetic distance for pairwise tissue comparisons for the first infection parasite population. For bone marrow and lymph nodes, two were taken per mouse. Each data point represents a pairwise comparison between two tissues. Colour indicates infection as shown in the key and is consistent across all panels. Mann-Whitney-Wilcoxon test. Significance of p-value: * p<0.05; **p<0.01; *** p<0.001: **** p<0.0001. Boxplots show the median, 25th and 75th quartiles, with whiskers representing all data-point within 1.5x Interquartile Range of the median; points above or below the whiskers represent outlying data values.

Together, these findings provide evidence for re-colonisation of the liver by latent parasites after re-infection of a chronically-infected mouse with *L. donovani*. These latent parasites appear to traffic to the liver from the bone marrow, spleen and possibly lymph nodes, whilst persistent parasites from the gut also appear to move to the lymph nodes on re-infection.

6.2.7 Patchy parasite colonisation with low diversity characterises skin dissemination in an immune-competent mouse

A key concern in *L. donovani*-caused VL infections is the potential development of PKDL, for which there is currently no mouse model system. The mechanisms for dissemination of parasites to the skin are unknown but one hypothesis is that re-infection may act as a trigger for this skin sequelae (Mukhopadhyay et al., 2014). Since REH parasites are able to be visualised, the skin from all mice infected with this parasite line (for 12 or 2 weeks and with or without re-infection) was imaged *ex vivo* to see whether parasites are visible under these infection conditions (Figure 6.11).

Figure 6.11A shows skin from short-term single REH promastigote infection and from mice re-infected after a prior barcoded parasite infection. Figure 6.11B shows skin from a single REH amastigote infection and from mice experiencing a long-term REH amastigote infection followed by re-infection with barcoded parasites. Figure 6.11C shows a control which was not infected with REH parasites and the scale for radiance, which is consistent for all images. Two mice which were excluded from all other analysis are shown here; exclusion criteria were based on a lack of evidence of systemic parasite burdens (see Methods), but one hypothesis for this occurrence was that the parasites could have sequestered in the skin for some hitherto unknown reason.

The results overall show a patchy distribution of parasites across three infection groups, with some mice exhibiting many patches and others a single patch. Patch intensity varies, as does the size of patches, both within and between groups. Parasites are visible at both 2 and 12 weeks post-infection. Notably, parasite patches are not clearly visible in any mice which received an REH infection after an initial infection with barcoded parasites, indicating that parasite burdens (if present) in the skin of these mice are much lower than in the other groups. Parasites were also not visible in the two otherwise excluded mice, indicating that they had not become established in the skin preferentially.

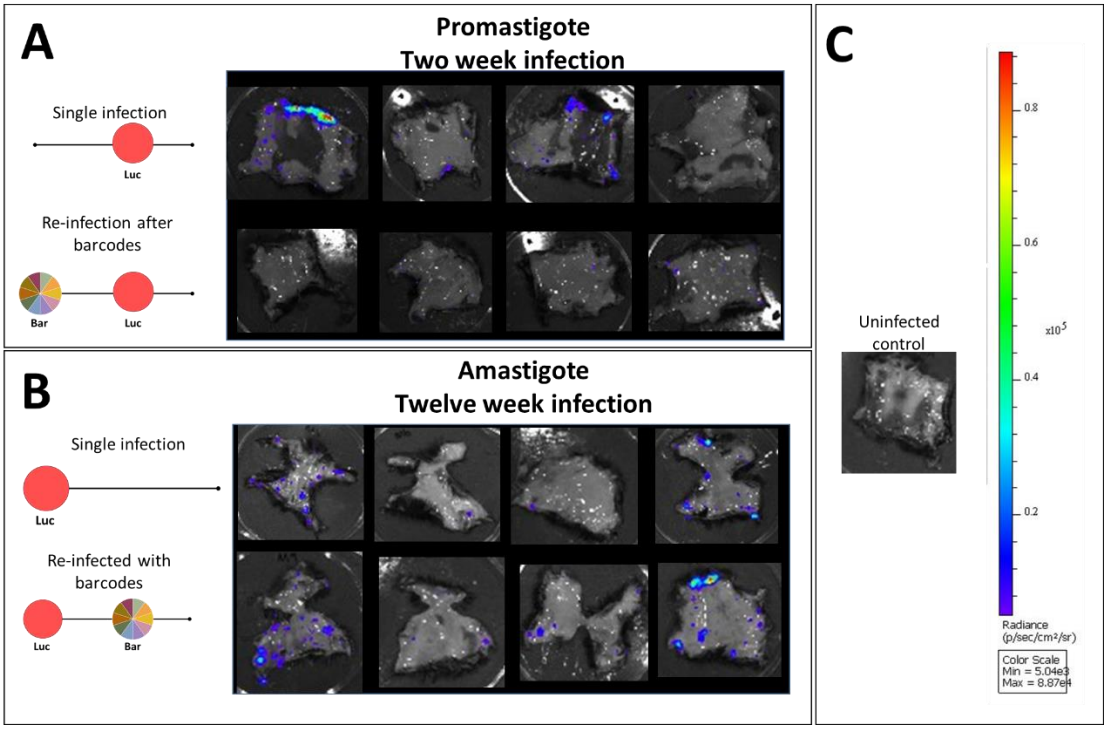


Figure 6.11: Patchy distribution of skin parasites in female C57BL/6J mice re-infected with promastigote *L. donovani* parasites after a prior amastigote infection. For single infection mice, female C57BL/6J mice were infected with 3x10^7 barcoded amastigote *L. donovani* or REH-*L. donovani* via tail IV for 12 weeks (n=4 per parasite line); for re-infection mice, the same protocol was followed, but mice were re-infected with the other stain of 4-5x10^7 stationary phase promastigote parasites via intravenous tail injection 10 weeks after initial infection (n=4 per parasite line). For promastigote single infection mice, the mice received only the promastigote REH parasites, with no prior infection (n=4). Mice were killed two weeks after the promastigote infection. Skin was removed, incubated with 30 mg/ml luciferin for 5 minutes in the dark, then imaged for 5 minutes using the IVIS. A). Skin from mice infected with REH-*L. donovani* promastigotes for 2 weeks with or without prior infection with barcoded parasites. The promastigote single infection mouse not included in other analysis (M20) is shown here. B). Skin from mice infected with REH-*L. donovani* promastigotes for 12 weeks with or without subsequent infection with barcoded parasites. The amastigote single infected mouse not included in other analysis (M7) is shown here. C). Control skin sample not infected with REH-*L. donovani* (this mouse was infected with barcoded *L. donovani* parasites). Scale indicates radiance in p/sec/cm^2/sr and applies to all images. Images were produced using Living Image software.

Since parasites could be visualised, DNA from five randomly selected locations in skin from mice infected with barcoded parasites was assessed to determine how

diverse the skin populations might be and the bottlenecks to skin colonisation (Figure 6.12A). This analysis indicated that a tight bottleneck appears to act in skin colonisation, with a very low FP for almost all patches across all mice (Figure 6.12A). A small number of patches had more diversity, but as patches were selected randomly, with no way to visualise the barcoded parasites, and were few in number per mouse, this could be unrepresentative of diversity across a mouse as a whole.

The relatedness of each skin patch to other tissues in the same mouse was also assessed, with relatedness in the long-term amastigote barcode infections shown in Figure 6.12B and for the shorter promastigote barcode infections in Figure 6.12C. The data show few differences between re-infection and single infection in the shorter promastigote infection, but in chronic amastigote infection, the gut (MW: $p=0.000136$) and lung (MW: $p=0.0181$) parasites are both more closely related to the skin populations in mice after re-infection, whilst the liver population is less closely related to skin parasites in re-infected mice (MW: $p=0.035$).

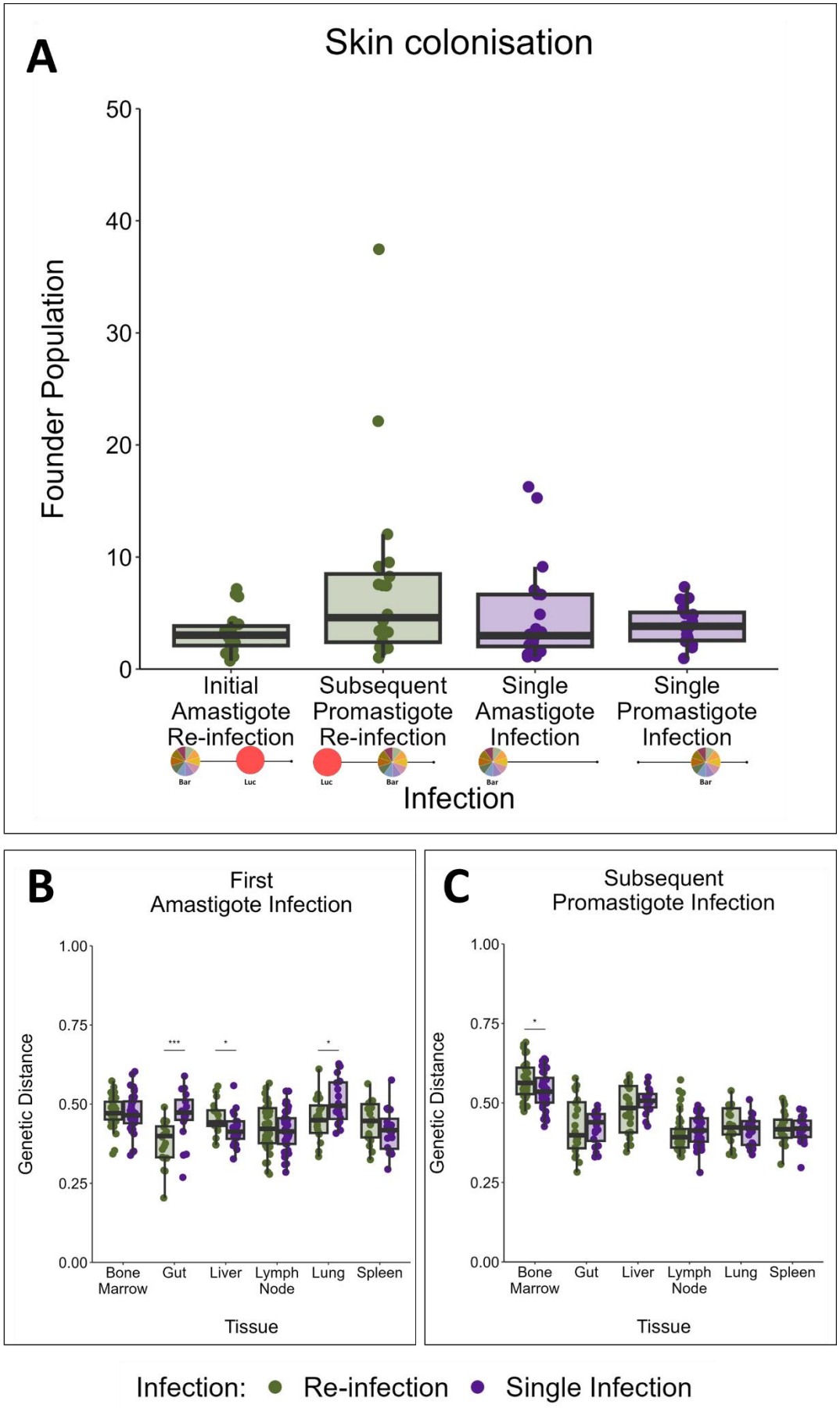


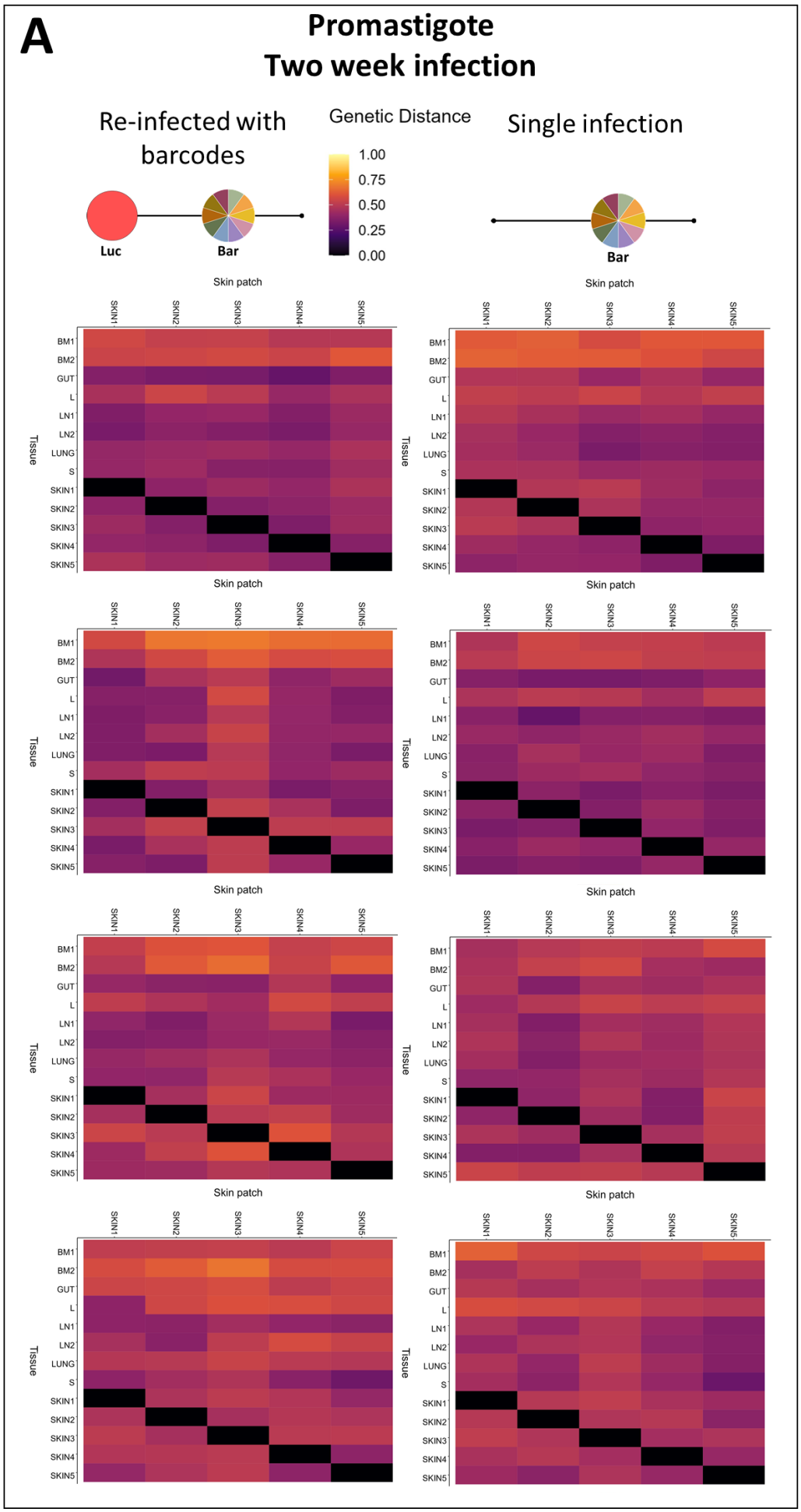
Figure 6.12: Skin colonisation in female C57BL/6J mice re-infected with promastigote *L. donovani* parasites after a prior amastigote infection. For amastigote single infection mice,

female C57BL/6J mice were infected with 3x10^7 barcoded amastigote *L. donovani* or REH-*L. donovani* via tail IV for 12 weeks (n=4 per parasite line); for re-infection mice, the same protocol was followed, but mice were re-infected with the other strain of 4-5x10^7 stationary phase promastigote parasites via intravenous tail injection 10 weeks after initial infection (n=4 per parasite line). For promastigote single infection mice, the mice received only the promastigote barcoded parasites, with no prior infection (n=4). Mice were killed two weeks after the promastigote infection. Skin was removed and five skin biopsies were taken per mouse from randomly selected patches and stored at -80 °C prior to DNA extraction. A). Founder populations for each skin patch. The parasite group referred to in the x-axis is the parasite population assessed (the barcoded population). Each data point represents one skin patch from one mouse, with 5 per mouse. B). Genetic distance for pairwise tissue comparisons between each skin patch and all other tissues from the same mouse for 12 week amastigote-infection with barcoded parasites, with or without subsequent infection with REH-*L. donovani*. C). Genetic distance for pairwise tissue comparisons between each skin patch and all other tissues from the same mouse for 2 week promastigote-infection with barcoded parasites, with or without prior infection with REH-*L. donovani*. For bone marrow and lymph nodes, two were taken per mouse, so there are two comparisons for each mouse for these tissues. . Each data point represents a pairwise comparison between two tissues. Colour indicates infection as shown in the key and is consistent across all panels. Mann-Whitney-Wilcoxon test (for GD only). Significance of p-value: * p<0.05; **p<0.01; *** p<0.001: **** p<0.0001. Boxplots show the median, 25th and 75th quartiles, with whiskers representing all data-point within 1.5x Interquartile Range of the median; points above or below the whiskers represent outlying data values.

The low FP indicates that colonisation is likely a stochastic event, so the relatedness of different tissues to skin patches, and the relatedness of skin patches themselves, was then compared per mouse, to see whether there was heterogeneity in skin from the same mouse, and whether different mice from the same group might show differing patterns in terms of skin dissemination.

Figure 6.13 shows individual heat-maps for each mouse, indicating how closely related skin patches are to other tissues. Blue indicates a close relationship, whilst yellow means populations are less related to each other. Interestingly, there is a striking trend of increased yellow for bone marrow relatedness, indicating that the bone marrow is not a key disseminator to skin. Other tissues show varying

relationships to skin patches. The relation between different skin patches also varies, with no obvious differences in trends between promastigote (Figure 6.13A) and amastigote (Figure 6.13B) infection, or between single and re-infection.


A
Promastigote
Two week infection
Re-infected with barcodes
Luc
Bar
Genetic Distance
1.00
0.75
0.50
0.25
0.00
Single infection
Bar
Skin patch
Tissue
SKIN1
SKIN2
SKIN3
SKIN4
SKIN5
BM1
BM2
GUT
L
LN1
LN2
LUNG
S

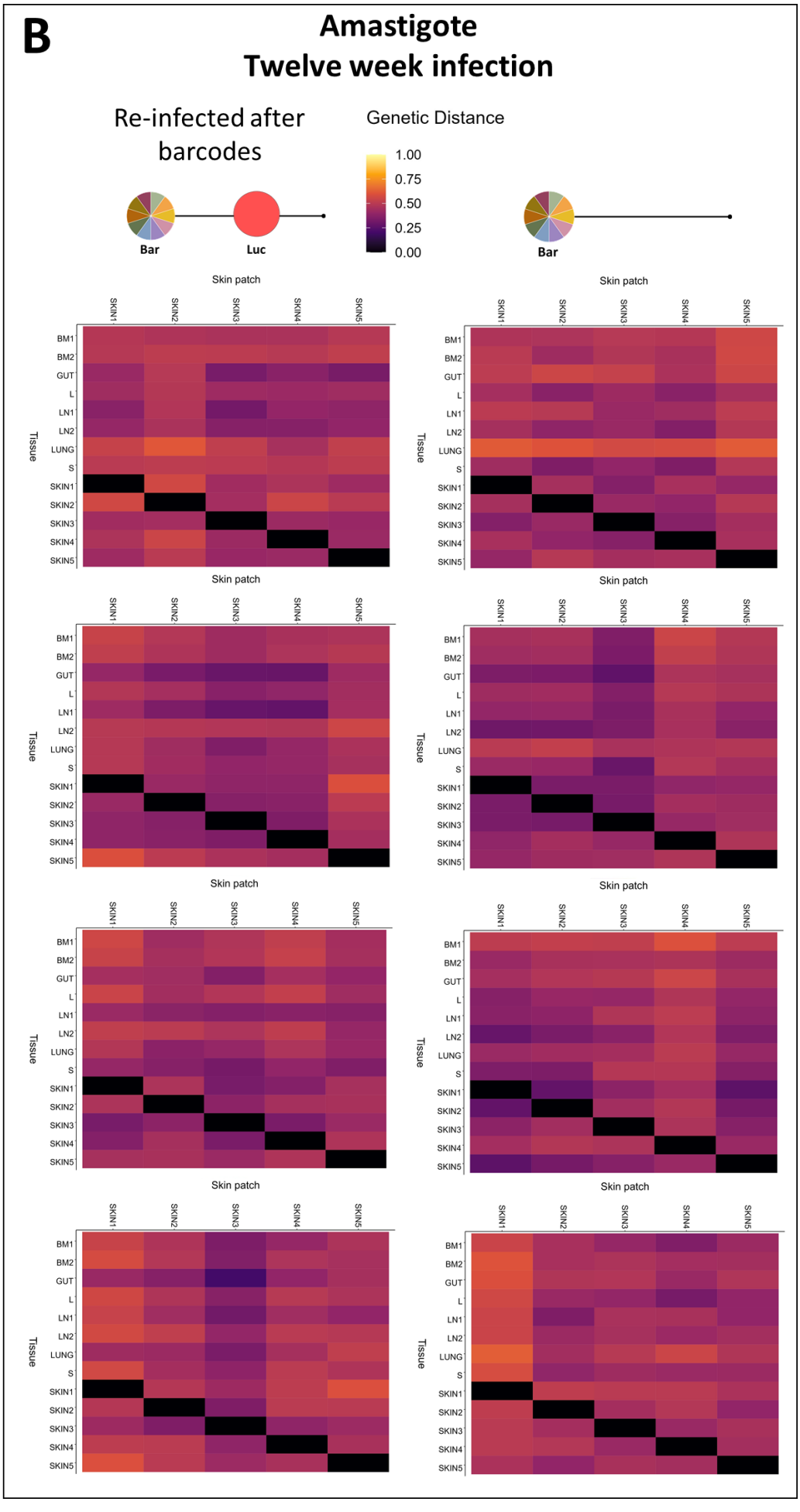

B
Amastigote
Twelve week infection
Re-infected after barcodes
Genetic Distance
1.00
0.75
0.50
0.25
0.00
Bar
Luc
Bar
Skin patch
Tissue
SKIN1
SKIN2
SKIN3
SKIN4
SKIN5
BM1
BM2
GUT
L
LN1
LN2
LUNG
S

Figure 6.13: Skin parasite population relatedness to other tissues in female C57BL/6J mice re-infected with promastigote *L. donovani* parasites after a prior amastigote infection. A). Genetic distances in each mouse for pairwise tissue comparisons between each skin patch and all other tissues from the same mouse for 2 week promastigote-infection with barcoded parasites. B). Genetic distances in each mouse for pairwise tissue comparisons between each skin patch and all other tissues from the same mouse for 12 week amastigote-infection with barcoded parasites. For bone marrow and lymph nodes, two were taken per mouse, so there are two comparisons for each mouse for these tissues. Colour indicates genetic distance as shown in the key and is consistent across all panels. BM1 and BM2 = Bone Marrow 1 and 2; LN1 and LN2 = Lymph Node 1 and 2. L = Liver. S = Spleen.

Together, these findings establish that a patchy parasite distribution occurs in skin in *L. donovani* infected C57BL/6J mice, and can be seen as early as 2 weeks post-infection using REH parasites. Parasites persist in the skin up to 12 weeks post-infection. The parasites appear to be stochastically seeded to the skin from various other tissues and are not necessarily closely related to other skin patches. Additionally, re-infection may trigger seeding of latent parasites; those which were previously resident in the gut and lung appear able to disseminate to the skin. Strikingly, the bone marrow seems to rarely act as a skin parasite source compared to other tissues.

6.2.8 Differences in host cell availability at 2 weeks post re-infection do not account for liver re-colonisation

In order to understand how bottlenecks to colonisation in the liver might be mediated in chronic and acute single infections and after re-infection, flow cytometry was used to immuno-phenotype the proportions of key immune cells in the livers of female mice. It was additionally hypothesised that differences in known host cell numbers in this tissue may help establish how parasites may be trafficking back to the liver in re-infected mice. To analyse this, mice which received either parasite strain infection were amalgamated for each group (i.e. mice which were infected for 12 weeks with either parasite were grouped, and the same for those infected for 2 weeks, whilst all re-infected mice, regardless of parasite order, were

grouped together). This was based on the assumption that any true effects of re-infection versus chronic infection would occur in both re-infection models.

The gating strategy for cell classification is shown in Figure 6.14. Live cells were gated based on LDAqua expression (Figure 6.14A), and these were then gated for NK1.1 (NK cells), CD3 (T cells), CD11b, CD20 (B cells), Ly6G (neutrophils), CD11c (dendritic cells), F4/80 (macrophages), Ly6C and CCR2 (monocytes) in the manner shown in Figure 6.14B. Macrophages were further categorised as F4/80^{low} CCR2+ or F4/80^{high} CCR2- to distinguish resident and newly monocyte-derived respectively. Neutrophils were classified as mature or immature based on high or low Ly6G expression respectively.

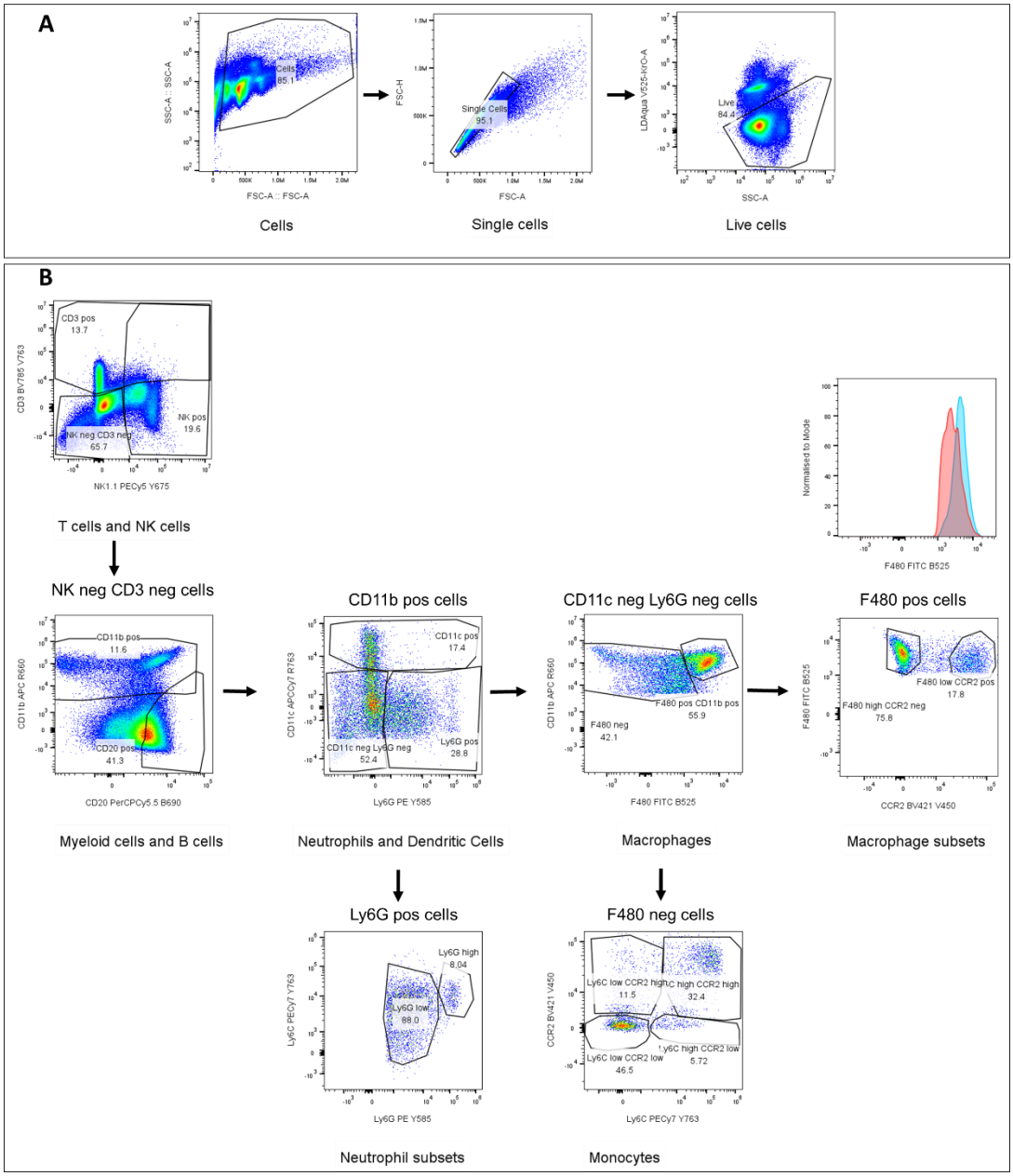


Figure 6.14: Gating strategy for flow cytometry analysis of liver immune cells in female C57BL/6J mice re-infected with promastigote *L. donovani* parasites after a prior amastigote infection. Liver cells from *L. donovani* infected and uninfected control (n=3) female C57BL/6J mice were processed to a single cell suspension and stained with various antibodies to classify them using flow cytometry. A). Gating for live cells excluded debris using FSC vs SSC, then characterised single cells using SSC-A vs SSC-H, then gated for live cells using LDAqua as shown. B). Live cells were characterised as T cells (CD3+), NK cells (NK1.1+), or double negative, then gated as B cells (CD20+). CD11b+ cells were further gated as neutrophils (Ly6G+) and dendritic cells (CD11c+). Double negatives were then gated as macrophages (F4/80+) or monocytes (F4/80-), with monocytes further categorised as inflammatory (CCR2+, Ly6Chigh) or patrolling (CCR2-, Ly6Clow). Macrophages were further gated as resident (F4/80high, CCR2-) or monocyte-derived (F4/80low, CCR2high).

Neutrophils were further gated as mature (Ly6Ghigh) or immature (Ly6Glow). All gating shown is from an uninfected control mouse.

Initially, cell types known to act as host cells for *Leishmania* were assessed, to determine any differences which may account for the findings detailed above. Figure 6.15 shows the proportions of CD11b+ cells which were categorised as CCR2+ inflammatory monocytes (Figure 6.15A), monocyte-derived macrophages (Figure 6.15B), dendritic cells (Figure 6.15C) and resident macrophages (Figure 6.15D). The data did not show any significant differences to uninfected mice for any infection groups, and no differences between re-infected and chronically infected mice were noted for any of these cell types.

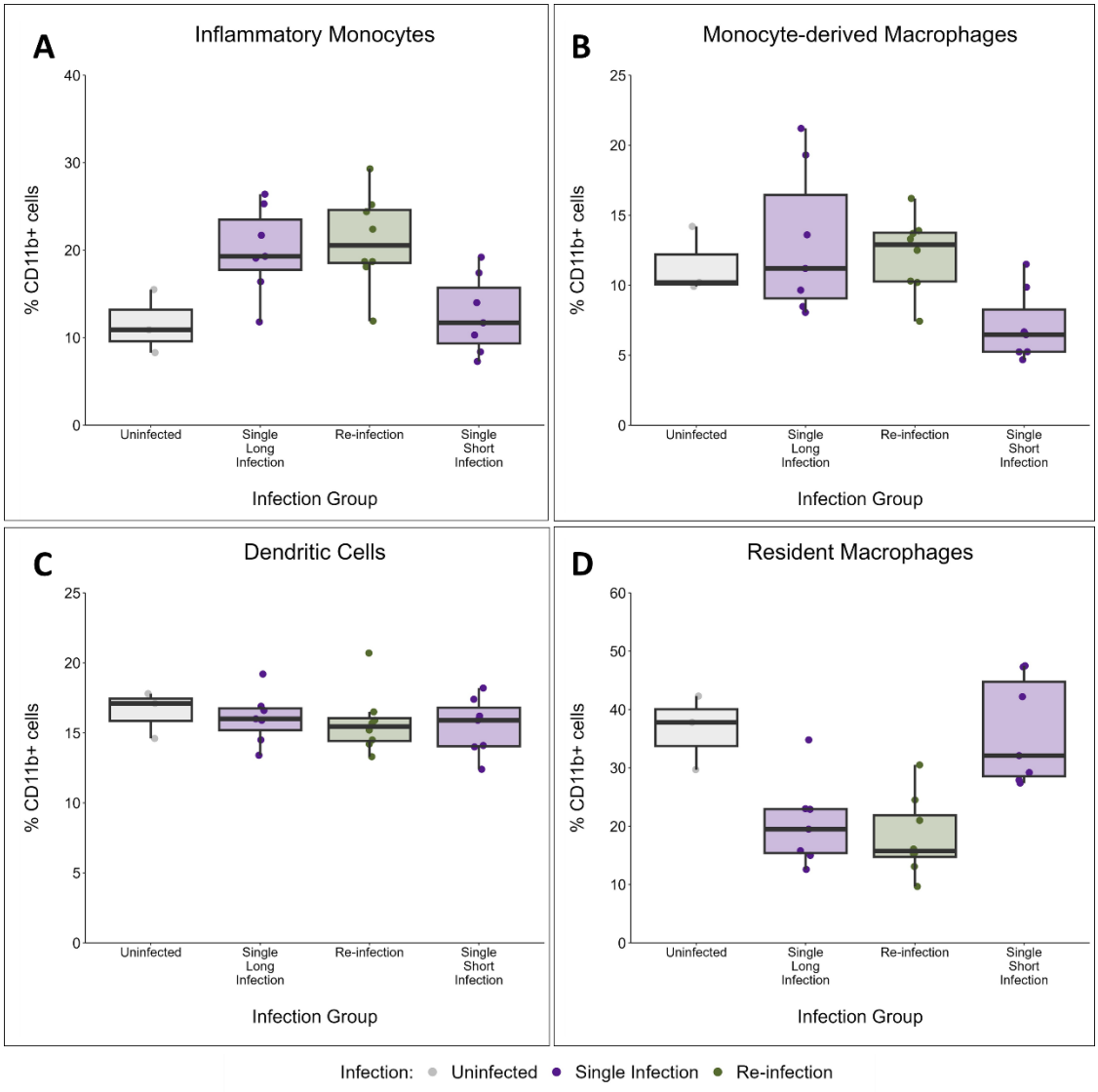


Figure 6.15: Monocytes, macrophages and dendritic cells in the livers of in female C57BL/6J mice re-infected with promastigote *L. donovani* parasites after a prior amastigote infection. Single long infection n=7 females; re-infection n=8 females; single short infection n=7 females. Control mice were uninfected (n=3). Livers were processed to a single cell suspension and assessed for various markers using flow cytometry. A). Inflammatory monocytes as a percentage of CD11b+ cells. B). Monocyte-derived macrophages as a percentage of CD11b+ cells. C). Dendritic cells as a percentage of CD11b+ cells. D). Resident macrophages as a percentage of CD11b+ cells. Each point represents one individual mouse liver. Colour indicates infection as shown in the key and is consistent across all panels. Mann-Whitney-Wilcoxon test with Bonferroni correction to compare infected groups to control reference, and Mann-Whitney-Wilcoxon test between single long infection and re-infection groups. Significance of (adjusted in comparisons to control) p-value: * $p<0.05$; ** $p<0.01$; *** $p<0.001$: **** $p<0.0001$. Boxplots show the median, 25th and 75th quartiles, with whiskers representing all data-point within 1.5x Interquartile Range of the median; points above or below the whiskers represent outlying data values.

Differences in neutrophil numbers (Figure 6.16A) and in the proportions of mature (Figure 6.16B) and immature (Figure 6.16C) neutrophils were also quantified. These cells can also act as host cells, but have a complex relationship with *Leishmania* parasites, sometimes exacerbating and sometimes abrogating disease. Again, no differences between the chronic infection and the re-infection groups were found for total neutrophils or either subset. When infected groups were compared to the uninfected control, only mature neutrophils showed any significant difference, with both long term (MW: $p=0.05$) and re-infected (MW: $p=0.036$) mice having higher levels of these cells than the control group. The short single infection did not significantly alter the numbers of these cells from the uninfected baseline.

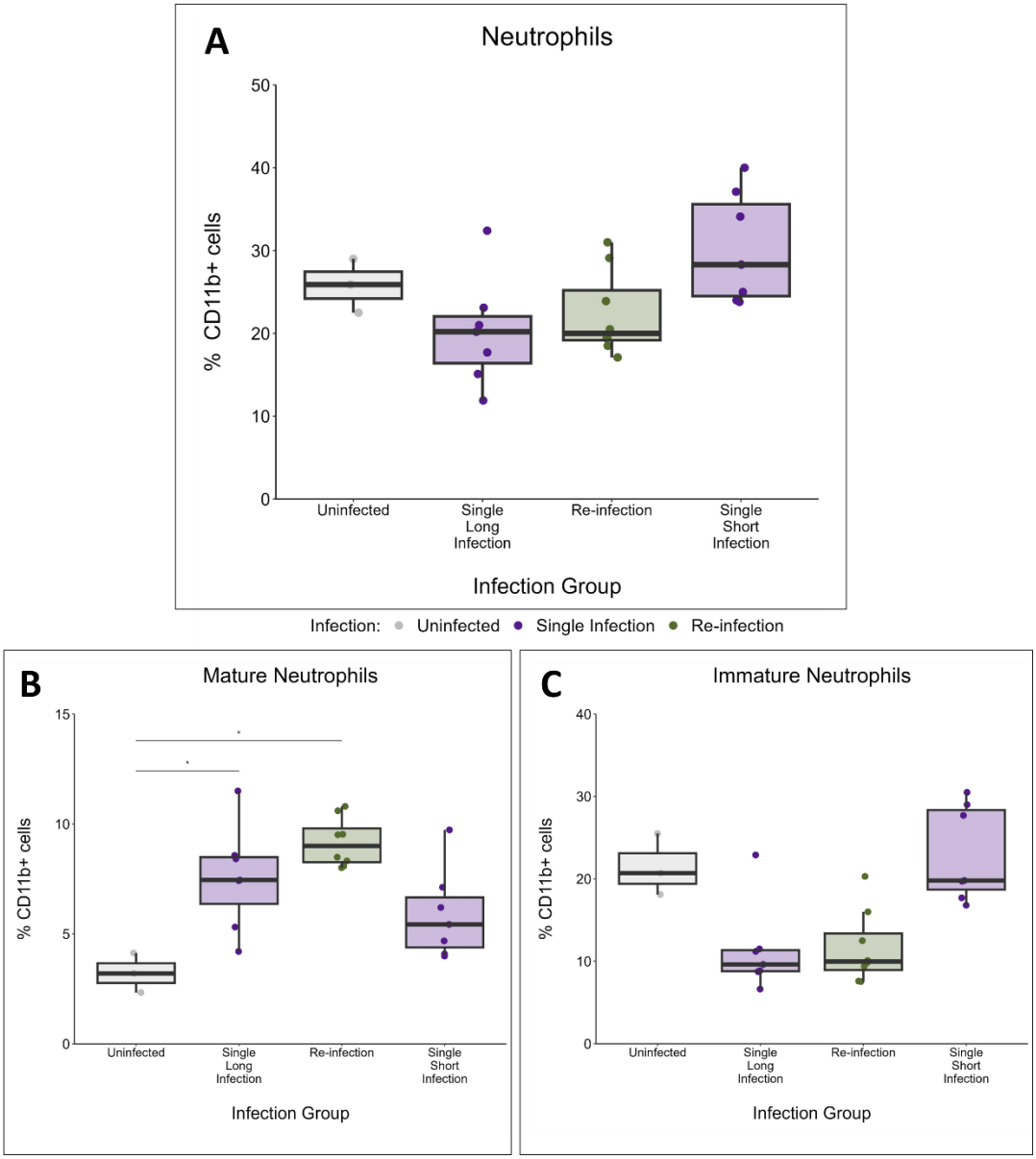


Figure 6.16: Neutrophils in the livers of in female C57BL/6J mice re-infected with promastigote *L. donovani* parasites after a prior amastigote infection. Single long infection n=7; re-infection n=8; single short infection n=7. Control mice were uninfected (n=3). Livers were processed to a single cell suspension and assessed for various markers using flow cytometry. A). Total neutrophils as a percentage of CD11b+ cells. B). Mature ($Ly6G^{high}$) neutrophils as a percentage of CD11b+ cells. C). Immature ($Ly6G^{low}$) neutrophils as a percentage of CD11b+ cells. Each point represents one individual mouse liver. Colour indicates infection as shown in the key and is consistent across all panels. Mann-Whitney-Wilcoxon test with Bonferroni correction to compare infected groups to control reference, and Mann-Whitney-Wilcoxon test between single long infection and re-infection groups. Significance of (adjusted in comparisons to control) p-value: * p<0.05; **p<0.01; *** p<0.001: **** p<0.0001. Boxplots show the median, 25th and 75th quartiles, with whiskers

representing all data-point within 1.5x Interquartile Range of the median; points above or below the whiskers represent outlying data values.

In contrast to the above known host cells for *Leishmania* parasites, the presence of T cells, B cells and NK cells may indicate a greater toxic bottleneck against colonisation of the liver. Therefore, the percentage of total cells comprising these cell types was measured (Figure 6.17).For B cells (Figure 6.17A), comparing to the baseline uninfected mouse found a significant decrease in B cells in both chronic (MW: p=0.05) and short term (MW: p=0.05) single infection groups. Re-infection also resulted in a higher level of B cells than seen in long-term single infection (MW: p=0.0401), as well as a greater spread of values across the mice in that group, indicating greater heterogeneity in re-infected mice. T cells differed from the uninfected baseline only for re-infected mice (MW: p=0.036), whilst NK cells differed from baseline in both long term (MW: p=0.05) and re-infected groups (MW: p=0.036). Neither cell type showed a significant change between the chronic single infection and the re-infected mice.

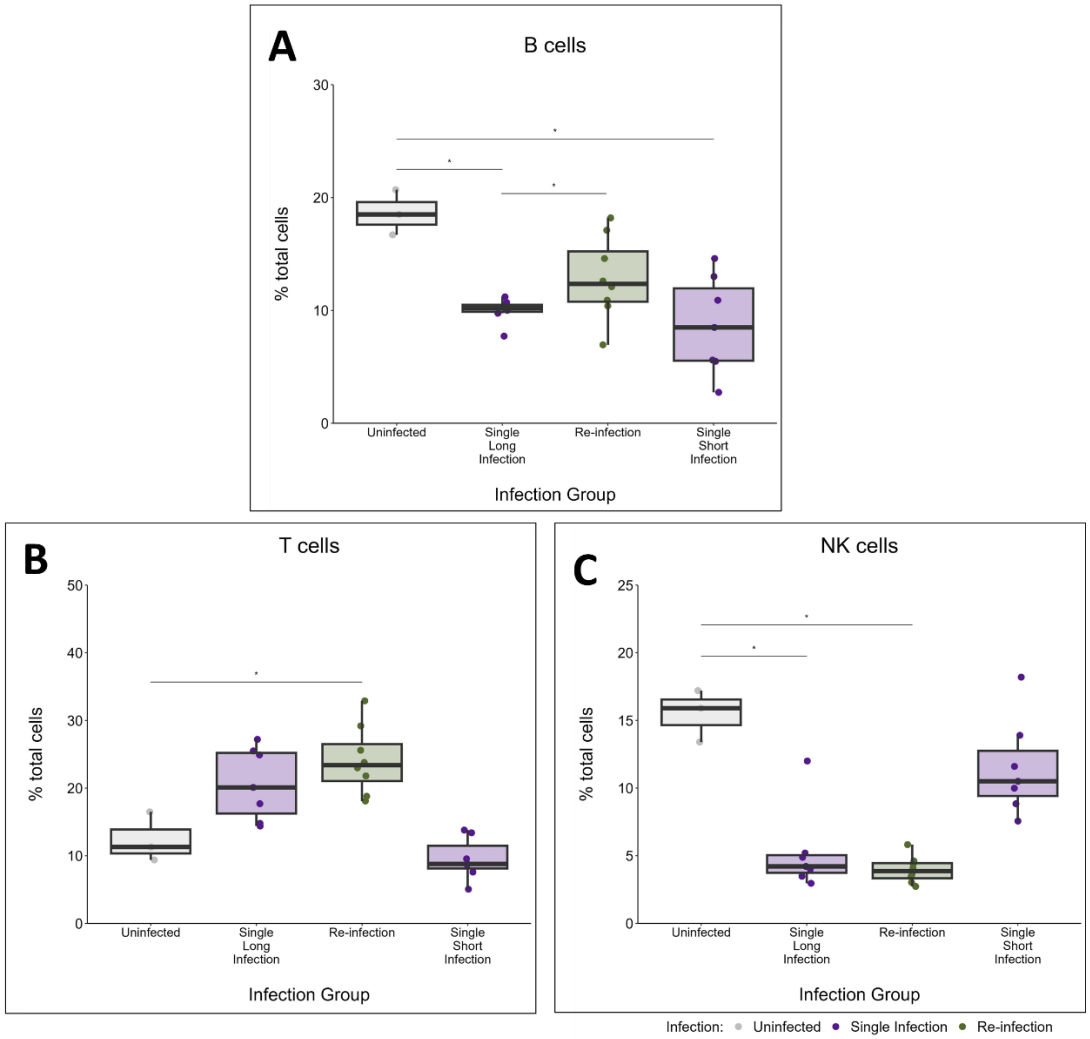


Figure 6.17: T cells, B cells and NK cells in the livers of in female C57BL/6J mice re-infected with promastigote *L. donovani* parasites after a prior amastigote infection. Single long infection n=7; re-infection n=8; single short infection n=7. Control mice were uninfected (n=3). Livers were processed to a single cell suspension and assessed for various markers using flow cytometry. A). B cells as a percentage of total cells. B). T cells as a percentage of total cells. C). NK cells as a percentage of total cells. Each point represents one individual mouse liver. Colour indicates infection as shown in the key and is consistent across all panels. Mann-Whitney-Wilcoxon test with Bonferroni correction to compare infected groups to control reference, and Mann-Whitney-Wilcoxon test between single long infection and re-infection groups. Significance of (adjusted in comparisons to control) p-value: * p<0.05; **p<0.01; *** p<0.001: **** p<0.0001. Boxplots show the median, 25th and 75th quartiles, with whiskers representing all data-point within 1.5x Interquartile Range of the median; points above or below the whiskers represent outlying data values.

Overall, these data show that there are no significant differences between re-infected and long-term single infection mouse livers in terms of macrophages, neutrophils, DCs or monocytes at 2 weeks post re-infection, providing little indication of how parasites may traffic back to the liver. The only cell type found to differ after re-infection was B cells, which appear in greater numbers in the liver when re-challenged. There is also evidence that long-term infection may alter the proportion of mature neutrophils present in the liver. Other differences in the adaptive immune response in the long-term infections indicate that a response to chronic infection is being mounted in the liver, but this does not result in parasite clearance from the liver, or in other tissues, when considered alongside the data presented earlier.

6.3 Discussion:

In this study, we sought to measure the effect of a prior infection on colonisation bottlenecks at the tissue-level in a mouse VL model. Recent work on re-infection has focussed mostly on cutaneous species of *Leishmania*, particularly *L. major*, but understanding the dynamics of re-infection in visceral species such as *L. donovani* is crucial for the development of effective treatments and vaccination approaches, as well as being of interest in terms of the biology of tissue-specific immunity.

There is evidence of re-activation of prior infection in *Leishmania* experimental models and patients, and it is thought that parasite persistence is important for maintenance of immunity. One potential mechanism which can trigger re-activation of latent parasites is re-infection (Glennie et al., 2015; Mukhopadhyay et al., 2014); it was hypothesised that this may occur in a tissue-specific way in VL, due to differences in the immunological microenvironment and the parasite populations present.

The model established here uses two different parasite lines, one of which is easy to visualise using IVIS, whilst the other provides a quantitative way to measure bottlenecks and dissemination patterns *in vivo*. Combining the two therefore provides a powerful approach to understanding how prior and subsequent

infections affect each other, with the ability to develop further hypotheses on dissemination and tissue-specific immunological responses based on the findings.

Bottleneck assessment indicated that parasite colonisation is reduced in the bone marrow, liver and spleen, but not in other tissues, on re-infection compared to a primary infection; this indicates that protection is tissue-specific, with the canonically parasite-burdened tissues being more difficult to re-colonise. The liver in particular shows a dramatically restricted bottleneck acting when a prior infection has occurred. Very few differences in tissue population relatedness were observed between the single and re-infection groups, which indicates that this protective effect is unlikely to be caused by decreased trafficking to the tissue, but rather by a toxic or resource-limiting bottleneck within the tissue itself. In the case of the liver, this is most probably a toxic effect of Th1-mediated parasite control (Bankoti and Stäger, 2011) (Figure 6.18A), whilst in the other tissues it could be a result of either cell-mediated killing or lack of resources for newcomers.

Comparing FP in the 12 week amastigote-infected mice to previous work in chapter 4 shows a progressive tightening of the bottleneck in the liver, causing a reduction in parasite FP over the 4 week to 12 week interval (Figure 6.18A). This fits with well-established data showing that the liver of C57BL/6J mice is able to control parasite burden effectively in the longer term (Bankoti and Stäger, 2011). In contrast, the bottleneck in the spleen seems to become progressively more permissive when compared to earlier time-points; again, this matches multiple previous studies (Bankoti and Stäger, 2011), indicating increasing dysregulation of the immune response, with concomitant increases in parasite burdens, in the spleen over time.

More intriguingly, the comparison of single and re-infection showed few tissues differed in response to a second challenge, but the liver and lymph nodes both exhibited an increase in FP after re-infection occurred. In the liver, this rise was particularly stark, and indicates that parasites which were residing in another tissue were able to traffic back to the liver in response to the second infection. Comparing GD showed that the most likely source for these parasites was the spleen or bone marrow, both of which were more closely related to the liver after re-infection. The lymph nodes were noted to be more closely related to the spleen, liver and gut after re-infection, perhaps indicating that parasites were trafficking from the gut to

the lymph nodes specifically. One hypothesis for this was that tissue-resident T cells may have recruited monocytes which harboured parasites to the liver on re-infection. This has been shown to occur in CL secondary infection (Glennie et al., 2017), whilst Romano et al (2021) found that bone marrow T cells can activate monocytes in the presence of *L. donovani*. Re-infection could act as a trigger for liver dissemination in a CCR2-dependent manner, and this may also occur in the spleen, which could result in the pattern observed here (Figure 6.18B).

Flow cytometry analysis of livers after single infection of short or long duration and post re-infection found no differences in the levels of canonical host cells such as monocytes, macrophages and DCs compared to uninfected controls. Neutrophils were also found not to vary overall, but interestingly, the mature neutrophil subset showed an increased abundance relative to the control in both the long-term single and re-infection groups.

The only cell type found to differ between the chronic and re-infected groups was B cells, which were higher in the re-infected mouse livers. With the caveat that the gating for CD20 expression was slightly ambiguous due to compensation effects in the flow cytometry panel used, this was an unexpected finding. Whilst the role of T cells in VL, and specifically in granulomas within the liver, is well-established, the role played by B cells is murkier. B cells are known to exacerbate infection, with depletion resulting in resistance to *L. donovani* (Smelt et al., 2000), and have been shown to interact with T cells and granulomas in the liver (Moore et al., 2012). An interesting new finding from Dirkx et al. (2024) describes extracellular attachment of *Leishmania* promastigotes to bone marrow and splenic B cells; parasites can differentiate into amastigotes on the cell surface *in vitro*, and this attachment was also observed in infected mice. These parasites are capable of being transferred into other host cells (Dirkx et al., 2024), indicating that B cells may be able to traffic parasites from one tissue to another *in vivo*. A study by Stogerer et al (2023) likewise found evidence of B cell-mediated transfer of parasites *in vivo*. Possibly, amastigotes in the bone marrow and spleen are ‘picked up’ by B cells, which then travel to the liver in response to chemokine signalling, depositing the parasites with resident macrophages and other host cells in granulomas (Figure 6.18). It is also possible that rather than an influx of B cells, the data here represent a trapping of B

cells within granulomas in the re-infected mice, which potentially reform more rapidly than the granulomas of an initial infection in naïve mice. This could explain the increase in numbers seen in the re-infected livers.

Follow-up studies which assess B cell location and the formation of granulomas after re-infection via imaging, as well as cytokine expression profile, are required. Additionally, analysis of the levels of key chemokines such as CXCL12 and CXCL13 (Barinov et al., 2017) could establish how trafficking might be occurring in response to re-infection. CD20 is not highly expressed on plasma cells, so this population does not represent a memory response; analysis of CD27 and CD19 expressing populations could elucidate whether memory cells are also increasing in re-infected livers, whilst antibody levels could be measured to assess functionality.

T cell numbers in the liver were not found to differ between chronic infection and re-infected mice; it is, however, likely that differences in the relative abundance of different subsets, and changes in function, occurred. Re-infection with *L. donovani* has been shown to increase CD8+ T cells in the liver after two weeks (McElrath et al., 1988). Studies of *L. major* re-infection have shown similar roles for IFNy secreting T cells in the skin inducing rapid recruitment of effector T cells (Glennie et al., 2015), whilst Müller et al (1993) found IFNy secretion by lymphoid CD8+ T cells increased in the early stages of re-infection. Therefore, further flow cytometry, transcriptomic and immunological phenotyping of CD4 and CD8 expression and cytokine profiles would be valuable. Assessing how T cells present respond to stimulation with amastigote versus promastigote parasites could also be interesting, with potential implications for both future re-infection studies and more widely in vaccine design considerations. Comparisons between the liver and other tissues could be informative in assessing the relative roles of tissue-resident memory cells in defence against re-infection. Tissue-resident memory cells and Tregs have both been implicated in *Leishmania* pathogenesis and protection (Glennie et al., 2015; Glennie et al., 2017; Scott, 2020). Analysis may therefore reveal hitherto unknown responses to re-infection at the tissue level.

Given the time-point at which livers were analysed was 2 weeks after the re-infection event, it is likely that differences in recruitment of other cell types would have been found if this analysis had been performed earlier in the re-infection time

course; inflammatory monocytes have been shown to home to the site of *L. major* re-infection within 72 hours (Glennie et al., 2017). Therefore, any cellular contributions to bottleneck events or dissemination differences observed are likely to have occurred rapidly after the new parasite influx and were missed by this analysis.

Ex vivo tissue imaging indicates that parasites can persist in all tissues assessed to some degree, up to 12 weeks post-infection. Whether persistent parasite presence itself contributes to bottlenecks in secondary infection, via a resource-limiting effect, cannot be determined from this work, but a follow-up study which uses drug clearance of chronic parasites prior to re-infection would be valuable. If parasite presence is necessary for continued immune protection, as is suggested (Glennie et al., 2015; Peters et al., 2014; Peters and Sacks, 2006; Romano et al., 2021; Scott et al, 2004), then the protective effect in some tissues would be abrogated in this model. Understanding this is vital for the development of effective strategies to both treat and vaccinate against *Leishmania*. Nonetheless, these data show that the liver does not fully clear parasites, even when apparently controlling them via granulomas and a Th1 response.

It is also possible that parasite burdens may initially vary after re-infection, and then become more stable as homeostasis returns, and this effect would have been missed by the *ex vivo* imaging performed here. Thus, a repeat experiment which assessed both parasite tissue burdens and the immunological response with a shorter time-frame would be valuable. Likewise, studying the impact over the longer term, to see differences in tissue pathology, with an emphasis on the liver and granuloma formation, would provide clues as to how secondary infection responses differ to those of a primary infection. Only liver cells were analysed in this study, but understanding how tissues which exhibit a contrasting response in terms of parasite regulation, such as the skin, spleen and bone marrow, would also provide useful insight into tissue-specific adaptive immune responses.

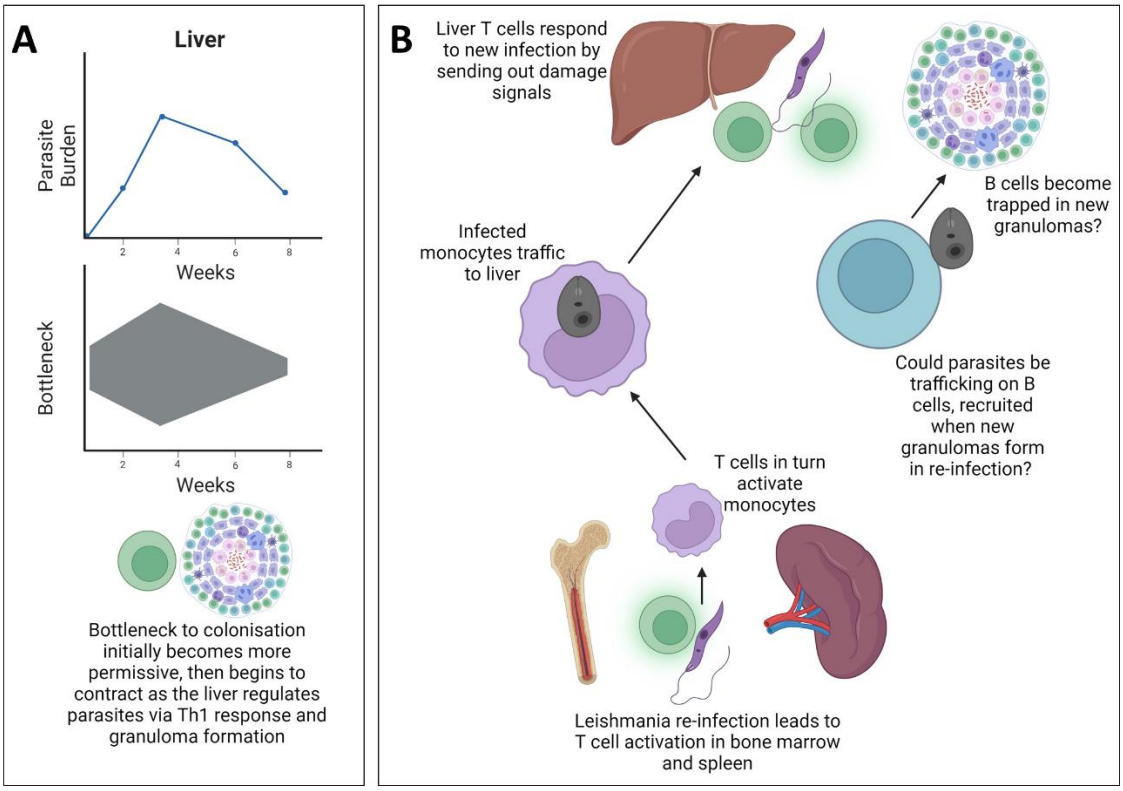


Figure 6.18: The liver in a re-infection C57BL/6J VL model. A). Liver bottlenecks over time and their relation to parasite burden. The bottleneck to colonisation becomes more permissive over the initial stages of infection, but as the liver begins to control parasite burdens via a Th1 response and granuloma formation, the bottleneck begins to tighten, reducing the FP. B). Proposed mechanism for liver re-colonisation. Latent parasites in the bone marrow and spleen may traffic to the liver after re-infection occurs via different mechanisms. One proposal is that liver-resident memory T cells recruit monocytes in response to the secondary infection, and these carry parasites to the liver from the spleen and bone marrow. Another plausible mechanism is that parasites may travel to the liver via B cells, on whose membrane they can adhere. These B cells may originate in the spleen and bone marrow, and in response to signalling when new granulomas begin to form, they traffic to the liver. Granuloma formation may occur rapidly in the case of re-infection, but needs further research. The B cells could become trapped in granuloma structures, resulting in their overall increase in the liver after re-infection. Figure created in Biorender.

The use of REH-*L. donovani* parasites in this study offered an opportunity to use *ex vivo* imaging to observe parasite burdens in the skin in short and long-term VL infection. Previous work using immunocompromised RAG mice (Doehl et al., 2021; Doehl et al., 2017) has noted a patchy distribution of parasites, but it was unclear whether any parasites would be visible in the skin of the immunocompetent

C57BL/6J mice used here. However, IVIS imaging showed that parasites are indeed present in the skin of these mice, both at 2 and 12 weeks post-infection, with a similar patchy distribution as previously seen in RAG mice (Doehl et al., 2017; Doehl et al., 2021). The same approach has been used in two other trypanosomatid parasites, with both *T. cruzi* (Ward et al., 2020) and *T. brucei* (Capewell et al., 2016) showing a similar patchy distribution; this supports the previous hypotheses of Doehl et al (2017 and 2021) that heterogeneous parasite dispersal patterns in the skin are an evolutionary adaptation to increase uptake by insect vectors. In their work, they used immuno-deficient mice and mathematical modelling to show that the heterogeneous skin landscape of *L. donovani* burden can increase transmission potential compared to a homogeneous parasite burden across the skin. In conjunction with the data on skin parasite localisation in Ward et al (2020) and Capewell et al (2016), this appears to indicate that there may be a conserved underlying mechanism driving this dispersal, despite the notable differences in other aspects of parasite behaviour, such as the extracellular nature of *T. brucei*; this raises the question of whether this phenomenon is driven by parasites themselves, rather than by host cell behaviour. However, it is possible the underlying proess may differ between each parasite species.

The mechanism for parasite homing to the skin is as yet unknown, but our study design enabled the analysis of skin FP and GD metrics for mice infected with the barcoded parasite library. These findings indicate that colonisation is likely a stochastic event, with most patches sequenced having a very low FP and variable relation to both each other and to other tissues in the same mouse. However, some patches were found to contain much greater levels of diversity, again reflecting the heterogeneity of the parasite landscape in the skin.

As our skin samples were randomly chosen, it is unknown how parasite diversity relates to parasite patch distribution, since our barcoded line could not be imaged via fluorescence or bioluminescence prior to biopsy site selection. A study with a parasite line which combines genetic barcoding and luciferase expression could examine this relationship in detail, to determine whether larger patches have greater initial diversity (higher FP) or simply grow more rapidly from a small number of progenitors. Additionally, it would allow investigation into how geographically

close patches relate to each other. Furthermore, we sampled only five patches per mouse, and as such our data may not reflect the full landscape of diversity across the skin – patches with higher diversity may be more common than our analysis suggests.

Whilst the low FPs observed may appear maladaptive in terms of diversity maintenance during wider transmission, it appears to be a commonality among many vector-borne pathogens that these organisms undergo severe bottlenecks during inter-host transmission, with multiple studies of arboviruses indicating repeated restrictions on population size occur during the viral life cycle (Weaver et al., 2021). Understanding how parasite diversity varies with patch distribution and how this relates to sand fly uptake could help to establish a model of parasite evolution and the retention of phenotypic diversity during the wider transmission cycle, by addressing whether sand flies are more liable to bite regions of higher diversity or whether parasites are experiencing a severe bottleneck during transmission, which would have implications for their future evolutionary capability. *Leishmania* parasites have been shown to rapidly and significantly alter their genome structure and ploidy during differentiation (Dumetz et al., 2017; Negreira et al., 2021) and it may be that this a compensatory mechanism to maintain diversity in the population if sand fly uptake from the skin is notably reducing this via a restrictive bottleneck.

One common thread was that parasites were not predominantly travelling to the skin from the bone marrow population, as almost all mice showed little relationship of skin patches to this tissue. The relationship between parasite diversity and patch size is as yet not established, but a study with a line which combined barcoding and luciferase expression could examine this in detail and determine whether larger patches have greater diversity or simply grow more rapidly from a small number of progenitors. Further imaging analysis could also establish how early in infection parasites are homing to the skin, and what the triggers for this might be, as well as predicting the cell types engaging in this skin trafficking. The development of a mouse model for skin dissemination in *L. donovani* could be hugely valuable for understanding and eventually preventing the development of PKDL, with its concomitant impacts on disease transmission.

In contrast to those mice with a single infection, mice which were re-infected with REH parasites showed an interesting absence of visible parasites in the skin, indicating that the skin has some protective immunity against recolonisation. Sequencing revealed that the secondary parasite population were not seeding the skin in an obviously different manner to those of a single infection. Notably, the initial infection did appear to be affected by re-infection, with increased relatedness of skin patches to the lung and gut, and a decreased relatedness to the liver, compared to those with a single chronic infection. This implies that a secondary infection can trigger the initial parasite population to traffic to the skin from the lung and gut, presumably as a by-product of movement by host cells attracted by skin cell-secreted chemokines (Figure 6.19). These results further show that parasites can still seed the skin, but apparently do not replicate, or replicate very slowly, on arrival. If the skin-resident CD4+ memory T cells which have been found in mice after *L. major* infection (Glennie et al., 2015) are also a feature of *L. donovani* infection then this may explain why parasites cannot establish in the skin (Figure 6.19); the prior parasite population may be somewhat protected by their host cells, but newcomers are exposed to T cell defences such as monocyte recruitment and subsequent parasite clearance (Glennie et al., 2017). The recruited monocytes may provide the vehicle for parasite infiltration before activation by the IFNγ-secreting CD4+ T cells triggers parasite killing activity, similar to that seen in the bone marrow (Romano et al., 2021); studying the skin landscape in a CCR2-depletion mouse model may provide answers about the role of these cells in this process. The differences in patch sizes could also simply be a function of when parasites first arrive in the skin, given the stochastic nature of this apparent homing behaviour. Alternatively, perhaps the regions where parasites are abundant enough to be visible via imaging are also deficient in these memory cells for some reason. Parasites would then exhibit the patchy distribution shown.

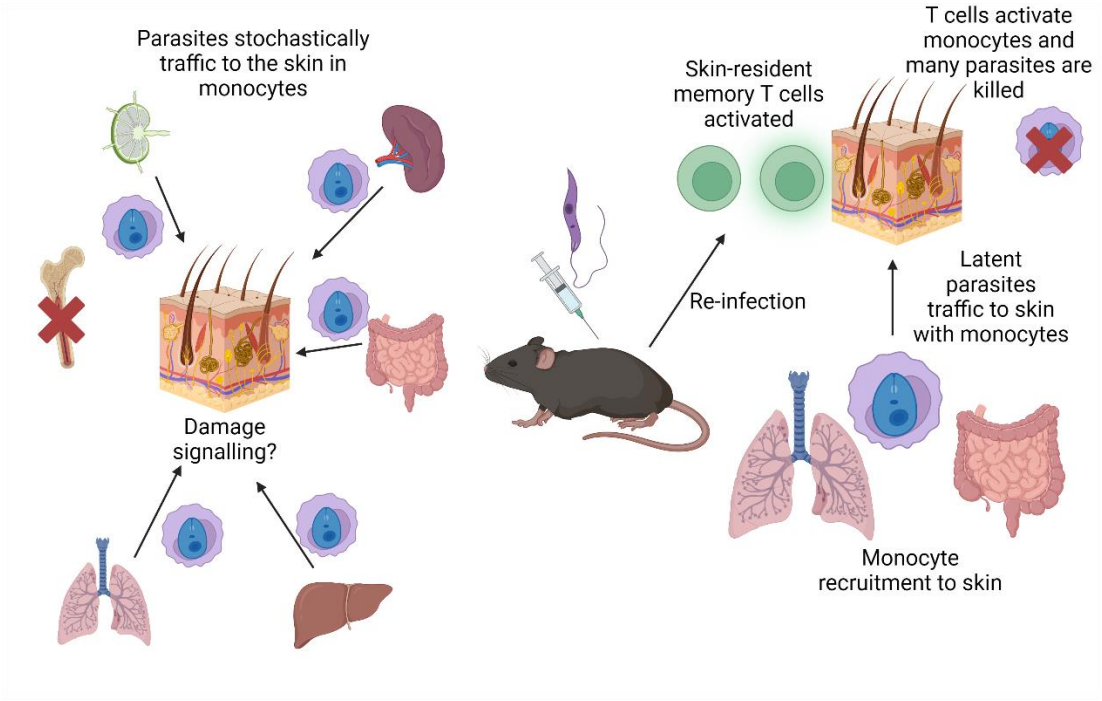


Figure 6.19: Skin trafficking proposal in a C57BL/6J VL model. The skin is colonised by parasites stochastically, with parasites seeding from various tissues, beginning early in infection. This results in a patchy distribution. Parasites can traffic to the skin within host cells such as monocytes, which may be attracted by skin-resident memory T cells signalling in response to damage. In re-infection, parasites appear to selectively move to the skin from the gut and lung. Figure created in Biorender.

This study utilised a chronic infection with amastigotes, followed by promastigote challenge; it is likely that differences might be present when re-infection is done using the same parasite stage. Therefore, repeating the study using an amastigote-amastigote and promastigote-promastigote model could be enlightening in terms of the differences in tissue-specific parasite behaviour when different life stages are used. Perhaps even more vital, this would allow elucidation of the mechanics of host response to promastigote versus amastigote parasites, which may assist in the design of future studies of re-infection and vaccine development. If the adaptive immune system responds differently to re-infection after initial promastigote challenge, then this has implications for most animal studies, which currently use amastigotes, rather than one which more closely resembles natural infection. Additionally, this may enlighten as to the tissue-specificity of recolonisation by the secondary infection; the data indicates that the lymph node, gut and lung are not protected from recolonisation, but this may be a result of low initial seeding and

subsequent immune response, as amastigotes do not appear to colonise these tissues particularly effectively (Chapter 4). Whether these tissues show a protective effect with prior promastigote infection could validate or negate this hypothesis.

Another key limitation of the work presented is that the predominant analysis was performed on females only. Males and females were noted to differ in their spleen pathology after re-infection specifically (data not shown), and previous data (Chapter 5; Klein and Flanagan, 2016) indicates differences in immune function and infection responses between the sexes. As such, repeating the above study to characterise the same prior infection population dynamics and immunological response to re-infection in male mice is an important follow-up experiment.

Furthermore, this model used a persistent infection followed by re-infection. Whether the same patterns are found when parasites have been cleared would be important to understand for establishing the mechanics of memory in VL. Ascertaining whether memory T cells and immunity from re-infection persevere without continued parasite persistence is an important step to determining the mechanisms underlying tissue-specific protection, as well as more widely for vaccine and treatment designs. Sterile clearance of *Leishmania* parasites has been notoriously difficult in animal models, even with drug regimens; using genetically-modified parasites which are highly susceptible to specific anti-microbial treatments, or which contain a life-limiting mutation (Titus et al., 1995) could provide a mechanism for this type of study.

Overall, these findings show that adaptive immune protection against VL is tissue-specific. These data reveal the novel finding that latent *L. donovani* parasites may traffic back to the liver in response to a secondary infection, moving from the spleen and bone marrow via a currently unknown mechanism. We have further shown that visceral parasites can colonise the skin of immunocompetent mice, and that these events appear to be stochastic; the skin appears to be protected against secondary population establishment. Further investigation into tissue-specific immunity is required, but these data provide a baseline for understanding how re-infection may trigger reactivation events, along with the preliminary results necessary to develop systems for PKDL research using mouse models.

Chapter 7: General Discussion:

In this work, we aimed to develop a method to analyse the population dynamics of parasites within different tissues in a murine VL model, using artificial diversity in the form of unique genetic barcodes. A library of 102 neutral barcodes were inserted into *L. donovani* promastigotes using CRISPR genome editing, and this was used successfully to measure the colonisation and dissemination dynamics of the parasite over a four week infection in C57BL/6J mice. The FP for various tissues was calculated after infection with either promastigote or amastigote parasites, with the results revealing key differences at the tissue level between these two infection models.

Together, these findings show that the initial events after infection are important in the subsequent outcomes of VL. Whilst all parasites would have differentiated to the amastigote form (or died) within two days of infection, the earliest time-point studied, there were noted differences between promastigote and amastigote infection models up to 4 weeks after the initial infection (Figure 7.1). The amastigote infection led to very little colonisation of some tissues, including the lung and gut, whilst promastigotes showed a high level of colonisation in these tissues. There were noted differences in dissemination patterns, which could be related to the differential responses of specific host cells to promastigote versus amastigote infection.

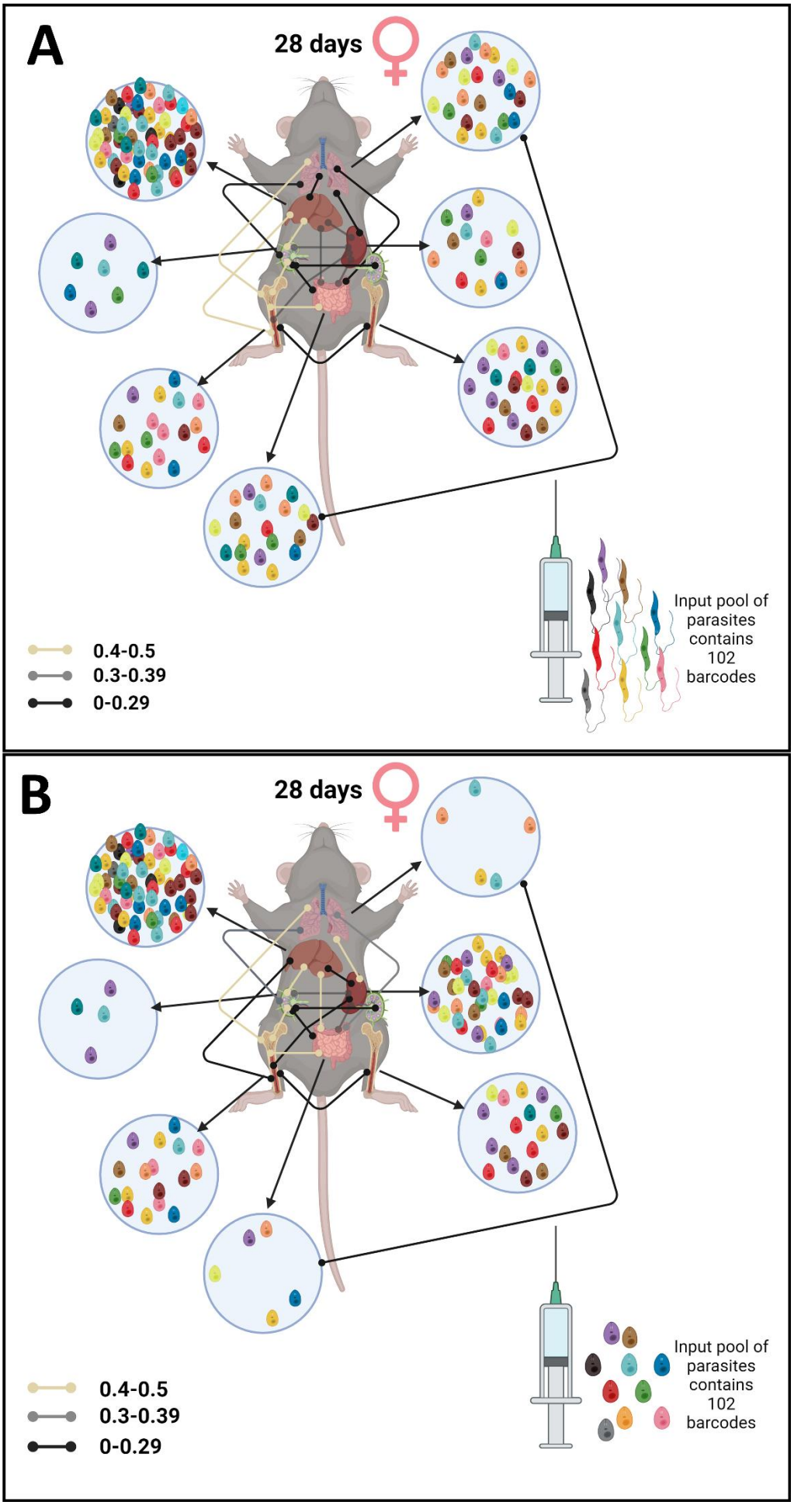

A
28 days
0.4-0.5
0.3-0.39
0-0.29
Input pool of parasites contains 102 barcodes
B
28 days
0.4-0.5
0.3-0.39
0-0.29
Input pool of parasites contains 102 barcodes

Figure 7.1: Illustration of colonising Founder Population and Genetic Distance to other tissues for the average female mouse 28 days post-infection with either amastigote or promastigote *L. donovani* parasites. A). Promastigote-infected (PRO) female at 28 days p.i, using mean founder population and genetic distance values for n=5 mice. B). Amastigote-infected (AMA) female at 28 days p.i, using mean founder population and genetic distance values for n=5 mice. Circles represent the founder population for each tissue, with number of amastigotes representing relative population size and colours indicating relative diversity (not to scale). Colours do not indicate individual barcodes and are not comparable between different mice. Lines linking two populations indicate the GD value for that pair (only values below 0.5 are shown), with black signifying a close relationship, grey an intermediate one and yellow a more distant relationship. Figure created in Biorender.

Analysis of the relationship between bottleneck (FP) and diversity (SI) metrics led to the insight that the relationship between the two apparently differed depending on the life stage at infection, with amastigotes showing the inverse of the asymptotic curve seen in the promastigote infection. One hypothesis to explain this phenomenon was a priming towards quiescence in amastigotes, due to their extensive *in vivo* passage. Quiescence has advantages in multiple microbial systems in terms of drug tolerance (as opposed to genetic mutations conferring drug resistance); pathogens can avoid succumbing to antimicrobial interventions in the short term by reducing their metabolic activity and replication rates (Barrett et al., 2019; Cabral et al., 2018; Ritterhaus et al., 2013). If amastigote parasites are induced to enter quiescence, this has implications for studies of drug resistance and tolerance, and may be useful to consider when designing experiments to investigate the mechanisms underlying these processes. Establishing how this process differs depending on tissue and local microenvironment conditions could provide valuable mechanistic information; this could involve metabolomics and single-cell dual transcriptomic or proteomic approaches to understand the relationship between quiescence and host cell behaviour, as well as environmental triggers of changes to parasite phenotype (Barrett et al., 2019). Likewise, the occurrence of PKDL and disease relapse are currently unsolved problems, with severe implications for patients and disease transmission – ascertaining how parasites enter into a drug-tolerant quiescent state, and the relationship to possible sanctuary sites (Kloehn et

al., 2021) and cells such as HSCs (Dirkx et al., 2024), could allow development of interventions to induce drug susceptibility again in drug tolerant populations, enabling parasite clearance in patients.

Dirkx et al (2024) reported that previously quiescent parasites were more infective *in vitro* and that this characteristic was maintained post-sand fly passage. Therefore, this process could also have consequences for transmission and parasite virulence, in terms of parasite evolution and disease outbreaks. Dirkx et al (2024) found that increased evolutionary adaptation was observed in quiescent parasites, and previous studies have indicated that dormancy can be important for retaining diversity in the long term, which facilitates a greater ability at the population level to respond to stressors effectively (Eisenreich et al., 2022; Rittershaus et al., 2013).

Another key insight was the influence of sex on immune regulation of parasite dissemination. Whilst females showed little effect of C3 depletion in the early stages of infection, male mice exhibited a greater impact as a result of this perturbation. It was unclear what the underlying mechanism was and this requires further investigation. The differential immune responses of males and females has been well documented (see Klein and Flannagan, 2016 for a brief review), and differences in susceptibility to leishmaniasis have likewise been established (Cloots et al., 2020; Lockard et al., 2019; Torres-Guerrero et al., 2017). It is crucial to understand how these differences are regulated and what the impact is in leishmaniasis infection, as this has implications for treatment and vaccine development. If, as hypothesised, hormonal influences are involved, this is of particular relevance in the case of patients who are pregnant or using hormonal contraceptives. Whilst VL is not noted to be especially common among those receiving hormone treatments for menopause, Polycystic Ovarian Syndrome, transgender identity or Disorders of Sexual Development, similar hormone-linked effects may be present in infections of more relevance to these populations.

Our investigation into re-infection with *L. donovani* found that expression of the protective adaptive immune response appears to be tissue-specific. Additionally, this work revealed evidence of possible re-colonisation of the liver by parasites from other tissues when mice experienced a secondary challenge. In particular, the data indicated that parasites may migrate to the liver from the spleen and bone marrow,

both tissues which have been implicated in chronic infection persistence and relapse susceptibility (Banoti and Stäger, 2011; Carter et al., 1989; Dirkx et al., 2022), as well as possible parasite quiescence (Dirkx et al., 2024); presumably, this movement is facilitated by host cell trafficking to the liver, with intracellular parasites as passengers. Our findings of potential quiescence in these two tissues, detailed in chapter 4, lead us to the hypothesis of a latent parasite population in both tissues, which can then be reactivated by immune perturbations such as a secondary infection. A latent pathogen presence which can re-activate in certain individuals is well described for the similarly macrophage-infective *M. tuberculosis* (Boom et al., 2021) and *Cryptococcus neoformans/C. gattii* (van der Torre et al., 2022). The underlying mechanism for this in our VL model is as yet unknown and a repeat study is essential to validate these findings, but this could be a reflection of the recruitment of specific immune cells which are either infected or can traffic parasites on their surface, as has recently been reported to be the case for B cells (Dirkx et al., 2024; Stogerer et al., 2023).

Further work to investigate parasite trafficking could focus on inflammatory monocytes, which have previously been implicated in parasite establishment in the liver and spleen (Terrezas et al., 2017), as well as playing a role in parasite control in the bone marrow (Romano et al., 2017). Further, Glennie et al (2017) described their indispensable function in the control of secondary *L. major* infections, via recruitment by tissue resident memory T cells. Therefore, investigating whether inflammatory monocytes can traffic to re-infected tissues and control parasite burdens in a similar way in VL infection would be prudent, and may explain the noted recolonisation of the liver discussed earlier.

Other future work of relevance in understanding dissemination could include using TNFα blockade to disrupt granuloma formation, which might assist in understanding the contribution of latent parasites to tissue-specific persistence. Utilising anti-CD4 antibodies could mimic HIV-VL co-infection and assess how this clinically important interaction alters parasite dissemination patterns and contributes to patient relapse, whilst B cell depletion could aid in ascertaining the role of these cells in tissue to tissue trafficking of parasites. However, B cells play a complicated and as

yet not fully understood role in VL, and as such, interpreting the findings of this type of study may be tricky.

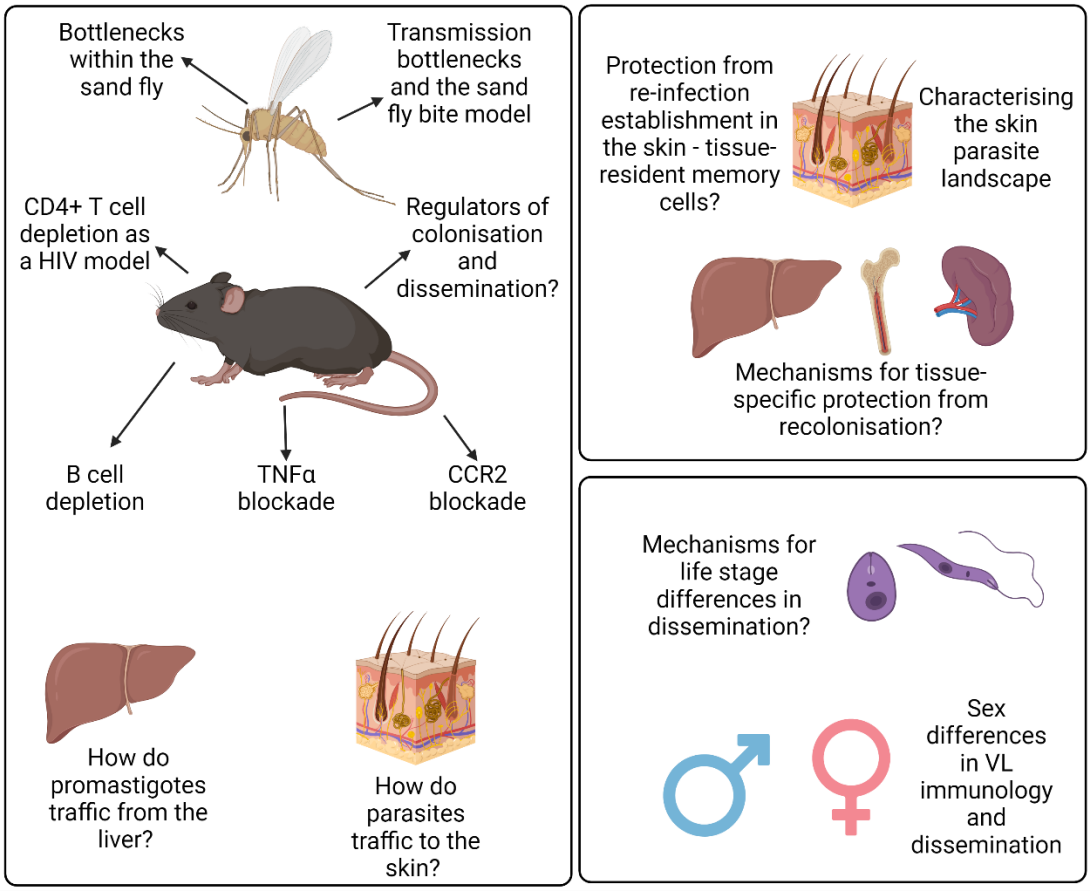


Figure 7.2: Future work to build on the findings presented. A schematic overview of follow-up experimental work, including future in vivo studies to investigate dissemination and immunological mechanisms regulating tissue bottlenecks, along with work to characterise tissue-specific immune protection to secondary challenge. Other additional work would characterise the differences in dissemination in male and female mice, and in studies using different life stages of parasite. Figure created in Biorender.

Of particular importance would be understanding how parasites traffic to the skin, as this has implications for transmission and in the development of PKDL. The use of skin homing markers to determine how immune cells are recruited would assist in developing further hypotheses for this mechanism. However, parasites are known to exhibit a patchy distribution in the skin in B6.*Rag2*$^{-/-}$ CD45.1 mice (Doehl et al., 2017; Doehl et al., 2021) and immunocompetent C57BL/6J mice (this study). Similar patterns have also been observed in other mouse models using the related parasites *T. brucei* (Capewell et al., 2016) and *T. cruzi* (Ward et al., 2020). This patchy colonisation may be an adaptive response to increase parasite uptake by sand flies (Doehl et al., 2021). Therefore, selecting the correct spots to assess for

parasite diversity may prove challenging, and could require the development of a new parasite library containing fluorescent or bioluminescent markers. Understanding how skin patch diversity relates to parasite burden would be valuable for assessing the evolutionary implications for wider disease transmission. This work could lay a framework to investigate how sand flies select a bite site (i.e. what aspect of the host skin is attractive to them?) and whether this relates in any way to the relative diversity or abundance of parasites within that region. This could have important implications for the longer-term evolution of parasites. Parasite infection may be able to induce specific metabolic signatures in host cells, which could attract insects to bite those locations containing more diverse parasites. Olfactory response to unique odours has been established to contribute to biting in mosquitos (Zhao et al., 2022), and similar mechanisms may operate in sand flies. Blood meals containing greater diversity would increase the adaptive capacity in the wider parasite population and enable more efficient responses to stressors.

One interesting new study has investigated the role of developmental endothelial locus-1 (DEL-1), a regulator of inflammation with established roles in neutrophil apoptosis and clearance (Kourtzelis et al., 2019), in VL. Lukazonek et al (2024) found that DEL-1 knock-out mice exhibited higher *L. donovani* parasite loads in the liver, but not the spleen, compared to wild-type mice. These mice also showed greater liver inflammation, which appears counter-intuitive in terms of the findings of increased parasite burden (Lukazonek et al., 2024). Population dynamics could provide a way to investigate the role played by DEL-1 in the tissue-specific immune response in more detail, using knock-out and over-expression mouse strains.

A population dynamics approach could be applied to understand how *L. donovani* and *L. infantum* dissemination patterns differ at the tissue level, as well as being applicable to CL-causative species in the future. Of particular interest would be whether the same skin homing is seen in *L. infantum*, as this species is noted not to cause PKDL. Understanding how these two differ may help enlighten the underlying mechanism for this disease.

Further, the Barseq technique has previously been applied to study gene function via genomic knock-out studies, including that by Baker et al. 2020 and Damianou et al., 2020. Combining population dynamics with this approach could aid in

understanding the relative roles played by different genes in tissue-specific virulence; using assembly process analysis could establish whether these differences are down to selective pressures or stochastic effects.

Utilising ecological methods to understand *Leishmania*-host interactions opens many potentially powerful avenues for further development. For example, combining tissue-specific dual single cell transcriptomics, metabolomics or proteomics with ecological modelling, whereby instead of barcodes the cluster signatures define the diversity within the population, could allow more in depth assessment of how the host cells interact with the parasite and how each adapts over time to the presence of the other.

Understanding parasite diversity on a spatial basis within tissues would be of great interest in elucidating how parasite replication and host cell interaction occurs in different microenvironments. How diverse parasite populations are and how they are radiating out in skin patches is important to understand for its implications in transmission and PKDL development. A key focus would be examining the distribution of parasites in prior and subsequent infections; the use of two different parasite lines, one expressing luciferase and G418 and the other expressing barcodes and BSD means this is possible via the use of antibodies specific for luciferase or G418 and BSD. For single-infection tissues, other strategies could help to determine the parasite landscape *in vivo*; the barcode insertion strategy used here was designed to allow for the possibility of spatial discrimination of barcodes via HCR, using the Zombie technique described in Askary et al., 2020. The barcode-containing repair cassettes also contained a reversed T3 promotor sequence, which enables *in vitro* transcription of barcodes for subsequent amplification using HCR probes, allowing different lines to be visualised *in situ*.

One key limitation in this work is the minimal investigation of different life stages. Whilst promastigote and amastigote infections are analysed in Chapter 4, most other analysis used promastigotes only. The culture of promastigotes did not enrich for any specific stage within that category, meaning the proportions of each culture comprising the highly infective metacyclic may have differed between experiments. Further studies could attempt to quantify the level of metacyclic versus other stages in the input culture, although this is currently difficult to do accurately, due to the

lack of stage-specific markers. Multiple studies have noted differences in gene expression profiles and infectivity between cultured and sand fly-differentiated promastigotes (reviewed in Alcolea et al., 2019). Concurrently, recent STAMP studies have assessed the role of infection route in *L. monocytogenes*, finding that oral administration versus injection results in divergent tissue colonisation patterns, which illustrates the care that must be taken when interpreting data derived using experimental models. Injection of parasites directly into a vein, as performed in our studies, is not truly reflective of the natural bite site, and this is likely to influence how parasites disseminate in the early stages of infection; our data indicates that this time is vital for determining the longer term patterns of dissemination. As such, studies of bottlenecks and dissemination dynamics using a sand fly bite as the route of infection would provide data of more direct relevance to clinical research. Using a sand fly infection would additionally provide opportunities to investigate the bottlenecks in both uptake and transmission of parasites between hosts, and to more fully appreciate the dynamics of parasite differentiation and bottleneck events within the sand fly (Figure 7.2), which are still not well understood.

Furthermore, whilst the mouse model is a standard experimental system for VL, it is not necessarily optimal. The Syrian hamster provides a closer recapitulation of late stages of human VL (Melby et al., 2001), and as such it would be prudent to apply this experimental protocol in that model system, which may be better able to identify the role of tissue tropism in disease outcomes. However, the baseline dataset presented here for the C57BL/6J model can aid in exploring the immunological mechanisms of disease, due to the wide availability of genetically-modified strains and the low genetic diversity within the inbred C57BL/6J strain.

Another limitation when interpreting the data presented was the use of different parasite doses, which, though unavoidable due to welfare concerns, means comparison between some experiments is caveated. Limited sample sizes, particularly for sex-split comparisons, also mean that some smaller effects and differences between groups may have been missed; the *ex vivo* organ imaging of re-infected and single infection mice is one example where a clear trend was observed but did not reach significance, exacerbated by the exclusion of mice for whom

infections did not establish. Since this work has provided estimates for expected effect sizes, future studies can utilise these data for more highly powered studies.

The lack of experimental confirmation of *in vivo* replication rates in *L. donovani* meant the calculation of FP relied on considering the bottleneck for each tissue as a single event. In reality, this is unlikely to be the case, particularly for longer-term infections, but generations were considered as being 1 for all bottleneck calculations. Conversely, the advantage of this method is that it provides a simple way to compare different tissues quantitatively. It also removes the problem of accounting for differences in individual replication rate, which has been evidenced to vary quite extensively in the case of amastigotes; multiple studies indicate subpopulations can enter a quiescent state with reduced replication (Dirkx et al., 2024; Kloehn et al., 2015; Kloehn et al., 2021; Mandell and Beverly, 2017). The use of computational modelling could assess the impact of successive bottleneck events on barcode distributions and compare these to experimental data for a more in depth understanding of the underlying mechanics of stressors and their role in population reduction *in vivo*.

Indeed, our work indicates that one barcoded line (B59) was potentially replicating more rapidly in the in vivo setting when compared to other lines, although this effect was not observed in culture. Whilst the removal of this barcode from the data does not hamper the computational analysis of the output (provided the removal is applied to all samples within an experiment including the input pool, which was the case for our pipeline), its inclusion in the input infection library may have altered the biological context in which the bottlenecks observed were occurring. For example, if B59 was at a selective advantage compared to other lines in specific tissue niches, this would inherently alter the colonisation and dissemination of the other lines by inducing more pressure on available resources in these tissues. This effect may not have been comparable across all tissues, time points or conditions, which would affect the interpretation of the conclusions drawn from the data. It is important to bear this possibility in mind when considering the findings detailed in the previous chapters.

Whilst some other STAMP studies in bacterial infection systems have exploited high-throughput techniques to generate much larger barcode libraries than ours

(Bachta et al., 2020; Campbell et al., 2023; Hotinger et al., 2024; Hullahalli and Waldor, 2021), the smaller library used here is sufficient to accurately determine the FP in the experiments performed. The simulations presented in chapter 3 show that small libraries can accurately calculate FP of over 10000; all FP values in all of the tissues assessed across all experiments here were smaller than this value. Larger libraries are vital for discrimination between more permissive bottlenecks (Abel et al., 2015b), but are unnecessary in systems where bottlenecks are relatively restrictive. Additionally, the library of 102 barcodes presented the advantage that it enabled application of community ecology analytic approaches, such as diversity and assembly process indices (Gotelli and Ulrich, 2012; Zhou and Ning, 2017).

In summary, this body of work provides an initial glimpse into the use of population dynamics to understand host-parasite interactions at the tissue level in a murine VL model. The data generated provide a foundation on which a variety of future studies can build, in order to understand both parasite and host factors which contribute to tissue colonisation and dissemination, as well as how these patterns can shape disease outcomes, transmission and treatment failure.

References:

Abel, S., Abel zur Wiesch, P., Chang, H.-H., *et al.* (2015) 'Sequence tag-based analysis of microbial population dynamics', *Nature methods*. Springer Science and Business Media LLC, 12(3), pp. 223–6, 3 p following 226. doi: 10.1038/nmeth.3253.

Abel, S., Abel zur Wiesch, P., Davis, B. M., *et al.* (2015b) 'Analysis of bottlenecks in experimental models of infection', *PLoS pathogens*. Public Library of Science (PLoS), 11(6), p. e1004823. doi: 10.1371/journal.ppat.1004823.

Abreu-Silva, A. L. *et al.* (2003) 'Central nervous system involvement in experimental infection with Leishmania (Leishmania) amazonensis', *The American journal of tropical medicine and hygiene*. American Society of Tropical Medicine and Hygiene, 68(6), pp. 661–665. doi: 10.4269/ajtmh.2003.68.661.

Abu-Dayyeh, I. *et al.* (2010) 'Comparative study of the ability of Leishmania mexicana promastigotes and amastigotes to alter macrophage signaling and functions', *Infection and immunity*. American Society for Microbiology, 78(6), pp. 2438–2445. doi: 10.1128/IAI.00812-09.

Akhoundi, M. *et al.* (2016) 'A historical overview of the classification, evolution, and dispersion of Leishmania parasites and sandflies', *PLoS neglected tropical diseases*. Public Library of Science (PLoS), 10(3), p. e0004349. doi: 10.1371/journal.pntd.0004349.

Akuffo, H. *et al.* (2018) 'New insights into leishmaniasis in the immunosuppressed', *PLoS neglected tropical diseases*. Public Library of Science (PLoS), 12(5), p. e0006375. doi: 10.1371/journal.pntd.0006375.

Alberdi, A. and Gilbert, M. T. P. (2019) 'A guide to the application of Hill numbers to DNA-based diversity analyses', *Molecular ecology resources*. Wiley, 19(4), pp. 804–817. doi: 10.1111/1755-0998.13014.

Alcolea, P. J. *et al.* (2014) 'Stage-specific differential gene expression in Leishmania infantum: from the foregut of Phlebotomus perniciosus to the human phagocyte', *BMC genomics*, 15, p. 849. doi: 10.1186/1471-2164-15-849.

Alcolea, P. J. *et al.* (2019) 'Functional genomics in sand fly-derived Leishmania promastigotes', *PLoS neglected tropical diseases*. Public Library of Science (PLoS), 13(5), p. e0007288. doi: 10.1371/journal.pntd.0007288.

Al-Sharif, W. Z. *et al.* (1998) 'Sea urchin coelomocytes specifically express a homologue of the complement component C3', *The journal of immunology*. The American Association of Immunologists, 160(6), pp. 2983–2997. doi: 10.4049/jimmunol.160.6.2983.

Alvarez-Jarreta, J. *et al.* (2024) 'VEuPathDB: the eukaryotic pathogen, vector and host bioinformatics resource center in 2023', *Nucleic acids research*. Oxford University Press (OUP), 52(D1), pp. D808–D816. doi: 10.1093/nar/gkad1003.

Al-Sharif, W. Z. *et al.* (1998) 'Sea urchin coelomocytes specifically express a homologue of the complement component C3', *The journal of immunology*. The American Association of Immunologists, 160(6), pp. 2983–2997. doi: 10.4049/jimmunol.160.6.2983.

Ambrosio, A. R. *et al.* (2021) 'The lectin pathway of complement and the initial recognition of Leishmania infantum promastigotes', *Life sciences*. Elsevier BV, 282(119793), p. 119793. doi: 10.1016/j.lfs.2021.119793.

Anstead, G. M. *et al.* (2001) 'Malnutrition alters the innate immune response and increases early visceralization following Leishmania donovani infection', *Infection and immunity*. American Society for Microbiology, 69(8), pp. 4709–4718. doi: 10.1128/IAI.69.8.4709-4718.2001.

Anthony, S. J. *et al.* (2015) 'Non-random patterns in viral diversity', *Nature communications*. Springer Science and Business Media LLC, 6(1), p. 8147. doi: 10.1038/ncomms9147.

Antonelli, L. R. V. *et al.* (2010) 'Intranasal Poly-IC treatment exacerbates tuberculosis in mice through the pulmonary recruitment of a pathogen-permissive monocyte/macrophage population', *The journal of clinical investigation*. American Society for Clinical Investigation, 120(5), pp. 1674–1682. doi: 10.1172/JCI40817.

Argueta-Donohué, J. *et al.* (2016) 'Differential phagocytosis of Leishmania mexicana promastigotes and amastigotes by monocyte-derived dendritic cells', *Microbiology and immunology*. Wiley, 60(6), pp. 369–381. doi: 10.1111/1348-0421.12325.

Askary, A. *et al.* (2020) 'In situ readout of DNA barcodes and single base edits facilitated by in vitro transcription', *Nature biotechnology*. Springer Science and Business Media LLC, 38(1), pp. 66–75. doi: 10.1038/s41587-019-0299-4.

Ato, M. *et al.* (2002) 'Defective CCR7 expression on dendritic cells contributes to the development of visceral leishmaniasis', *Nature immunology*. Springer Science and Business Media LLC, 3(12), pp. 1185–1191. doi: 10.1038/ni861.

Azumi, K. *et al.* (2003) 'Genomic analysis of immunity in a Urochordate and the emergence of the vertebrate immune system: "waiting for Godot"', *Immunogenetics*. Springer Science and Business Media LLC, 55(8), pp. 570–581. doi: 10.1007/s00251-003-0606-5.

Bachta, K. E. R. *et al.* (2020) 'Systemic infection facilitates transmission of Pseudomonas aeruginosa in mice', *Nature communications*. Springer Science and Business Media LLC, 11(1), p. 543. doi: 10.1038/s41467-020-14363-4.

Badia-I-Mompel, P. *et al.* (2023) 'Gene regulatory network inference in the era of single-cell multi-omics', *Nature reviews. Genetics*, 24(11), pp. 739–754. doi: 10.1038/s41576-023-00618-5.

Bahr, V. *et al.* (1993) 'Expression of lipophosphoglycan, high-molecular weight phosphoglycan and glycoprotein 63 in promastigotes and amastigotes of Leishmania mexicana', *Molecular and biochemical parasitology*. Elsevier BV, 58(1), pp. 107–121. doi: 10.1016/0166-6851(93)90095-f.

Baker, N. *et al.* (2021) 'Systematic functional analysis of Leishmania protein kinases identifies regulators of differentiation or survival', *Nature communications*. Springer Science and Business Media LLC, 12(1), p. 1244. doi: 10.1038/s41467-021-21360-8.

Baldridge, M. T. *et al.* (2010) 'Quiescent haematopoietic stem cells are activated by IFN-gamma in response to chronic infection', *Nature*. Springer Science and Business Media LLC, 465(7299), pp. 793–797. doi: 10.1038/nature09135.

Banerjee, S. *et al.* (2016) 'Attenuated Leishmania induce pro-inflammatory mediators and influence leishmanicidal activity by p38 MAPK dependent phagosome maturation in Leishmania donovani co-infected macrophages', *Scientific reports*. Springer Science and Business Media LLC, 6(1), p. 22335. doi: 10.1038/srep22335.

Bankoti, R. *et al.* (2012) 'Marginal zone B cells regulate antigen-specific T cell responses during infection', *The journal of immunology*. The American Association of Immunologists, 188(8), pp. 3961–3971. doi: 10.4049/jimmunol.1102880.

Bankoti, R. and Stäger, S. (2012) 'Differential regulation of the immune response in the spleen and liver of mice infected with Leishmania donovani', *Journal of tropical medicine*. Hindawi Limited, 2012, p. 639304. doi: 10.1155/2012/639304.

Barinov, A. *et al.* (2017) 'Essential role of immobilized chemokine CXCL12 in the regulation of the humoral immune response', *Proceedings of the National Academy of Sciences of the United States of America*. Proceedings of the National Academy of Sciences, 114(9), pp. 2319–2324. doi: 10.1073/pnas.1611958114.

Barrett, M. P. *et al.* (2019) 'Protozoan persister-like cells and drug treatment failure', *Nature reviews. Microbiology*. Springer Science and Business Media LLC, 17(10), pp. 607–620. doi: 10.1038/s41579-019-0238-x.

Batista, A. F. *et al.* (2024) 'The importance of complement-mediated immune signaling in Alzheimer's disease pathogenesis', *International journal of molecular sciences*, 25(2). doi: 10.3390/ijms25020817.

Beattie, L. *et al.* (2010) 'Dynamic imaging of experimental Leishmania donovani-induced hepatic granulomas detects Kupffer cell-restricted antigen presentation to antigen-specific CD8 T cells', *PLoS pathogens*. Public Library of Science (PLoS), 6(3), p. e1000805. doi: 10.1371/journal.ppat.1000805.

Beattie, L. *et al.* (2011) 'Interferon regulatory factor 7 contributes to the control of Leishmania donovani in the mouse liver', *Infection and immunity*. American Society for Microbiology, 79(3), pp. 1057–1066. doi: 10.1128/IAI.00633-10.

Becerra-Díaz, M. *et al.* (2021) 'Sex differences in M2 polarization, chemokine and IL-4 receptors in monocytes and macrophages from asthmatics', *Cellular immunology*. Elsevier BV, 360(104252), p. 104252. doi: 10.1016/j.cellimm.2020.104252.

Becker, I. *et al.* (2003) 'Leishmania lipophosphoglycan (LPG) activates NK cells through toll-like receptor-2', *Molecular and biochemical parasitology*. Elsevier BV, 130(2), pp. 65–74. doi: 10.1016/s0166-6851(03)00160-9.

Belkaid, Y., Butcher, B. and Sacks, D. L. (1998) 'Analysis of cytokine production by inflammatory mouse macrophages at the single-cell level: selective impairment of IL-12 induction inLeishmania -infected cells', *European journal of immunology*. Wiley, 28(4), pp. 1389–1400. doi: 10.1002/(sici)1521-4141(199804)28:04<1389::aid-immu1389>3.0.co;2-1.

Belkaid, Y. *et al.* (2002) 'CD8+ T cells are required for primary immunity in C57BL/6 mice following low-dose, intradermal challenge with Leishmania major', *The journal*

of immunology. The American Association of Immunologists, 168(8), pp. 3992–4000. doi: 10.4049/jimmunol.168.8.3992.

Belo, V. S. *et al.* (2014) 'Risk factors for adverse prognosis and death in American visceral leishmaniasis: a meta-analysis', *PLoS neglected tropical diseases*. Public Library of Science (PLoS), 8(7), p. e2982. doi: 10.1371/journal.pntd.0002982.

Beneke, T. *et al.* (2019) 'Genetic dissection of a Leishmania flagellar proteome demonstrates requirement for directional motility in sand fly infections', *PLoS pathogens*. Public Library of Science (PLoS), 15(6), p. e1007828. doi: 10.1371/journal.ppat.1007828.

Beneke, T. and Gluenz, E. (2020) 'Bar-seq strategies for the LeishGEdit toolbox', *Molecular and biochemical parasitology*. Elsevier BV, 239(111295), p. 111295. doi: 10.1016/j.molbiopara.2020.111295.

Beneke, T. *et al.* (2017) 'A CRISPR Cas9 high-throughput genome editing toolkit for kinetoplastids', *Royal Society open science*. The Royal Society, 4(5), p. 170095. doi: 10.1098/rsos.170095.

Bennis, I. *et al.* (2018) 'Psychosocial burden of localised cutaneous Leishmaniasis: a scoping review', *BMC public health*. Springer Science and Business Media LLC, 18(1). doi: 10.1186/s12889-018-5260-9.

Biggs, T. E. *et al.* (2001) 'Nramp1 modulates iron homoeostasis in vivo and in vitro: evidence for a role in cellular iron release involving de-acidification of intracellular vesicles', *European journal of immunology*. Wiley, 31(7), pp. 2060–2070. doi: 10.1002/1521-4141(200107)31:7<2060::aid-immu2060>3.0.co;2-l.

Bharati, K. (2022) 'Human genetic polymorphism and Leishmaniasis', *Infection, genetics and evolution: journal of molecular epidemiology and evolutionary genetics in infectious diseases*. Elsevier BV, 98(105203), p. 105203. doi: 10.1016/j.meegid.2021.105203.

Bispo, A. J. B. *et al.* (2020) 'Pulmonary involvement in human visceral leishmaniasis: Clinical and tomographic evaluation', *PloS one*. Public Library of Science (PLoS), 15(1), p. e0228176. doi: 10.1371/journal.pone.0228176.

Boom, W.H., Schaible, U.E. and Achkar, J.M. (2021) 'The knowns and unknowns of latent Mycobacterium tuberculosis infection', *The journal of clinical investigation*, 131(3). Available at: https://doi.org/10.1172/JCI136222.

Blackwell, J. M. *et al.* (2009) 'Genetics and visceral leishmaniasis: of mice and man', *Parasite immunology*. Wiley, 31(5), pp. 254–266. doi: 10.1111/j.1365-3024.2009.01102.x.

Blazkova, J. *et al.* (2017) 'Multicenter systems analysis of human blood reveals immature neutrophils in males and during pregnancy', *The journal of immunology*. The American Association of Immunologists, 198(6), pp. 2479–2488. doi: 10.4049/jimmunol.1601855.

Bloom, A. C. *et al.* (2016) 'Deletion of the membrane complement inhibitor CD59a drives age and gender-dependent alterations to bone phenotype in mice', *Bone*. Elsevier BV, 84, pp. 253–261. doi: 10.1016/j.bone.2015.12.014.

Bogdan, C., Mattner, J. and Schleicher, U. (2004) 'The role of type I interferons in non-viral infections', *Immunological reviews*. Wiley, 202(1), pp. 33–48. doi: 10.1111/j.0105-2896.2004.00207.x.

Bradley, D. J. (1977) 'Regulation of Leishmania populations within the host. II. genetic control of acute susceptibility of mice to Leishmania donovani infection', *Clinical and experimental immunology*, 30(1), pp. 130–140. Available at: https://www.ncbi.nlm.nih.gov/pubmed/606434.

Brandes, U. (2001) ‘A faster algorithm for betweenness centrality’, *The journal of mathematical sociology*. Informa UK Limited, 25(2), pp. 163–177. doi: 10.1080/0022250x.2001.9990249.

Brandl, K., Kumar, V. and Eckmann, L. (2017) ‘Gut-liver axis at the frontier of host-microbial interactions’, *American journal of physiology. Gastrointestinal and liver physiology*. American Physiological Society, 312(5), pp. G413–G419. doi: 10.1152/ajpgi.00361.2016.

Brittingham, A. *et al.* (1995) ‘Role of the Leishmania surface protease gp63 in complement fixation, cell adhesion, and resistance to complement-mediated lysis’, *The journal of immunology*, 155(6), pp. 3102–3111. Available at: https://www.ncbi.nlm.nih.gov/pubmed/7673725.

Brittingham, A. and Mosser, D. M. (1996) ‘Exploitation of the complement system by Leishmania promastigotes’, *Parasitology today (Personal ed.)*. Elsevier BV, 12(11), pp. 444–447. doi: 10.1016/0169-4758(96)10067-3.

Bucheton, B. *et al.* (2003) ‘Genetic control of visceral leishmaniasis in a Sudanese population: candidate gene testing indicates a linkage to the NRAMP1 region’, *Genes and immunity*. Springer Science and Business Media LLC, 4(2), pp. 104–109. doi: 10.1038/sj.gene.6363927.

Bucheton, B. *et al.* (2007) ‘Identification of a novel G245R polymorphism in the IL-2 receptor beta membrane proximal domain associated with human visceral leishmaniasis’, *Genes and immunity*. Springer Science and Business Media LLC, 8(1), pp. 79–83. doi: 10.1038/sj.gene.6364355.

Bunn, P. T. *et al.* (2014) ‘Tissue requirements for establishing long-term CD4+ T cell-mediated immunity following Leishmania donovani infection’, *The journal of immunology*. The American Association of Immunologists, 192(8), pp. 3709–3718. doi: 10.4049/jimmunol.1300768.

Burge, R. J. *et al.* (2020) 'Leishmania differentiation requires ubiquitin conjugation mediated by a UBC2-UEV1 E2 complex', *PLoS pathogens*. Public Library of Science (PLoS), 16(10), p. e1008784. doi: 10.1371/journal.ppat.1008784.

Burns, A. R. *et al.* (2016) 'Contribution of neutral processes to the assembly of gut microbial communities in the zebrafish over host development', *The ISME journal*. Oxford University Press (OUP), 10(3), pp. 655–664. doi: 10.1038/ismej.2015.142.

Burza, S., Croft, S. L. and Boelaert, M. (2018) 'Leishmaniasis', *Lancet*. Elsevier BV, 392(10151), pp. 951–970. doi: 10.1016/S0140-6736(18)31204-2.

Bushell, E. *et al.* (2017) 'Functional profiling of a Plasmodium genome reveals an abundance of essential genes', *Cell*. Elsevier BV, 170(2), pp. 260-272.e8. doi: 10.1016/j.cell.2017.06.030.

C Falcão, S. *et al.* (2012) 'The presence of Tregs does not preclude immunity to reinfection with Leishmania braziliensis', *International journal for parasitology*. Elsevier BV, 42(8), pp. 771–780. doi: 10.1016/j.ijpara.2012.05.006.

Cabral, D., Wurster, J. and Belenky, P. (2018) 'Antibiotic persistence as a metabolic adaptation: Stress, metabolism, the host, and new directions', *Pharmaceuticals (Basel, Switzerland)*. MDPI AG, 11(1), p. 14. doi: 10.3390/ph11010014.

Camelo, G. M. A. *et al.* (2023) 'Schistosoma and Leishmania: An untold story of coinfection', *Tropical medicine and infectious disease*, 8(8). doi: 10.3390/tropicalmed8080383.

Campbell, I. W. *et al.* (2023) 'Quantitative dose-response analysis untangles host bottlenecks to enteric infection', *Nature communications*. Springer Science and Business Media LLC, 14(1), p. 456. doi: 10.1038/s41467-023-36162-3.

Cañeda-Guzmán, I. C. *et al.* (2014) 'NK cell activity differs between patients with localized and diffuse cutaneous leishmaniasis infected with Leishmania mexicana: a

comparative study of TLRs and cytokines', *PloS one*. Public Library of Science (PLoS), 9(11), p. e112410. doi: 10.1371/journal.pone.0112410.

Capewell, P. *et al.* (2016) 'The skin is a significant but overlooked anatomical reservoir for vector-borne African trypanosomes', *eLife*. eLife Sciences Publications, Ltd, 5. doi: 10.7554/elife.17716.

Carlsen, E. D. *et al.* (2015) 'Permissive and protective roles for neutrophils in leishmaniasis', *Clinical and experimental immunology*. Oxford University Press (OUP), 182(2), pp. 109–118. doi: 10.1111/cei.12674.

Carnielli, J. B. T. *et al.* (2022) '3'Nucleotidase/nuclease is required for Leishmania infantum clinical isolate susceptibility to miltefosine', *EBioMedicine*. Elsevier BV, 86(104378), p. 104378. doi: 10.1016/j.ebiom.2022.104378.

Carrera, L. *et al.* (1996) 'Leishmania promastigotes selectively inhibit interleukin 12 induction in bone marrow-derived macrophages from susceptible and resistant mice', *The journal of experimental medicine*. Rockefeller University Press, 183(2), pp. 515–526. doi: 10.1084/jem.183.2.515.

Carroll, M. C. (2008) 'Complement and humoral immunity', *Vaccine*. Elsevier BV, 26, pp. I28–I33. doi: 10.1016/j.vaccine.2008.11.022.

Carroll, M. C. (2004) 'The complement system in B cell regulation', *Molecular immunology*. Elsevier BV, 41(2–3), pp. 141–146. doi: 10.1016/j.molimm.2004.03.017.

Carter, K. C. *et al.* (1989) 'Visceral leishmaniasis: resistance to reinfection in the liver following chemotherapy in the BALB/c mouse', *Experimental parasitology*. Elsevier BV, 68(4), pp. 375–381. doi: 10.1016/0014-4894(89)90122-7.

Castleman, M. J. *et al.* (2018) 'Innate sex bias of staphylococcus aureus skin infection is driven by α-hemolysin', *The journal of immunology*. The American

Association of Immunologists, 200(2), pp. 657–668. doi: 10.4049/jimmunol.1700810.

Cavalcante, R. R., Pereira, M. H. and Gontijo, N. F. (2003) 'Anti-complement activity in the saliva of phlebotomine sand flies and other haematophagous insects', *Parasitology*. Cambridge University Press (CUP), 127(1), pp. 87–93. doi: 10.1017/s0031182003003329.

Cavalli-Sforza, L. L. and Edwards, A. W. (1967) 'Phylogenetic analysis. Models and estimation procedures', *The American Journal of Human Genetics*, 19(3 Pt 1), pp. 233–257. Available at: https://www.ncbi.nlm.nih.gov/pubmed/6026583.

Chan, A. *et al.* (2021) 'The role of Leishmania GP63 in the modulation of innate inflammatory response to Leishmania major infection', *PloS one*. Public Library of Science (PLoS), 16(12), p. e0262158. doi: 10.1371/journal.pone.0262158.

Chang, K. P. (1981) 'Leishmanicidal mechanisms of human polymorphonuclear phagocytes', *The American journal of tropical medicine and hygiene*. American Society of Tropical Medicine and Hygiene, 30(2), pp. 322–333. doi: 10.4269/ajtmh.1981.30.322.

Chaparro, V. *et al.* (2022) 'Transcriptional profiling of macrophages reveals distinct parasite stage-driven signatures during early infection by Leishmania donovani', *Scientific reports*. Springer Science and Business Media LLC, 12(1), p. 6369. doi: 10.1038/s41598-022-10317-6.

Chaparro, V. *et al.* (2020) 'Translational profiling of macrophages infected with Leishmania donovani identifies mTOR- and eIF4A-sensitive immune-related transcripts', *PLoS pathogens*. Public Library of Science (PLoS), 16(6), p. e1008291. doi: 10.1371/journal.ppat.1008291.

Chase, J. M. *et al.* (2011) 'Using null models to disentangle variation in community dissimilarity from variation in α-diversity', *Ecosphere (Washington, D.C)*. Wiley, 2(2), p. art24. doi: 10.1890/es10-00117.1.

Chattopadhyay, A. *et al.* (2020) 'Intestinal leishmaniasis', *Clinical microbiology and infection: the official publication of the European Society of Clinical Microbiology and Infectious Diseases*. Elsevier BV, 26(10), pp. 1345–1346. doi: 10.1016/j.cmi.2020.05.003.

Chaudhary, N. *et al.* (2022) 'A single-cell lung atlas of complement genes identifies the mesothelium and epithelium as prominent sources of extrahepatic complement proteins', *Mucosal immunology*. Elsevier BV, 15(5), pp. 927–939. doi: 10.1038/s41385-022-00534-7.

Chevée, V. *et al.* (2024) 'Temporal and spatial dynamics of Listeria monocytogenes central nervous system infection in mice', *Proceedings of the National Academy of Sciences of the United States of America*, 121(17), p. e2320311121. doi: 10.1073/pnas.2320311121.

Chi, L. *et al.* (2024) 'Sexual dimorphism in skin immunity is mediated by an androgen-ILC2-dendritic cell axis', *Science (New York, N.Y.)*, 384(6692), p. eadk6200. doi: 10.1126/science.adk6200.

Chng, K. R. *et al.* (2020) 'Metagenome-wide association analysis identifies microbial determinants of post-antibiotic ecological recovery in the gut', *Nature ecology & evolution*. Springer Science and Business Media LLC, 4(9), pp. 1256–1267. doi: 10.1038/s41559-020-1236-0.

Cira, N. J., Pearce, M. T. and Quake, S. R. (2018) 'Neutral and selective dynamics in a synthetic microbial community', *Proceedings of the National Academy of Sciences of the United States of America*. Proceedings of the National Academy of Sciences, 115(42), pp. E9842–E9848. doi: 10.1073/pnas.1808118115.

Cloots, K. *et al.* (2020) 'Male predominance in reported Visceral Leishmaniasis cases: Nature or nurture? A comparison of population-based with health facility-reported data', *PLoS neglected tropical diseases*. Public Library of Science (PLoS), 14(1), p. e0007995. doi: 10.1371/journal.pntd.0007995.

Contreras, I. *et al.* (2014) 'Impact of Leishmania mexicana infection on dendritic cell signaling and functions', *PLoS neglected tropical diseases*. Public Library of Science (PLoS), 8(9), p. e3202. doi: 10.1371/journal.pntd.0003202.

Costa, C. H. N. *et al.* (2010) 'Is severe visceral leishmaniasis a systemic inflammatory response syndrome? A case control study', *Revista da Sociedade Brasileira de Medicina Tropical*. FapUNIFESP (SciELO), 43(4), pp. 386–392. doi: 10.1590/s0037-86822010000400010.

Cotterell, S. E., Engwerda, C. R. and Kaye, P. M. (2000) 'Enhanced hematopoietic activity accompanies parasite expansion in the spleen and bone marrow of mice infected with Leishmania donovani', *Infection and immunity*. American Society for Microbiology, 68(4), pp. 1840–1848. doi: 10.1128/IAI.68.4.1840-1848.2000.

Ćuruvija, I. *et al.* (2017) 'Sex differences in macrophage functions in middle-aged rats: Relevance of estradiol level and macrophage estrogen receptor expression', *Inflammation*, 40(3), pp. 1087–1101. doi: 10.1007/s10753-017-0551-3.

Dalakas, E. *et al.* (2005) 'Hematopoietic stem cell trafficking in liver injury', *FASEB journal: official publication of the Federation of American Societies for Experimental Biology*. Wiley, 19(10), pp. 1225–1231. doi: 10.1096/fj.04-2604rev.

Damianou, A. *et al.* (2020) 'Essential roles for deubiquitination in Leishmania life cycle progression', *PLoS pathogens*. Public Library of Science (PLoS), 16(6), p. e1008455. doi: 10.1371/journal.ppat.1008455.

Dang, A. T. and Marsland, B. J. (2019) 'Microbes, metabolites, and the gut-lung axis', *Mucosal immunology*. Elsevier BV, 12(4), pp. 843–850. doi: 10.1038/s41385-019-0160-6.

Dantas-Torres, F. (2007) 'The role of dogs as reservoirs of Leishmania parasites, with emphasis on Leishmania (Leishmania) infantum and Leishmania (Viannia) braziliensis', *Veterinary parasitology*. Elsevier BV, 149(3–4), pp. 139–146. doi: 10.1016/j.vetpar.2007.07.007.

Das, G. *et al.* (1999) 'Leishmania donovani infection of a susceptible host results in CD4+ T-cell apoptosis and decreased Th1 cytokine production', *Scandinavian journal of immunology*. Wiley, 49(3), pp. 307–310. doi: 10.1046/j.1365-3083.1999.00486.x.

de Araújo Albuquerque, L. P. *et al.* (2021) 'Influence of sex hormones on the immune response to leishmaniasis', *Parasite immunology*. Wiley, 43(10–11), p. e12874. doi: 10.1111/pim.12874.

de Carvalho, R. V. H. *et al.* (2019) 'Leishmania lipophosphoglycan triggers caspase-11 and the non-canonical activation of the NLRP3 inflammasome', *Cell reports*. Elsevier BV, 26(2), pp. 429-437.e5. doi: 10.1016/j.celrep.2018.12.047.

de Menezes, J. P., Saraiva, E. M. and da Rocha-Azevedo, B. (2016) 'The site of the bite: Leishmania interaction with macrophages, neutrophils and the extracellular matrix in the dermis', *Parasites & vectors*. Springer Science and Business Media LLC, 9(1), p. 264. doi: 10.1186/s13071-016-1540-3.

Debray, R. *et al.* (2022b) 'Priority effects in microbiome assembly', *Nature reviews. Microbiology*. Springer Science and Business Media LLC, 20(2), pp. 109–121. doi: 10.1038/s41579-021-00604-w.

Debray, R. *et al.* (2022a) 'Author Correction: Priority effects in microbiome assembly', *Nature reviews. Microbiology*. Springer Science and Business Media LLC, 20(2), p. 122. doi: 10.1038/s41579-021-00643-3.

Delobel, P. *et al.* (2022) 'CXCR2 intrinsically drives the maturation and function of neutrophils in mice', *Frontiers in immunology*. Frontiers Media SA, 13, p. 1005551. doi: 10.3389/fimmu.2022.1005551.

Deniset, J. F. *et al.* (2017) 'Splenic Ly6Ghigh mature and Ly6Gint immature neutrophils contribute to eradication of S. pneumoniae', *The journal of experimental medicine*. Rockefeller University Press, 214(5), pp. 1333–1350. doi: 10.1084/jem.20161621.

de Oca, M. M. *et al.* (2023) 'UVB modifies skin immune-stroma cross-talk and promotes effector T cell recruitment during cryptic*Leishmania donovani*infection', *bioRxiv*. doi: 10.1101/2023.02.03.526940.

Desjardins, M. and Descoteaux, A. (1997) 'Inhibition of phagolysosomal biogenesis by the *Leishmania* lipophosphoglycan', *The journal of experimental medicine*. Rockefeller University Press, 185(12), pp. 2061–2068. doi: 10.1084/jem.185.12.2061.

Desjardins, M. and Descoteaux, A. (1998) 'Survival strategies of Leishmania donovani in mammalian host macrophages', *Research in immunology*. Elsevier BV, 149(7–8), pp. 689–692. doi: 10.1016/s0923-2494(99)80040-6.

Detsika, M. G. *et al.* (2024) 'The complement cascade in lung injury and disease', *Respiratory research*, 25(1), p. 20. doi: 10.1186/s12931-023-02657-2.

Dias, B. T. *et al.* (2019) 'Neutrophil elastase promotes Leishmania donovani infection via interferon-β', *FASEB journal: official publication of the Federation of American Societies for Experimental Biology*. Wiley, 33(10), pp. 10794–10807. doi: 10.1096/fj.201900524R.

Didelot, X. *et al.* (2016) 'Within-host evolution of bacterial pathogens', *Nature reviews. Microbiology*. Springer Science and Business Media LLC, 14(3), pp. 150–162. doi: 10.1038/nrmicro.2015.13.

Dirkx, L. *et al.* (2022) 'Long-term hematopoietic stem cells as a parasite niche during treatment failure in visceral leishmaniasis', *Communications biology*. Springer Science and Business Media LLC, 5(1), p. 626. doi: 10.1038/s42003-022-03591-7.

Dirkx, L., Loyens, M., *et al.* (2024) 'Effect of Leishmania infantum infection on B cell lymphopoiesis and memory in the bone marrow and spleen', *FASEB journal: official publication of the Federation of American Societies for Experimental Biology*. Wiley, 38(16), p. e23893. doi: 10.1096/fj.202400715R.

Dirkx, L., Van Acker, S. I., *et al.* (2024) 'Long-term hematopoietic stem cells trigger quiescence in Leishmania parasites', *PLoS pathogens*. Public Library of Science (PLoS), 20(4), p. e1012181. doi: 10.1371/journal.ppat.1012181.

Doehl, J. S. P. *et al.* (2021) 'Spatial point pattern analysis identifies mechanisms shaping the skin parasite landscape in Leishmania donovani infection', *Frontiers in immunology*. Frontiers Media SA, 12, p. 795554. doi: 10.3389/fimmu.2021.795554.

Doehl, J. S. P. *et al.* (2017) 'Skin parasite landscape determines host infectiousness in visceral leishmaniasis', *Nature communications*. Springer Science and Business Media LLC, 8(1). doi: 10.1038/s41467-017-00103-8.

Dos Santos, A. G. A. *et al.* (2021) 'A new target organ of Leishmania (Viannia) braziliensis chronic infection: The intestine', *Frontiers in cellular and infection microbiology*. Frontiers Media SA, 11, p. 687499. doi: 10.3389/fcimb.2021.687499.

Dos Santos, P. F. *et al.* (2018) 'ISG15-induced IL-10 is a novel anti-inflammatory myeloid axis disrupted during active tuberculosis', *The journal of immunology*. The American Association of Immunologists, 200(4), pp. 1434–1442. doi: 10.4049/jimmunol.1701120.

Druzian, A. F. *et al.* (2015) 'Risk factors for death from visceral leishmaniasis in an urban area of Brazil', *PLoS neglected tropical diseases*. Public Library of Science (PLoS), 9(8), p. e0003982. doi: 10.1371/journal.pntd.0003982.

Duarte, M. I. *et al.* (1989) 'Interstitial pneumonitis in human visceral leishmaniasis', *Transactions of the Royal Society of Tropical Medicine and Hygiene*. Oxford University Press (OUP), 83(1), pp. 73–76. doi: 10.1016/0035-9203(89)90712-8.

Dumetz, F. *et al.* (2017) 'Modulation of aneuploidy in *Leishmania donovani* during adaptation to different *in vitro* and *in vivo* environments and its impact on gene expression', *mBio*. American Society for Microbiology, 8(3). doi: 10.1128/mbio.00599-17.

Dunkelberger, J. R. and Song, W.-C. (2010) 'Complement and its role in innate and adaptive immune responses', *Cell research*. Springer Science and Business Media LLC, 20(1), pp. 34–50. doi: 10.1038/cr.2009.139.

Ebert, D. (1998) 'Experimental evolution of parasites', *Science (New York, N.Y.)*. American Association for the Advancement of Science (AAAS), 282(5393), pp. 1432–1436. doi: 10.1126/science.282.5393.1432.

Eisenreich, W. *et al.* (2022) 'Link between antibiotic persistence and antibiotic resistance in bacterial pathogens', *Frontiers in cellular and infection microbiology*. Frontiers Media SA, 12, p. 900848. doi: 10.3389/fcimb.2022.900848.

Engstler, M. and Beneke, T. (2023) 'Gene editing and scalable functional genomic screening in Leishmania species using the CRISPR/Cas9 cytosine base editor toolbox LeishBASEedit', *eLife*. eLife Sciences Publications, Ltd, 12. doi: 10.7554/eLife.85605.

Engwerda, C. R. *et al.* (2002) 'A role for tumor necrosis factor-α in remodeling the splenic marginal zone during Leishmania donovani infection', *The American journal*

of pathology. Elsevier BV, 161(2), pp. 429–437. doi: 10.1016/s0002-9440(10)64199-5.

Engwerda, C. R. *et al.* (1998) 'Neutralization of IL-12 demonstrates the existence of discrete organ-specific phases in the control ofLeishmania donovani', *European journal of immunology*. Wiley, 28(2), pp. 669–680. doi: 10.1002/(sici)1521-4141(199802)28:02<669::aid-immu669>3.0.co;2-n.

Fang, C. *et al.* (2007) 'Complement-dependent enhancement of CD8+ T cell immunity to lymphocytic choriomeningitis virus infection in decay-accelerating factor-deficient mice', *The journal of immunology*. The American Association of Immunologists, 179(5), pp. 3178–3186. doi: 10.4049/jimmunol.179.5.3178.

Faust, K. and Raes, J. (2012) 'Microbial interactions: from networks to models', *Nature reviews. Microbiology*. Springer Science and Business Media LLC, 10(8), pp. 538–550. doi: 10.1038/nrmicro2832.

Fernandez-Ruiz, R. and Niewold, T. B. (2022) 'Type I interferons in autoimmunity', *The journal of investigative dermatology*. Elsevier BV, 142(3 Pt B), pp. 793–803. doi: 10.1016/j.jid.2021.11.031.

Ferreira, V. P. *et al.* (2016) 'SALO, a novel classical pathway complement inhibitor from saliva of the sand fly Lutzomyia longipalpis', *Scientific reports*. Springer Science and Business Media LLC, 6(1), p. 19300. doi: 10.1038/srep19300.

Filho, A. A. P. *et al.* (2021) 'Evasion of the complement system by Leishmania through the uptake of factor H, a complement regulatory protein', *Acta tropica*. Elsevier BV, 224(106152), p. 106152. doi: 10.1016/j.actatropica.2021.106152.

Fischer, D. *et al.* (2011) 'Combining climatic projections and dispersal ability: a method for estimating the responses of sandfly vector species to climate change', *PLoS neglected tropical diseases*. Public Library of Science (PLoS), 5(11), p. e1407. doi: 10.1371/journal.pntd.0001407.

Fischer, W. H. and Hugli, T. E. (1997) 'Regulation of B cell functions by C3a and C3a(desArg): suppression of TNF-alpha, IL-6, and the polyclonal immune response', *The journal of immunology*. The American Association of Immunologists, 159(9), pp. 4279–4286. doi: 10.4049/jimmunol.159.9.4279.

Forrester, N. L. *et al.* (2012) 'Vector-borne transmission imposes a severe bottleneck on an RNA virus population', *PLoS pathogens*. Public Library of Science (PLoS), 8(9), p. e1002897. doi: 10.1371/journal.ppat.1002897.

Forster, D. *et al.* (2015) 'Testing ecological theories with sequence similarity networks: marine ciliates exhibit similar geographic dispersal patterns as multicellular organisms', *BMC biology*. Springer Science and Business Media LLC, 13(1), p. 16. doi: 10.1186/s12915-015-0125-5.

Foster, J. L. and Fogleman, J. C. (1994) 'Bacterial succession in necrotic tissue of agria cactus (Stenocereus gummosus)', *Applied and environmental microbiology*. American Society for Microbiology, 60(2), pp. 619–625. doi: 10.1128/aem.60.2.619-625.1994.

Franke, E. D. *et al.* (1985) 'Growth cycle-dependent generation of complement-resistant Leishmania promastigotes', *The journal of immunology*. The American Association of Immunologists, 134(4), pp. 2713–2718. doi: 10.4049/jimmunol.134.4.2713.

Franssen, S. U. *et al.* (2021) 'Diversity and within-host evolution of Leishmania donovani from Visceral leishmaniasis patients with and without HIV coinfection in northern Ethiopia', *mBio*. American Society for Microbiology, 12(3), p. e0097121. doi: 10.1128/mBio.00971-21.

Freeman, L. C. (1979) 'Centrality in Social Networks I: Conceptual Clarification', *Social Networks*, 1, pp. 215–239.

Fukami, T. and Morin, P. J. (2003) 'Productivity-biodiversity relationships depend on the history of community assembly', *Nature*. Springer Science and Business Media LLC, 424(6947), pp. 423–426. doi: 10.1038/nature01785.

Gabriel, C. *et al.* (2010) 'Leishmania donovani promastigotes evade the antimicrobial activity of neutrophil extracellular traps', *The journal of immunology*. The American Association of Immunologists, 185(7), pp. 4319–4327. doi: 10.4049/jimmunol.1000893.

Garcia-Romo, G. S. *et al.* (2011) 'Netting neutrophils are major inducers of type I IFN production in pediatric systemic lupus erythematosus', *Science translational medicine*. American Association for the Advancement of Science (AAAS), 3(73), p. 73ra20. doi: 10.1126/scitranslmed.3001201.

Gatewood, J. D. *et al.* (2006) 'Sex chromosome complement and gonadal sex influence aggressive and parental behaviors in mice', *The Journal of neuroscience: the official journal of the Society for Neuroscience*. Society for Neuroscience, 26(8), pp. 2335–2342. doi: 10.1523/JNEUROSCI.3743-05.2006.

Genois, M.-M. *et al.* (2014) 'DNA repair pathways in trypanosomatids: from DNA repair to drug resistance', *Microbiology and molecular biology reviews: MMBR*. American Society for Microbiology, 78(1), pp. 40–73. doi: 10.1128/MMBR.00045-13.

Giang, J. *et al.* (2018) 'Complement activation in inflammatory skin diseases', *Frontiers in immunology*. Frontiers Media SA, 9. doi: 10.3389/fimmu.2018.00639.

Glennie, N. D., Volk, S. W. and Scott, P. (2017) 'Skin-resident CD4+ T cells protect against Leishmania major by recruiting and activating inflammatory monocytes', *PLoS pathogens*. Public Library of Science (PLoS), 13(4), p. e1006349. doi: 10.1371/journal.ppat.1006349.

Glennie, N. D. *et al.* (2015) 'Skin-resident memory CD4+ T cells enhance protection against Leishmania major infection', *The journal of experimental medicine*. Rockefeller University Press, 212(9), pp. 1405–1414. doi: 10.1084/jem.20142101.

Golubchik, T. *et al.* (2013) 'Within-host evolution of Staphylococcus aureus during asymptomatic carriage', *PloS one*. Public Library of Science (PLoS), 8(5), p. e61319. doi: 10.1371/journal.pone.0061319.

Gomes, A. R. *et al.* (2015) 'A genome-scale vector resource enables high-throughput reverse genetic screening in a malaria parasite', *Cell host & microbe*. Elsevier BV, 17(3), pp. 404–413. doi: 10.1016/j.chom.2015.01.014.

Gonzalez, A., Rayfield, B. and Lindo, Z. (2011) 'The disentangled bank: how loss of habitat fragments and disassembles ecological networks', *American journal of botany*. Wiley, 98(3), pp. 503–516. doi: 10.3732/ajb.1000424.

Gossage, S. M., Rogers, M. E. and Bates, P. A. (2003) 'Two separate growth phases during the development of Leishmania in sand flies: implications for understanding the life cycle', *International journal for parasitology*. Elsevier BV, 33(10), pp. 1027–1034. doi: 10.1016/s0020-7519(03)00142-5.

Gotelli, N. J. and Ulrich, W. (2012) 'Statistical challenges in null model analysis', *Oikos (Copenhagen, Denmark)*. Wiley, 121(2), pp. 171–180. doi: 10.1111/j.1600-0706.2011.20301.x.

Graham, E. B. *et al.* (2016) 'Coupling spatiotemporal community assembly processes to changes in microbial metabolism', *Frontiers in microbiology*. Frontiers Media SA, 7, p. 1949. doi: 10.3389/fmicb.2016.01949.

Graham, E. B. *et al.* (2017) 'Deterministic influences exceed dispersal effects on hydrologically-connected microbiomes', *Environmental microbiology*. Wiley, 19(4), pp. 1552–1567. doi: 10.1111/1462-2920.13720.

Grenfell, B. T. *et al.* (2004) 'Unifying the epidemiological and evolutionary dynamics of pathogens', *Science (New York, N.Y.)*. American Association for the Advancement of Science (AAAS), 303(5656), pp. 327–332. doi: 10.1126/science.1090727.

Guimarães-Costa, A. B. *et al.* (2014) '3'-nucleotidase/nuclease activity allows Leishmania parasites to escape killing by neutrophil extracellular traps', *Infection and immunity*. American Society for Microbiology, 82(4), pp. 1732–1740. doi: 10.1128/IAI.01232-13.

Guimarães-Costa, A. B. *et al.* (2009) 'Leishmania amazonensis promastigotes induce and are killed by neutrophil extracellular traps', *Proceedings of the National Academy of Sciences of the United States of America*. Proceedings of the National Academy of Sciences, 106(16), pp. 6748–6753. doi: 10.1073/pnas.0900226106.

Gupta, A. K. *et al.* (2017) 'Leishmania donovani inhibits inflammasome-dependent macrophage activation by exploiting the negative regulatory proteins A20 and UCP2', *FASEB journal: official publication of the Federation of American Societies for Experimental Biology*. Wiley, 31(11), pp. 5087–5101. doi: 10.1096/fj.201700407r.

Haile, S. and Papadopoulou, B. (2007) 'Developmental regulation of gene expression in trypanosomatid parasitic protozoa', *Current opinion in microbiology*. Elsevier BV, 10(6), pp. 569–577. doi: 10.1016/j.mib.2007.10.001.

Halliday, F. W. *et al.* (2020) 'Facilitative priority effects drive parasite assembly under coinfection', *Nature ecology & evolution*. Springer Science and Business Media LLC, 4(11), pp. 1510–1521. doi: 10.1038/s41559-020-01289-9.

Han, T. X. *et al.* (2010) 'Global fitness profiling of fission yeast deletion strains by barcode sequencing', *Genome biology*. Springer Nature, 11(6), p. R60. doi: 10.1186/gb-2010-11-6-r60.

Hanson, C. A. *et al.* (2012) 'Beyond biogeographic patterns: processes shaping the microbial landscape', *Nature reviews. Microbiology*. Springer Science and Business Media LLC, 10(7), pp. 497–506. doi: 10.1038/nrmicro2795.

Harrison, L. H. *et al.* (1986) 'Reciprocal relationships between undernutrition and the parasitic disease visceral leishmaniasis', *Clinical infectious diseases: an official publication of the Infectious Diseases Society of America*. Oxford University Press (OUP), 8(3), pp. 447–453. doi: 10.1093/clinids/8.3.447.

Hassan, M. F. *et al.* (2006) 'The Schistosoma mansoni hepatic egg granuloma provides a favorable microenvironment for sustained growth of Leishmania donovani', *The American journal of pathology*. Elsevier BV, 169(3), pp. 943–953. doi: 10.2353/ajpath.2006.051319.

Hatherell, H.-A. *et al.* (2016) 'Interpreting whole genome sequencing for investigating tuberculosis transmission: a systematic review', *BMC medicine*. Springer Science and Business Media LLC, 14(1), p. 21. doi: 10.1186/s12916-016-0566-x.

Hawksworth, O. A. *et al.* (2018) 'Complement in stem cells and development', *Seminars in immunology*. Elsevier BV, 37, pp. 74–84. doi: 10.1016/j.smim.2018.02.009.

Helmy, K. Y. *et al.* (2006) 'CRIg: a macrophage complement receptor required for phagocytosis of circulating pathogens', *Cell*. Elsevier BV, 124(5), pp. 915–927. doi: 10.1016/j.cell.2005.12.039.

Hermida, M. D. R. *et al.* (2014) 'Leishmania amazonensis infection impairs dendritic cell migration from the inflammatory site to the draining lymph node', *BMC infectious diseases*. Springer Science and Business Media LLC, 14(1), p. 450. doi: 10.1186/1471-2334-14-450.

Holt, D. S. *et al.* (2001) 'Targeted deletion of the CD59 gene causes spontaneous intravascular hemolysis and hemoglobinuria', *Blood*. American Society of Hematology, 98(2), pp. 442–449. doi: 10.1182/blood.v98.2.442.

Hill, M. O. (1973) 'Diversity and evenness: A unifying notation and its consequences', *Ecology*. Wiley, 54(2), pp. 427–432. doi: 10.2307/1934352.

Hoover, D. L. *et al.* (1984) 'Killing of Leishmania tropica amastigotes by factors in normal human serum', *The journal of immunology*. The American Association of Immunologists, 132(2), pp. 893–897. doi: 10.4049/jimmunol.132.2.893.

Hotinger, J. A. *et al.* (2024) 'Quantification of*salmonella enterica*serovar Typhimurium population dynamics in Murine infection using a highly diverse barcoded library', *bioRxiv*. doi: 10.1101/2024.06.28.601246.

Hsiao, C.-H. C. *et al.* (2008) 'The major surface protease (MSP or GP63) in the intracellular amastigote stage of Leishmania chagasi', *Molecular and biochemical parasitology*. Elsevier BV, 157(2), pp. 148–159. doi: 10.1016/j.molbiopara.2007.10.008.

Hullahalli, K. and Waldor, M. K. (2021) 'Pathogen clonal expansion underlies multiorgan dissemination and organ-specific outcomes during murine systemic infection', *eLife*. eLife Sciences Publications, Ltd, 10. doi: 10.7554/eLife.70910.

Hurrell, B.P. *et al.* (2015) 'Rapid sequestration of Leishmania mexicana by neutrophils contributes to the development of chronic lesion', *PLoS pathogens*, 11(5), p. e1004929. Available at: https://doi.org/10.1371/journal.ppat.1004929.

Iantorno, S. A. *et al.* (2017) 'Gene expression in *Leishmania* is regulated predominantly by gene dosage', *mBio*. American Society for Microbiology, 8(5). doi: 10.1128/mbio.01393-17.

Imamura, H. *et al.* (2016) 'Evolutionary genomics of epidemic visceral leishmaniasis in the Indian subcontinent', *eLife*. eLife Sciences Publications, Ltd, 5. doi: 10.7554/eLife.12613.

Inbar, E. *et al.* (2017) 'The transcriptome of*Leishmania major*developmental stages in their natural sand fly vector', *mBio*. American Society for Microbiology, 8(2). doi: 10.1128/mbio.00029-17.

Ivashkiv, L. B. and Donlin, L. T. (2014) 'Regulation of type I interferon responses', *Nature reviews. Immunology*. Springer Science and Business Media LLC, 14(1), pp. 36–49. doi: 10.1038/nri3581.

Jacobs, T. *et al.* (2005) 'Complement C3 is required for the progression of cutaneous lesions and neutrophil attraction in Leishmania major infection', *Medical microbiology and immunology*. Springer Science and Business Media LLC, 194(3), pp. 143–149. doi: 10.1007/s00430-004-0229-y.

Jalili, A. *et al.* (2010) 'Fifth complement cascade protein (C5) cleavage fragments disrupt the SDF-1/CXCR4 axis: further evidence that innate immunity orchestrates the mobilization of hematopoietic stem/progenitor cells', *Experimental hematology*. Elsevier BV, 38(4), pp. 321–332. doi: 10.1016/j.exphem.2010.02.002.

Jambulingam, P. *et al.* (2017) 'Domestic dogs as reservoir hosts for Leishmania donovani in the southernmost Western Ghats in India', *Acta tropica*. Elsevier BV, 171, pp. 64–67. doi: 10.1016/j.actatropica.2017.03.006.

Jang, M. J. *et al.* (2023) 'Spatial transcriptomics for profiling the tropism of viral vectors in tissues', *Nature biotechnology*. Springer Science and Business Media LLC, 41(9), pp. 1272–1286. doi: 10.1038/s41587-022-01648-w.

Jara, M. *et al.* (2022) 'Transcriptional shift and metabolic adaptations during Leishmania quiescence using stationary phase and drug pressure as

models', *Microorganisms*. MDPI AG, 10(1), p. 97. doi: 10.3390/microorganisms10010097.

Jasinska, W. *et al.* (2020) 'Chromosomal barcoding of E. coli populations reveals lineage diversity dynamics at high resolution', *Nature ecology & evolution*. Springer Science and Business Media LLC, 4(3), pp. 437–452. doi: 10.1038/s41559-020-1103-z.

Jawed, J. J., Dutta, S. and Majumdar, S. (2019) 'Functional aspects of T cell diversity in visceral leishmaniasis', *Biomedecine & pharmacotherapie [Biomedicine & pharmacotherapy]*. Elsevier BV, 117(109098), p. 109098. doi: 10.1016/j.biopha.2019.109098.

Jensen, J. A. *et al.* (1981) 'The complement system of the nurse shark: hemolytic and comparative characteristics', *Science (New York, N.Y.)*. American Association for the Advancement of Science (AAAS), 214(4520), pp. 566–569. doi: 10.1126/science.7291995.

Jiang, G. *et al.* (2024) 'Effects of plant tissue permeability on invasion and population bottlenecks of a phytopathogen', *Nature communications*, 15(1), p. 62. doi: 10.1038/s41467-023-44234-7.

Johnson, M. S., Venkataram, S. and Kryazhimskiy, S. (2023) 'Best practices in designing, sequencing, and identifying random DNA barcodes', *Journal of molecular evolution*. Springer Science and Business Media LLC, 91(3), pp. 263–280. doi: 10.1007/s00239-022-10083-z.

Jones, S. E. and Lennon, J. T. (2010) 'Dormancy contributes to the maintenance of microbial diversity', *Proceedings of the National Academy of Sciences of the United States of America*. Proceedings of the National Academy of Sciences, 107(13), pp. 5881–5886. doi: 10.1073/pnas.0912765107.

Jones, N. G. *et al.* (2022) 'Bromodomain factor 5 is an essential regulator of transcription in Leishmania', *Nature communications*. Springer Science and Business Media LLC, 13(1), p. 4071. doi: 10.1038/s41467-022-31742-1.

Joshi, T. *et al.* (2009) 'B7-H1 blockade increases survival of dysfunctional CD8(+) T cells and confers protection against Leishmania donovani infections', *PLoS pathogens*. Public Library of Science (PLoS), 5(5), p. e1000431. doi: 10.1371/journal.ppat.1000431.

Karagiannis, K. *et al.* (2023) 'Dual-scRNA-seq analysis reveals rare and uncommon parasitized cell populations in chronic L. donovani infection', *Cell reports*. Elsevier BV, 42(9), p. 113097. doi: 10.1016/j.celrep.2023.113097.

Kariyawasam, U. L. *et al.* (2017) 'Genetic diversity of Leishmania donovani that causes cutaneous leishmaniasis in Sri Lanka: a cross sectional study with regional comparisons', *BMC infectious diseases*. Springer Science and Business Media LLC, 17(1), p. 791. doi: 10.1186/s12879-017-2883-x.

Karmakar, S. *et al.* (2012) 'TLR4 and NKT cell synergy in immunotherapy against visceral leishmaniasis', *PLoS pathogens*. Public Library of Science (PLoS), 8(4), p. e1002646. doi: 10.1371/journal.ppat.1002646.

Karmakar, S. *et al.* (2021) 'Preclinical validation of a live attenuated dermotropic Leishmania vaccine against vector transmitted fatal visceral leishmaniasis', *Communications biology*. Springer Science and Business Media LLC, 4(1), p. 929. doi: 10.1038/s42003-021-02446-x.

Karplus, T. M. *et al.* (2002) 'Association between the tumor necrosis factor locus and the clinical outcome of Leishmania chagasi infection', *Infection and immunity*. American Society for Microbiology, 70(12), pp. 6919–6925. doi: 10.1128/IAI.70.12.6919-6925.2002.

Kaushal, H. *et al.* (2014) 'Role of CD8(+) T cells in protection against Leishmania donovani infection in healed Visceral Leishmaniasis individuals', *BMC infectious diseases*. Springer Science and Business Media LLC, 14(1), p. 653. doi: 10.1186/s12879-014-0653-6.

Kaye, P. and Scott, P. (2011) 'Leishmaniasis: complexity at the host-pathogen interface', *Nature reviews. Microbiology*. Springer Science and Business Media LLC, 9(8), pp. 604–615. doi: 10.1038/nrmicro2608.

Kaye, P. M. *et al.* (2020) 'Leishmaniasis immunopathology-impact on design and use of vaccines, diagnostics and drugs', *Seminars in immunopathology*. Springer Science and Business Media LLC, 42(3), pp. 247–264. doi: 10.1007/s00281-020-00788-y.

Kaye, P. M. *et al.* (2021) 'Overcoming roadblocks in the development of vaccines for leishmaniasis', *Expert review of vaccines*. Informa UK Limited, 20(11), pp. 1419–1430. doi: 10.1080/14760584.2021.1990043.

Kaye, P. M., Curry, A. J. and Blackwell, J. M. (1991) 'Differential production of Th1- and Th2-derived cytokines does not determine the genetically controlled or vaccine-induced rate of cure in murine visceral leishmaniasis', *The journal of immunology*. The American Association of Immunologists, 146(8), pp. 2763–2770. doi: 10.4049/jimmunol.146.8.2763.

Kennedy, A. T. *et al.* (2016) 'Recruitment of factor H as a novel complement evasion strategy for blood-stage Plasmodium falciparum infection', *The journal of immunology*. The American Association of Immunologists, 196(3), pp. 1239–1248. doi: 10.4049/jimmunol.1501581.

Khouri, R. *et al.* (2009) 'IFN-beta impairs superoxide-dependent parasite killing in human macrophages: evidence for a deleterious role of SOD1 in cutaneous leishmaniasis', *The journal of immunology*. The American Association of Immunologists, 182(4), pp. 2525–2531. doi: 10.4049/jimmunol.0802860.

Kim, A. H. J. *et al.* (2004) 'Complement C5a receptor is essential for the optimal generation of antiviral CD8+ T cell responses', *The journal of immunology*. The American Association of Immunologists, 173(4), pp. 2524–2529. doi: 10.4049/jimmunol.173.4.2524.

Kim, M., Lu, R. J. and Benayoun, B. A. (2022) 'Single-cell RNA-seq of primary bone marrow neutrophils from female and male adult mice', *Scientific data*. Springer Science and Business Media LLC, 9(1), p. 442. doi: 10.1038/s41597-022-01544-7.

Kima, P. E. and Soong, L. (2013) 'Interferon gamma in leishmaniasis', *Frontiers in immunology*. Frontiers Media SA, 4, p. 156. doi: 10.3389/fimmu.2013.00156.

Klein, S. L. and Flanagan, K. L. (2016) 'Sex differences in immune responses', *Nature reviews. Immunology*, 16(10), pp. 626–638. doi: 10.1038/nri.2016.90.

Kloehn, J. *et al.* (2021) 'Identification of metabolically quiescent Leishmania mexicana parasites in peripheral and cured dermal granulomas using stable isotope tracing imaging mass spectrometry', *mBio*. American Society for Microbiology, 12(2). doi: 10.1128/mBio.00129-21.

Kloehn, J. *et al.* (2015) 'Characterization of metabolically quiescent Leishmania parasites in murine lesions using heavy water labeling', *PLoS pathogens*. Public Library of Science (PLoS), 11(2), p. e1004683. doi: 10.1371/journal.ppat.1004683.

Kotimaa, J. *et al.* (2016) 'Sex matters: Systemic complement activity of female C57BL/6J and BALB/cJ mice is limited by serum terminal pathway components', *Molecular immunology*. Elsevier BV, 76, pp. 13–21. doi: 10.1016/j.molimm.2016.06.004.

Kotov, D. I. *et al.* (2023) 'Early cellular mechanisms of type I interferon-driven susceptibility to tuberculosis', *Cell*. Elsevier BV, 186(25), pp. 5536-5553.e22. doi: 10.1016/j.cell.2023.11.002.

Kourtzelis, I. *et al.* (2019) 'DEL-1 promotes macrophage efferocytosis and clearance of inflammation', *Nature immunology*. Springer Science and Business Media LLC, 20(1), pp. 40–49. doi: 10.1038/s41590-018-0249-1.

Kovats, S. (2015) 'Estrogen receptors regulate innate immune cells and signaling pathways', *Cellular immunology*. Elsevier BV, 294(2), pp. 63–69. doi: 10.1016/j.cellimm.2015.01.018.

Krimbas, C. B. and Tsakas, S. (1971) 'The genetics of Dacus oleae. V. Changes of esterase polymorphism in a natural population following insecticide control-selection or drift?', *Evolution; international journal of organic evolution*. Wiley, 25(3), pp. 454–460. doi: 10.1111/j.1558-5646.1971.tb01904.x.

Kropf, P. *et al.* (2004) 'Toll-like receptor 4 contributes to efficient control of infection with the protozoan parasite*Leishmania major*', *Infection and immunity*. American Society for Microbiology, 72(4), pp. 1920–1928. doi: 10.1128/iai.72.4.1920-1928.2004.

Kwon-Chung, K. J. and Chang, Y. C. (2012) 'Aneuploidy and drug resistance in pathogenic fungi', *PLoS pathogens*. Public Library of Science (PLoS), 8(11), p. e1003022. doi: 10.1371/journal.ppat.1003022.

Laffont, S. *et al.* (2014) 'X-Chromosome complement and estrogen receptor signaling independently contribute to the enhanced TLR7-mediated IFN-α production of plasmacytoid dendritic cells from women', *The journal of immunology*. The American Association of Immunologists, 193(11), pp. 5444–5452. doi: 10.4049/jimmunol.1303400.

Laguna, F. *et al.* (1994) 'Gastrointestinal leishmaniasis in human immunodeficiency virus-infected patients: report of five cases and review', *Clinical infectious diseases: an official publication of the Infectious Diseases Society of America*. Oxford University Press (OUP), 19(1), pp. 48–53. doi: 10.1093/clinids/19.1.48.

Lahav, T. *et al.* (2011) 'Multiple levels of gene regulation mediate differentiation of the intracellular pathogen*Leishmania*', *FASEB journal: official publication of the Federation of American Societies for Experimental Biology*. Wiley, 25(2), pp. 515–525. doi: 10.1096/fj.10-157529.

Laskay, T. *et al.* (1995) 'Early parasite containment is decisive for resistance to Leishmania major infection', *European journal of immunology*. Wiley, 25(8), pp. 2220–2227. doi: 10.1002/eji.1830250816.

Laufs, H. *et al.* (2002) 'Intracellular survival of Leishmania major in neutrophil granulocytes after uptake in the absence of heat-labile serum factors', *Infection and immunity*. American Society for Microbiology, 70(2), pp. 826–835. doi: 10.1128/IAI.70.2.826-835.2002.

Lebel, M.-È. *et al.* (2017) 'Complement component 3 regulates IFN-α production by plasmacytoid dendritic cells following TLR7 activation by a plant virus-like nanoparticle', *The journal of immunology*. The American Association of Immunologists, 198(1), pp. 292–299. doi: 10.4049/jimmunol.1601271.

Lee, S. M. *et al.* (2013) 'Bacterial colonization factors control specificity and stability of the gut microbiota', *Nature*. Springer Science and Business Media LLC, 501(7467), pp. 426–429. doi: 10.1038/nature12447.

Lefèvre, L. *et al.* (2013) 'The C-type lectin receptors dectin-1, MR, and SIGNR3 contribute both positively and negatively to the macrophage response to Leishmania infantum', *Immunity*. Elsevier BV, 38(5), pp. 1038–1049. doi: 10.1016/j.immuni.2013.04.010.

Lehmann, J. *et al.* (2000) 'The capacity to produce IFN-gamma rather than the presence of interleukin-4 determines the resistance and the degree of susceptibility to Leishmania donovani infection in mice', *Journal of interferon & cytokine research: the official journal of the International Society for Interferon and Cytokine Research*. Mary Ann Liebert Inc, 20(1), pp. 63–77. doi: 10.1089/107999000312748.

Lennon, J. T. and Jones, S. E. (2011) 'Microbial seed banks: the ecological and evolutionary implications of dormancy', *Nature reviews. Microbiology*. Springer Science and Business Media LLC, 9(2), pp. 119–130. doi: 10.1038/nrmicro2504.

Leprohon, P. *et al.* (2009) 'Gene expression modulation is associated with gene amplification, supernumerary chromosomes and chromosome loss in antimony-resistant Leishmania infantum', *Nucleic acids research*. Oxford University Press (OUP), 37(5), pp. 1387–1399. doi: 10.1093/nar/gkn1069.

Lestinova, T. *et al.* (2017) 'Insights into the sand fly saliva: Blood-feeding and immune interactions between sand flies, hosts, and Leishmania', *PLoS neglected tropical diseases*. Public Library of Science (PLoS), 11(7), p. e0005600. doi: 10.1371/journal.pntd.0005600.

Levin, B. R., Perrot, V. and Walker, N. (2000) 'Compensatory mutations, antibiotic resistance and the population genetics of adaptive evolution in bacteria', *Genetics*. Oxford University Press (OUP), 154(3), pp. 985–997. doi: 10.1093/genetics/154.3.985.

Levy, S. F. *et al.* (2015) 'Quantitative evolutionary dynamics using high-resolution lineage tracking', *Nature*. Springer Science and Business Media LLC, 519(7542), pp. 181–186. doi: 10.1038/nature14279.

Lewis, M. D. *et al.* (2020) 'Fatal progression of experimental visceral leishmaniasis is associated with intestinal parasitism and secondary infection by commensal bacteria, and is delayed by antibiotic prophylaxis', *PLoS pathogens*. Public Library of Science (PLoS), 16(4), p. e1008456. doi: 10.1371/journal.ppat.1008456.

Lieberman, T. D. *et al.* (2014) 'Genetic variation of a bacterial pathogen within individuals with cystic fibrosis provides a record of selective pressures', *Nature genetics*. Springer Science and Business Media LLC, 46(1), pp. 82–87. doi: 10.1038/ng.2848.

Liese, J., Schleicher, U. and Bogdan, C. (2007) 'TLR9 signaling is essential for the innate NK cell response in murine cutaneous leishmaniasis', *European journal of immunology*. Wiley, 37(12), pp. 3424–3434. doi: 10.1002/eji.200737182.

Lima-Junior, D. S. *et al.* (2013) 'Inflammasome-derived IL-1β production induces nitric oxide-mediated resistance to Leishmania', *Nature medicine*. Springer Science and Business Media LLC, 19(7), pp. 909–915. doi: 10.1038/nm.3221.

Lipoldová, M. and Demant, P. (2006) 'Genetic susceptibility to infectious disease: lessons from mouse models of leishmaniasis', *Nature reviews. Genetics*. Springer Science and Business Media LLC, 7(4), pp. 294–305. doi: 10.1038/nrg1832.

Lockard, R. D., Wilson, M. E. and Rodríguez, N. E. (2019) 'Sex-related differences in immune response and symptomatic manifestations to infection with Leishmania species', *Journal of immunology research*. Hindawi Limited, 2019, p. 4103819. doi: 10.1155/2019/4103819.

Lodge, R., Diallo, T. O. and Descoteaux, A. (2006) 'Leishmania donovani lipophosphoglycan blocks NADPH oxidase assembly at the phagosome membrane', *Cellular microbiology*. Hindawi Limited, 8(12), pp. 1922–1931. doi: 10.1111/j.1462-5822.2006.00758.x.

Loeuillet, C., Bañuls, A.-L. and Hide, M. (2016) 'Study of Leishmania pathogenesis in mice: experimental considerations', *Parasites & vectors*. Springer Science and Business Media LLC, 9(1), p. 144. doi: 10.1186/s13071-016-1413-9.

Lopes, M. E. M. *et al.* (2016) 'Indigenous microbiota and leishmaniasis', *Parasite immunology*. Wiley, 38(1), pp. 37–44. doi: 10.1111/pim.12279.

Loría-Cervera, E. N. and Andrade-Narváez, F. J. (2014) 'Animal models for the study of leishmaniasis immunology', *Revista do Instituto de Medicina Tropical de Sao*

Paulo. FapUNIFESP (SciELO), 56(1), pp. 1–11. doi: 10.1590/S0036-46652014000100001.

Lou, X. and Zhang, Y. (2013) 'Mechanism studies on nanoPCR and applications of gold nanoparticles in genetic analysis', *ACS applied materials & interfaces*. American Chemical Society (ACS), 5(13), pp. 6276–6284. doi: 10.1021/am4013209.

Louie, A. *et al.* (2019) 'A multiorgan trafficking circuit provides purifying selection of listeria monocytogenes virulence genes', *mBio*. American Society for Microbiology, 10(6). doi: 10.1128/mBio.02948-19.

Louradour, I. *et al.* (2020) 'In vitro generation of Leishmania hybrids', *Cell reports*. Elsevier BV, 31(2), p. 107507. doi: 10.1016/j.celrep.2020.03.071.

Lu, R. J. *et al.* (2021) 'Multi-omic profiling of primary mouse neutrophils predicts a pattern of sex and age-related functional regulation', *Nature aging*. Springer Science and Business Media LLC, 1(8), pp. 715–733. doi: 10.1038/s43587-021-00086-8.

Lukaszonek, J. *et al.* (2024) 'Macrophage-derived developmental endothelial locus 1 (DEL-1) expression promotes an immunoprotective phenotype in experimental visceral leishmaniasis', *bioRxiv*. doi: 10.1101/2024.04.24.587740.

Lukow, D. A. and Sheltzer, J. M. (2022) 'Chromosomal instability and aneuploidy as causes of cancer drug resistance', *Trends in cancer*. Elsevier BV, 8(1), pp. 43–53. doi: 10.1016/j.trecan.2021.09.002.

Luz, Y. *et al.* (2023) 'Leishmania infection alters macrophage and dendritic cell migration in a three-dimensional environment', *Frontiers in cell and developmental biology*, 11, p. 1206049. doi: 10.3389/fcell.2023.1206049.

Lye, L.-F. *et al.* (2010) 'Retention and loss of RNA interference pathways in trypanosomatid protozoans', *PLoS pathogens*. Public Library of Science (PLoS), 6(10), p. e1001161. doi: 10.1371/journal.ppat.1001161.

Lypaczewski, P., Zhang, W.-W. and Matlashewski, G. (2021) 'Evidence that a naturally occurring single nucleotide polymorphism in the RagC gene of Leishmania donovani contributes to reduced virulence', *PLoS neglected tropical diseases*. Public Library of Science (PLoS), 15(2), p. e0009079. doi: 10.1371/journal.pntd.0009079.

Mackey, J. B. G. *et al.* (2021) 'Maturation, developmental site, and pathology dictate murine neutrophil function'. bioRxiv. doi: 10.1101/2021.07.21.453108.

Mahoney, A. B. *et al.* (1999) 'Intra-species and stage-specific polymorphisms in lipophosphoglycan structure control Leishmania donovani-sand fly interactions', *Biochemistry*. American Chemical Society (ACS), 38(31), pp. 9813–9823. doi: 10.1021/bi990741g.

Malafaia, G. *et al.* (2009) 'Protein-energy malnutrition decreases immune response to Leishmania chagasi vaccine in BALB/c mice', *Parasite immunology*. Wiley, 31(1), pp. 41–49. doi: 10.1111/j.1365-3024.2008.01069.x.

Malarkey, D. E. *et al.* (2005) 'New insights into functional aspects of liver morphology', *Toxicologic pathology*. SAGE Publications, 33(1), pp. 27–34. doi: 10.1080/01926230590881826.

Malo, D. *et al.* (1994) 'Haplotype mapping and sequence analysis of the mouse Nramp gene predict susceptibility to infection with intracellular parasites', *Genomics*. Elsevier BV, 23(1), pp. 51–61. doi: 10.1006/geno.1994.1458.

Mandell, M. A. and Beverley, S. M. (2017) 'Continual renewal and replication of persistent*Leishmania major*parasites in concomitantly immune hosts', *Proceedings of the National Academy of Sciences of the United States of America*. Proceedings of

the National Academy of Sciences, 114(5), pp. E801–E810. doi: 10.1073/pnas.1619265114.

Mannaert, A. *et al.* (2012) 'Adaptive mechanisms in pathogens: universal aneuploidy in Leishmania', *Trends in parasitology*. Elsevier BV, 28(9), pp. 370–376. doi: 10.1016/j.pt.2012.06.003.

Marchiori, M. and Latora, V. (2000) 'Harmony in the small-World', *arXiv [cond-mat.stat-mech]*. Available at: http://arxiv.org/abs/cond-mat/0008357.

Mariotti, S. *et al.* (2004) 'Mycobacterium tuberculosis diverts alpha interferon-induced monocyte differentiation from dendritic cells into immunoprivileged macrophage-like host cells', *Infection and immunity*. American Society for Microbiology, 72(8), pp. 4385–4392. doi: 10.1128/IAI.72.8.4385-4392.2004.

Marriott, I., Bost, K. L. and Huet-Hudson, Y. M. (2006) 'Sexual dimorphism in expression of receptors for bacterial lipopolysaccharides in murine macrophages: a possible mechanism for gender-based differences in endotoxic shock susceptibility', *Journal of reproductive immunology*. Elsevier BV, 71(1), pp. 12–27. doi: 10.1016/j.jri.2006.01.004.

Martel, D. *et al.* (2017) 'Characterisation of casein kinase 1.1 inLeishmania donovaniUsing the CRISPR Cas9 toolkit', *BioMed research international*. Hindawi Limited, 2017, pp. 1–11. doi: 10.1155/2017/4635605.

Martínez, I. *et al.* (2018) 'Experimental evaluation of the importance of colonization history in early-life gut microbiota assembly', *eLife*. eLife Sciences Publications, Ltd, 7. doi: 10.7554/eLife.36521.

Masuda, C. *et al.* (2018) 'Sex-based differences in CD103+ dendritic cells promote female-predominant Th2 cytokine production during allergic asthma', *Clinical and experimental allergy: journal of the British Society for Allergy and Clinical Immunology*. Wiley, 48(4), pp. 379–393. doi: 10.1111/cea.13081.

Matte, C., Arango Duque, G. and Descoteaux, A. (2021) 'Leishmania donovani metacyclic promastigotes impair phagosome properties in inflammatory monocytes', *Infection and immunity*. American Society for Microbiology, 89(7), p. e0000921. doi: 10.1128/IAI.00009-21.

Matte, C. *et al.* (2016) 'Leishmania major promastigotes evade LC3-associated phagocytosis through the action of GP63', *PLoS pathogens*. Public Library of Science (PLoS), 12(6), p. e1005690. doi: 10.1371/journal.ppat.1005690.

Mattner, J. *et al.* (2000) 'Regulation of type 2 nitric oxide synthase by type 1 interferons in macrophages infected withLeishmania major', *European journal of immunology*. Wiley, 30(8), pp. 2257–2267. doi: 10.1002/1521-4141(2000)30:8<2257::aid-immu2257>3.0.co;2-u.

Mattner, J. *et al.* (2004) 'Protection against progressive leishmaniasis by IFN-beta', *The journal of immunology*. The American Association of Immunologists, 172(12), pp. 7574–7582. doi: 10.4049/jimmunol.172.12.7574.

McCall, L.-I., Zhang, W.-W. and Matlashewski, G. (2013) 'Determinants for the development of visceral leishmaniasis disease', *PLoS pathogens*. Public Library of Science (PLoS), 9(1), p. e1003053. doi: 10.1371/journal.ppat.1003053.

McConville, M. J. and Blackwell, J. M. (1991) 'Developmental changes in the glycosylated phosphatidylinositols of Leishmania donovani. Characterization of the promastigote and amastigote glycolipids', *The journal of biological chemistry*. Elsevier BV, 266(23), pp. 15170–15179. doi: 10.1016/s0021-9258(18)98600-x.

McElrath, M. J., Murray, H. W. and Cohn, Z. A. (1988) 'The dynamics of granuloma formation in experimental visceral leishmaniasis', *The journal of experimental medicine*. Rockefeller University Press, 167(6), pp. 1927–1937. doi: 10.1084/jem.167.6.1927.

McElrath, M. J., Murray, H. W. and Cohn, Z. A. (1987) 'The local cellular immune response in experimental leishmaniasis', *Annales de l'Institut Pasteur. Immunology*. Elsevier BV, 138(5), pp. 787–790. doi: 10.1016/s0769-2625(87)80039-9.

McFarlane, E. *et al.* (2019) 'IL-4 mediated resistance of BALB/c mice to visceral leishmaniasis is independent of IL-4Rα signaling via T cells', *Frontiers in immunology*. Frontiers Media SA, 10, p. 1957. doi: 10.3389/fimmu.2019.01957.

McNab, F. W. *et al.* (2014) 'Type I IFN induces IL-10 production in an IL-27-independent manner and blocks responsiveness to IFN-γ for production of IL-12 and bacterial killing in Mycobacterium tuberculosis-infected macrophages', *The journal of immunology*. The American Association of Immunologists, 193(7), pp. 3600–3612. doi: 10.4049/jimmunol.1401088.

Medina-Acosta, E. *et al.* (1989) 'The promastigote surface protease (gp63) of Leishmania is expressed but differentially processed and localized in the amastigote stage', *Molecular and biochemical parasitology*. Elsevier BV, 37(2), pp. 263–273. doi: 10.1016/0166-6851(89)90158-8.

Melby, P. C. *et al.* (2001) 'The hamster as a model of human visceral leishmaniasis: progressive disease and impaired generation of nitric oxide in the face of a prominent Th1-like cytokine response', *The journal of immunology*. The American Association of Immunologists, 166(3), pp. 1912–1920. doi: 10.4049/jimmunol.166.3.1912.

Melo, G. D. *et al.* (2017) 'Unveiling Cerebral Leishmaniasis: parasites and brain inflammation in Leishmania donovani infected mice', *Scientific reports*. Springer Science and Business Media LLC, 7(1), p. 8454. doi: 10.1038/s41598-017-09085-5.

Mendez, S. *et al.* (2004) 'Role for CD4(+) CD25(+) regulatory T cells in reactivation of persistent leishmaniasis and control of concomitant immunity', *The journal of experimental medicine*. Rockefeller University Press, 200(2), pp. 201–210. doi: 10.1084/jem.20040298.

Mengesha, B. *et al.* (2014) 'Prevalence of malnutrition and associated risk factors among adult visceral leishmaniasis patients in Northwest Ethiopia: a cross sectional study', *BMC research notes*. Springer Science and Business Media LLC, 7(1), p. 75. doi: 10.1186/1756-0500-7-75.

Metzemaekers, M., Gouwy, M. and Proost, P. (2020) 'Neutrophil chemoattractant receptors in health and disease: double-edged swords', *Cellular & molecular immunology*. Springer Science and Business Media LLC, 17(5), pp. 433–450. doi: 10.1038/s41423-020-0412-0.

Mevorach, D. *et al.* (1998) 'Complement-dependent clearance of apoptotic cells by human macrophages', *The journal of experimental medicine*. Rockefeller University Press, 188(12), pp. 2313–2320. doi: 10.1084/jem.188.12.2313.

Michalsky, E. M. *et al.* (2007) 'Infectivity of seropositive dogs, showing different clinical forms of leishmaniasis, to Lutzomyia longipalpis phlebotomine sand flies', *Veterinary parasitology*. Elsevier BV, 147(1–2), pp. 67–76. doi: 10.1016/j.vetpar.2007.03.004.

Miller, D. J., Hemmrich, G. and Ball, E. E. (2007) 'The innate immune repertoire in cnidaria-ancestral complexity and stochastic gene loss', *Genome Biol*, 8.

Mishra, M. *et al.* (2023) 'TIM-3 increases the abundance of type-2 dendritic cells during Leishmania donovani infection by enhancing IL-10 production via STAT3', *Cell death & disease*, 14(5), p. 331. doi: 10.1038/s41419-023-05848-3.

Mishra, J., Carpenter, S. and Singh, S. (2010) 'Low serum zinc levels in an endemic area of visceral leishmaniasis in Bihar, India', *The Indian journal of medical research*, 131, pp. 793–798. Available at: https://www.ncbi.nlm.nih.gov/pubmed/20571168.

Mohrs, M. *et al.* (1999) 'Differences between IL-4- and IL-4 receptor alpha-deficient mice in chronic leishmaniasis reveal a protective role for IL-13 receptor

signaling', *The journal of immunology*, 162(12), pp. 7302–7308. Available at: https://www.ncbi.nlm.nih.gov/pubmed/10358179.

Molloy, E. J. *et al.* (2003) 'Sex-specific alterations in neutrophil apoptosis: the role of estradiol and progesterone', *Blood*. American Society of Hematology, 102(7), pp. 2653–2659. doi: 10.1182/blood-2003-02-0649.

Montes de Oca, M., Engwerda, C. R. and Kaye, P. M. (2020) 'Cytokines and splenic remodelling during Leishmania donovani infection', *Cytokine: X*. Elsevier BV, 2(4), p. 100036. doi: 10.1016/j.cytox.2020.100036.

McNab, F. *et al.* (2015) 'Type I interferons in infectious disease', *Nature reviews. Immunology*. Springer Science and Business Media LLC, 15(2), pp. 87–103. doi: 10.1038/nri3787.

Moody, S. F. *et al.* (1993) 'The structure of Leishmania major amastigote lipophosphoglycan', *The journal of biological chemistry*. Elsevier BV, 268(25), pp. 18457–18466. doi: 10.1016/s0021-9258(17)46648-8.

Moore, J. W. J. *et al.* (2012) 'B cell: T cell interactions occur within hepatic granulomas during experimental visceral leishmaniasis', *PloS one*. Public Library of Science (PLoS), 7(3), p. e34143. doi: 10.1371/journal.pone.0034143.

Moore, J. W. J. *et al.* (2013) 'Functional complexity of the Leishmania granuloma and the potential of in silico modeling', *Frontiers in immunology*. Frontiers Media SA, 4, p. 35. doi: 10.3389/fimmu.2013.00035.

Moreira, P. O. L., Nogueira, P. M. and Monte-Neto, R. L. (2023) 'Next-generation leishmanization: Revisiting molecular targets for selecting genetically engineered live-attenuated Leishmania', *Microorganisms*, 11(4). doi: 10.3390/microorganisms11041043.

Mosser, D. M. and Brittingham, A. (1997) 'Leishmania, macrophages and complement: a tale of subversion and exploitation', *Parasitology*. Cambridge University Press (CUP), 115(7), pp. 9–23. doi: 10.1017/s0031182097001789.

Mosser, D. M. and Edelson, P. J. (1984) 'Activation of the alternative complement pathway by Leishmania promastigotes: parasite lysis and attachment to macrophages', *The journal of immunology*. The American Association of Immunologists, 132(3), pp. 1501–1505. doi: 10.4049/jimmunol.132.3.1501.

Mosser, D. M. *et al.* (1986) 'Leishmania species: mechanisms of complement activation by five strains of promastigotes', *Experimental parasitology*. Elsevier BV, 62(3), pp. 394–404. doi: 10.1016/0014-4894(86)90048-2.

Mueller, M. C. *et al.* (2009) 'Relapsing cutaneous leishmaniasis in a patient with ankylosing spondylitis treated with infliximab', *The American journal of tropical medicine and hygiene*. American Society of Tropical Medicine and Hygiene, 81(1), pp. 52–54. doi: 10.4269/ajtmh.2009.81.52.

Mukhopadhyay, D. *et al.* (2014) 'Post kala-azar dermal leishmaniasis: an unresolved mystery', *Trends in parasitology*. Elsevier BV, 30(2), pp. 65–74. doi: 10.1016/j.pt.2013.12.004.

Mullaly, S. C. and Kubes, P. (2007) 'Mast cell-expressed complement receptor, not TLR2, is the main detector of zymosan in peritonitis', *European journal of immunology*. Wiley, 37(1), pp. 224–234. doi: 10.1002/eji.200636405.

Müller, I. *et al.* (1993) 'Gamma interferon response in secondary Leishmania major infection: role of CD8+ T cells', *Infection and immunity*. American Society for Microbiology, 61(9), pp. 3730–3738. doi: 10.1128/iai.61.9.3730-3738.1993.

Müller, I. *et al.* (1994) 'Expansion of gamma interferon-producing CD8+ T cells following secondary infection of mice immune to Leishmania major', *Infection and*

immunity. American Society for Microbiology, 62(6), pp. 2575–2581. doi: 10.1128/iai.62.6.2575-2581.1994.

Murphy, M. L. *et al.* (2001) 'IL-10 mediates susceptibility toLeishmania donovani infection', *European journal of immunology*. Wiley, 31(10), pp. 2848–2856. doi: 10.1002/1521-4141(2001010)31:10<2848::aid-immu2848>3.0.co;2-t.

Murray, H. W. *et al.* (2013) 'Regulatory actions of Toll-like receptor 2 (TLR2) and TLR4 in Leishmania donovani infection in the liver', *Infection and immunity*. American Society for Microbiology, 81(7), pp. 2318–2326. doi: 10.1128/IAI.01468-12.

Murray, H. W. *et al.* (2000) 'Visceral leishmaniasis in mice devoid of tumor necrosis factor and response to treatment', *Infection and immunity*. American Society for Microbiology, 68(11), pp. 6289–6293. doi: 10.1128/IAI.68.11.6289-6293.2000.

Naik, S. *et al.* (2012) 'Compartmentalized control of skin immunity by resident commensals', *Science (New York, N.Y.)*. American Association for the Advancement of Science (AAAS), 337(6098), pp. 1115–1119. doi: 10.1126/science.1225152.

Nakanishi, H., Horii, Y. and Fujita, K. (1989) 'The role of macrophages on the expression of sex difference in the susceptibility to*Brugia pahangi*infection in C57BL/6 mice', *Journal of helminthology*. Cambridge University Press (CUP), 63(3), pp. 213–217. doi: 10.1017/s0022149x00009007.

Nandan, D., Lo, R. and Reiner, N. E. (1999) 'Activation of phosphotyrosine phosphatase activity attenuates mitogen-activated protein kinase signaling and inhibits c-FOS and nitric oxide synthase expression in macrophages infected with Leishmania donovani', *Infection and immunity*. American Society for Microbiology, 67(8), pp. 4055–4063. doi: 10.1128/IAI.67.8.4055-4063.1999.

Nandan, D. and Reiner, N. E. (1995) 'Attenuation of gamma interferon-induced tyrosine phosphorylation in mononuclear phagocytes infected with Leishmania

donovani: selective inhibition of signaling through Janus kinases and Stat1', *Infection and immunity*. American Society for Microbiology, 63(11), pp. 4495–4500. doi: 10.1128/iai.63.11.4495-4500.1995.

Napier, G. *et al.* (2022) 'Characterisation of drug-resistant Mycobacterium tuberculosis mutations and transmission in Pakistan', *Scientific reports*. Springer Science and Business Media LLC, 12(1), p. 7703. doi: 10.1038/s41598-022-11795-4.

Negreira, G. H. *et al.* (2022) 'High throughput single-cell genome sequencing gives insights into the generation and evolution of mosaic aneuploidy in Leishmania donovani', *Nucleic acids research*. Oxford University Press (OUP), 50(1), pp. 293–305. doi: 10.1093/nar/gkab1203.

Nei, M. and Tajima, F. (1981) 'Genetic drift and estimation of effective population size', *Genetics*, 98(3), pp. 625–640. doi: 10.1093/genetics/98.3.625.

Nemergut, D. R. *et al.* (2013) 'Patterns and processes of microbial community assembly', *Microbiology and molecular biology reviews: MMBR*. American Society for Microbiology, 77(3), pp. 342–356. doi: 10.1128/MMBR.00051-12.

Neves, B. M. *et al.* (2010) 'Activation of Phosphatidylinositol 3-Kinase/Akt and Impairment of Nuclear Factor-κB: Molecular Mechanisms Behind the Arrested Maturation/Activation State of Leishmania infantum-Infected Dendritic Cells', *The American Journal of Pathology*, 177, pp. 2898–2911.

Ning, D. *et al.* (2020) 'A quantitative framework reveals ecological drivers of grassland microbial community assembly in response to warming', *Nature communications*. Springer Science and Business Media LLC, 11(1), p. 4717. doi: 10.1038/s41467-020-18560-z.

Nweze, J. A., Nweze, E. I. and Onoja, U. S. (2020) 'Nutrition, malnutrition, and leishmaniasis', *Nutrition (Burbank, Los Angeles County, Calif.)*. Elsevier BV, 73(110712), p. 110712. doi: 10.1016/j.nut.2019.110712.

O'Keeffe, K. R. *et al.* (2021) 'Parasites, niche modification and the host microbiome: A field survey of multiple parasites', *Molecular ecology*. Wiley, 30(10), pp. 2404–2416. doi: 10.1111/mec.15892.

Oliveira, T. K. F. *et al.* (2023) 'Leishmania infantum axenic amastigotes induce human neutrophil extracellular traps and resist NET-mediated killing', *Tropical medicine and infectious disease*, 8(7). doi: 10.3390/tropicalmed8070336.

Omachi, S. *et al.* (2017) 'B-cell activating factor deficiency suppresses splenomegaly during Leishmania donovani infection', *Biochemical and biophysical research communications*. Elsevier BV, 489(4), pp. 528–533. doi: 10.1016/j.bbrc.2017.06.005.

Ong, H. B. *et al.* (2020) 'Establishment, optimisation and quantitation of a bioluminescent murine infection model of visceral leishmaniasis for systematic vaccine screening', *Scientific reports*. Springer Science and Business Media LLC, 10(1), p. 4689. doi: 10.1038/s41598-020-61662-3.

Ong, H. B. *et al.* (2021) 'Author Correction: Establishment, optimisation and quantitation of a bioluminescent murine infection model of visceral leishmaniasis for systematic vaccine screening', *Scientific reports*. Springer Science and Business Media LLC, 11(1), p. 8015. doi: 10.1038/s41598-021-87190-2.

O'Neil, K. M. *et al.* (1988) 'Role of C3 in humoral immunity. Defective antibody production in C3-deficient dogs', *The journal of immunology*. The American Association of Immunologists, 140(6), pp. 1939–1945. doi: 10.4049/jimmunol.140.6.1939.

Oualha, R. *et al.* (2019) 'Infection of human neutrophils with Leishmania infantum or Leishmania major strains triggers activation and differential cytokines release', *Frontiers in cellular and infection microbiology*. Frontiers Media SA, 9, p. 153. doi: 10.3389/fcimb.2019.00153.

Pace, D. (2014) ‘Leishmaniasis’, *The Journal of infection*. Elsevier BV, 69 Suppl 1, pp. S10-8. doi: 10.1016/j.jinf.2014.07.016.

Parker, S. E., Bellingham, M. C. and Woodruff, T. M. (2022) ‘Complement drives circuit modulation in the adult brain’, *Progress in neurobiology*. Elsevier BV, 214(102282), p. 102282. doi: 10.1016/j.pneurobio.2022.102282.

Parkash, V. *et al.* (2024) ‘Safety and reactogenicity of a controlled human infection model of sand fly-transmitted cutaneous leishmaniasis’, *Nature medicine*. Springer Science and Business Media LLC. doi: 10.1038/s41591-024-03146-9.

Pearson, R. D. *et al.* (1992) ‘Visceral leishmaniasis: a model for infection-induced cachexia’, *The American journal of tropical medicine and hygiene*, 47(1 Pt 2), pp. 8–15. doi: 10.4269/ajtmh.1992.47.8.

Pearson, R. D. and Steigbigel, R. T. (1981) ‘Phagocytosis and killing of the protozoan Leishmania donovani by human polymorphonuclear leukocytes’, *The journal of immunology*. The American Association of Immunologists, 127(4), pp. 1438–1443. doi: 10.4049/jimmunol.127.4.1438.

Peay, K. G., Belisle, M. and Fukami, T. (2012) ‘Phylogenetic relatedness predicts priority effects in nectar yeast communities’, *Proceedings. Biological sciences*. The Royal Society, 279(1729), pp. 749–758. doi: 10.1098/rspb.2011.1230.

Peters, C. *et al.* (1995) ‘The role of macrophage receptors in adhesion and uptake of Leishmania mexicana amastigotes’, *Journal of cell science*. The Company of Biologists, 108 (Pt 12)(12), pp. 3715–3724. doi: 10.1242/jcs.108.12.3715.

Peters, N. and Sacks, D. (2006) ‘Immune privilege in sites of chronic infection: Leishmania and regulatory T cells’, *Immunological reviews*. Wiley, 213(1), pp. 159–179. doi: 10.1111/j.1600-065X.2006.00432.x.

Peters, N. C. *et al.* (2014) 'Chronic parasitic infection maintains high frequencies of short-lived Ly6C+CD4+ effector T cells that are required for protection against re-infection', *PLoS pathogens*. Public Library of Science (PLoS), 10(12), p. e1004538. doi: 10.1371/journal.ppat.1004538.

Phillips, R. *et al.* (2010) 'Innate killing of Leishmania donovani by macrophages of the splenic marginal zone requires IRF-7', *PLoS pathogens*. Public Library of Science (PLoS), 6(3), p. e1000813. doi: 10.1371/journal.ppat.1000813.

Pokhrel, S. *et al.* (2020) 'Complement receptor 3 contributes to the sexual dimorphism in neutrophil killing of Staphylococcus aureus', *The journal of immunology*. The American Association of Immunologists, 205(6), pp. 1593–1600. doi: 10.4049/jimmunol.2000545.

Polando, R. *et al.* (2013) 'The roles of complement receptor 3 and Fcγ receptors during Leishmania phagosome maturation', *Journal of leukocyte biology*. Oxford University Press (OUP), 93(6), pp. 921–932. doi: 10.1189/jlb.0212086.

Pollak, E. (1983) 'A new method for estimating the effective population size from allele frequency changes', *Genetics*. Oxford University Press (OUP), 104(3), pp. 531–548. doi: 10.1093/genetics/104.3.531.

Ponte-Sucre, A. *et al.* (2017) 'Drug resistance and treatment failure in leishmaniasis: A 21st century challenge', *PLoS neglected tropical diseases*. Public Library of Science (PLoS), 11(12), p. e0006052. doi: 10.1371/journal.pntd.0006052.

Posch, W. *et al.* (2021) 'Complement potentiates immune sensing of HIV-1 and early type I interferon responses', *mBio*. American Society for Microbiology, 12(5), p. e0240821. doi: 10.1128/mBio.02408-21.

Potnis, N. *et al.* (2014) 'Xanthomonas perforans colonization influences Salmonella enterica in the tomato phyllosphere', *Applied and environmental microbiology*.

American Society for Microbiology, 80(10), pp. 3173–3180. doi: 10.1128/AEM.00345-14.

Potvin, J. E. *et al.* (2021) 'Mutations in an Aquaglycerporin as a Proven Marker of Antimony Clinical Resistance in the Parasite Leishmania donovani', *Clinical Infectious Diseases*, 72(10), pp. e526–e532.

Poudel, B. *et al.* (2020) 'Acute IL-4 Governs Pathogenic T Cell Responses during Leishmania major Infection', *ImmunoHorizons*. The American Association of Immunologists, 4(9), pp. 546–560. doi: 10.4049/immunohorizons.2000076.

Poza-Carrion, C., Suslow, T. and Lindow, S. (2013) 'Resident bacteria on leaves enhance survival of immigrant cells of Salmonella enterica', *Phytopathology*. Scientific Societies, 103(4), pp. 341–351. doi: 10.1094/PHYTO-09-12-0221-FI.

Pradhan, S. *et al.* (2021) 'Linking membrane fluidity with defective antigen presentation in leishmaniasis', *Parasite immunology*. Wiley, 43(7), p. e12835. doi: 10.1111/pim.12835.

Prasad, L. S. and Sen, S. (1996) 'Migration of Leishmania donovani amastigotes in the cerebrospinal fluid', *The American journal of tropical medicine and hygiene*. American Society of Tropical Medicine and Hygiene, 55(6), pp. 652–654. doi: 10.4269/ajtmh.1996.55.652.

Prat-Luri, B. *et al.* (2022) 'The C5a-C5aR1 complement axis is essential for neutrophil recruitment to draining lymph nodes via high endothelial venules in cutaneous leishmaniasis', *Cell reports*. Elsevier BV, 39(5), p. 110777. doi: 10.1016/j.celrep.2022.110777.

Puentes, S. M. *et al.* (1990) 'Serum resistance of metacyclic stage Leishmania major promastigotes is due to release of C5b-9', *The journal of immunology*. The American Association of Immunologists, 145(12), pp. 4311–4316. doi: 10.4049/jimmunol.145.12.4311.

Pujantell, M. *et al.* (2023) 'Sex-dependent differences in type I IFN-induced natural killer cell activation', *Frontiers in immunology*, 14, p. 1277967. doi: 10.3389/fimmu.2023.1277967.

Queffeulou, M. *et al.* (2024) 'CRISPR-Cas9 high-throughput screening to study drug resistance in Leishmania infantum', *mBio*. American Society for Microbiology, 15(7), p. e0047724. doi: 10.1128/mbio.00477-24.

Quinn, J. J. *et al.* (2021) 'Single-cell lineages reveal the rates, routes, and drivers of metastasis in cancer xenografts', *Science (New York, N.Y.)*. American Association for the Advancement of Science (AAAS), 371(6532), p. eabc1944. doi: 10.1126/science.abc1944.

Quintela-Carvalho, G. *et al.* (2022) 'Leishmania infantum defective in lipophosphoglycan biosynthesis interferes with activation of human neutrophils', *Frontiers in cellular and infection microbiology*. Frontiers Media SA, 12, p. 788196. doi: 10.3389/fcimb.2022.788196.

Rai, K. *et al.* (2013) 'Relapse after treatment with miltefosine for visceral leishmaniasis is associated with increased infectivity of the infecting Leishmania donovani strain', *mBio*. American Society for Microbiology, 4(5), pp. e00611-13. doi: 10.1128/mBio.00611-13.

Ratajczak, M. Z. *et al.* (2006) 'Modulation of the SDF-1-CXCR4 axis by the third complement component (C3)--implications for trafficking of CXCR4+ stem cells', *Experimental hematology*. Elsevier BV, 34(8), pp. 986–995. doi: 10.1016/j.exphem.2006.03.015.

Ratz, M. *et al.* (2022) 'Clonal relations in the mouse brain revealed by single-cell and spatial transcriptomics', *Nature neuroscience*. Springer Science and Business Media LLC, 25(3), pp. 285–294. doi: 10.1038/s41593-022-01011-x.

Ray, M. *et al.* (2000) 'Inhibition of interferon-gamma signaling by Leishmania donovani', *The journal of infectious diseases*. Oxford University Press (OUP), 181(3), pp. 1121–1128. doi: 10.1086/315330.

Rayamajhi, M. *et al.* (2010) 'Induction of IFN-alphabeta enables Listeria monocytogenes to suppress macrophage activation by IFN-gamma', *The journal of experimental medicine*. Rockefeller University Press, 207(2), pp. 327–337. doi: 10.1084/jem.20091746.

Raup, D. M. and Crick, R. E. (1979) 'Measurement of faunal similarity in paleontology', *J Paleontol*, 53, pp. 1213–1227.

Rey, L. C. *et al.* (2005) 'American visceral leishmaniasis (kala-azar) in hospitalized children from an endemic area', *Jornal de pediatria*. Jornal de Pediatria, 81(1), pp. 73–78. doi: 10.2223/jped.1286.

Ribeiro-Gomes, F. L. and Sacks, D. (2012) 'The influence of early neutrophil-Leishmania interactions on the host immune response to infection', *Frontiers in cellular and infection microbiology*. Frontiers Media SA, 2, p. 59. doi: 10.3389/fcimb.2012.00059.

Rijal, S. *et al.* (2013) 'Increasing failure of miltefosine in the treatment of Kala-azar in Nepal and the potential role of parasite drug resistance, reinfection, or noncompliance', *Clinical infectious diseases: an official publication of the Infectious Diseases Society of America*. Oxford University Press (OUP), 56(11), pp. 1530–1538. doi: 10.1093/cid/cit102.

Rochael, N. C. *et al.* (2015) 'Classical ROS-dependent and early/rapid ROS-independent release of Neutrophil Extracellular Traps triggered by Leishmania parasites', *Scientific reports*. Springer Science and Business Media LLC, 5(1), p. 18302. doi: 10.1038/srep18302.

Rochette, A. *et al.* (2008) 'Genome-wide gene expression profiling analysis of Leishmania major and Leishmania infantum developmental stages reveals substantial differences between the two species', *BMC genomics*. Springer Science and Business Media LLC, 9(1), p. 255. doi: 10.1186/1471-2164-9-255.

Rodrigues, A. *et al.* (2017) 'Leishmania infantum antigens modulate memory cell subsets of liver resident T lymphocyte', *Immunobiology*. Elsevier BV, 222(2), pp. 409–422. doi: 10.1016/j.imbio.2016.08.009.

Rodríguez, N. E. *et al.* (2018) 'Epidemiological and experimental evidence for sex-dependent differences in the outcome of Leishmania infantum infection', *The American journal of tropical medicine and hygiene*. American Society of Tropical Medicine and Hygiene, 98(1), pp. 142–145. doi: 10.4269/ajtmh.17-0563.

Rodrigues, L. S. *et al.* (2021) 'Multifunctional, TNF-α and IFN-γ-secreting CD4 and CD8 T cells and CD8High T cells are associated with the cure of human visceral leishmaniasis', *Frontiers in immunology*. Frontiers Media SA, 12, p. 773983. doi: 10.3389/fimmu.2021.773983.

Rogers, M. B. *et al.* (2011) 'Chromosome and gene copy number variation allow major structural change between species and strains of Leishmania', *Genome research*. Cold Spring Harbor Laboratory, 21(12), pp. 2129–2142. doi: 10.1101/gr.122945.111.

Romano, A. *et al.* (2021) 'Interferon-γ-producing CD4+ T cells drive monocyte activation in the bone marrow during experimental Leishmania donovani infection', *Frontiers in immunology*. Frontiers Media SA, 12, p. 700501. doi: 10.3389/fimmu.2021.700501.

Romano, A. *et al.* (2017) 'Divergent roles for Ly6C+CCR2+CX3CR1+ inflammatory monocytes during primary or secondary infection of the skin with the intra-phagosomal pathogen Leishmania major', *PLoS pathogens*. Public Library of Science (PLoS), 13(6), p. e1006479. doi: 10.1371/journal.ppat.1006479.

Romano, A. *et al.* (2014) 'Cross-species genetic exchange between visceral and cutaneous strains of Leishmania in the sand fly vector', *Proceedings of the National Academy of Sciences of the United States of America*. Proceedings of the National Academy of Sciences, 111(47), pp. 16808–16813. doi: 10.1073/pnas.1415109111.

Ronet, C. *et al.* (2010) 'Regulatory B cells shape the development of Th2 immune responses in BALB/c mice infected with Leishmania major through IL-10 production', *The journal of immunology*. The American Association of Immunologists, 184(2), pp. 886–894. doi: 10.4049/jimmunol.0901114.

Rovira-Clavé, X. *et al.* (2022) 'Spatial epitope barcoding reveals clonal tumor patch behaviors', *Cancer cell*. Elsevier BV, 40(11), pp. 1423-1439.e11. doi: 10.1016/j.ccell.2022.09.014.

Russell, D. G. and Wright, S. D. (1988) 'Complement receptor type (CR3) binds to an Arg-Gly-Asp containing rcgion of the major surface glycoprotein, gp63, of Leishmania promastigotes', *J. Exp. Med*, 168, pp. 279–292.

Russell, D. G. and Wilhelm, H. (1986) 'The involvement of the major surface glycoprotein, gp63, of Leishmania promastigotes in attachment to macrophages', *J. Immunol*, 136, pp. 2613–2620.

Sacks, D. and Sher, A. (2002) 'Evasion of innate immunity by parasitic protozoa', *Nature immunology*. Springer Science and Business Media LLC, 3(11), pp. 1041–1047. doi: 10.1038/ni1102-1041.

Sacks, D. L. *et al.* (1995) 'Stage-specific binding of Leishmania donovani to the sand fly vector midgut is regulated by conformational changes in the abundant surface lipophosphoglycan', *The journal of experimental medicine*. Rockefeller University Press, 181(2), pp. 685–697. doi: 10.1084/jem.181.2.685.

Sacks, D. L. *et al.* (1994) 'The role of the lipophosphoglycan of*Leishmania*in vector competence', *Parasitology*. Cambridge University Press (CUP), 108(S1), pp. S55–S62. doi: 10.1017/s0031182000075727.

Salei, N. *et al.* (2017) 'Enhanced survival of Leishmania major in neutrophil granulocytes in the presence of apoptotic cells', *PloS one*. Public Library of Science (PLoS), 12(2), p. e0171850. doi: 10.1371/journal.pone.0171850.

Sampaio, M. J. A. de Q. *et al.* (2010) 'Risk factors for death in children with visceral leishmaniasis', *PLoS neglected tropical diseases*. Public Library of Science (PLoS), 4(11), p. e877. doi: 10.1371/journal.pntd.0000877.

Sänger, C. *et al.* (2015) 'Intrahepatic vascular anatomy in rats and mice--variations and surgical implications', *PloS one*. Public Library of Science (PLoS), 10(11), p. e0141798. doi: 10.1371/journal.pone.0141798.

Santos, G. de A. *et al.* (2023) 'Systematic review of treatment failure and clinical relapses in leishmaniasis from a multifactorial perspective: Clinical aspects, factors associated with the parasite and host', *Tropical medicine and infectious disease*, 8(9). doi: 10.3390/tropicalmed8090430.

Satoskar, A., Bluethmann, H. and Alexander, J. (1995) 'Disruption of the murine interleukin-4 gene inhibits disease progression during Leishmania mexicana infection but does not increase control of Leishmania donovani infection', *Infection and immunity*. American Society for Microbiology, 63(12), pp. 4894–4899. doi: 10.1128/iai.63.12.4894-4899.1995.

Schartz, N. D. *et al.* (2023) 'Mice deficient in complement C3 are protected against recognition memory deficits and astrogliosis induced by status epilepticus', *Frontiers in molecular neuroscience*. Frontiers Media SA, 16, p. 1265944. doi: 10.3389/fnmol.2023.1265944.

Schleicher, U. *et al.* (2007) 'NK cell activation in visceral leishmaniasis requires TLR9, myeloid DCs, and IL-12, but is independent of plasmacytoid DCs', *The journal of experimental medicine*. Rockefeller University Press, 204(4), pp. 893–906. doi: 10.1084/jem.20061293.

Scianimanico, S. *et al.* (1999) 'Impaired recruitment of the small GTPase rab7 correlates with the inhibition of phagosome maturation by Leishmania donovani promastigotes', *Cellular microbiology*. Hindawi Limited, 1(1), pp. 19–32. doi: 10.1046/j.1462-5822.1999.00002.x.

Scotland, R. S. *et al.* (2011) 'Sex differences in resident immune cell phenotype underlie more efficient acute inflammatory responses in female mice', *Blood*. American Society of Hematology, 118(22), pp. 5918–5927. doi: 10.1182/blood-2011-03-340281.

Scott, P. *et al.* (1988) 'Immunoregulation of cutaneous leishmaniasis. T cell lines that transfer protective immunity or exacerbation belong to different T helper subsets and respond to distinct parasite antigens', *The journal of experimental medicine*. Rockefeller University Press, 168(5), pp. 1675–1684. doi: 10.1084/jem.168.5.1675.

Scott, P. (2020) 'Long-lived skin-resident memory T cells contribute to concomitant immunity in cutaneous leishmaniasis', *Cold Spring Harbor perspectives in biology*. Cold Spring Harbor Laboratory, 12(10), p. a038059. doi: 10.1101/cshperspect.a038059.

Scott, P., Caspar, P. and Sher, A. (1990) 'Protection against Leishmania major in BALB/c mice by adoptive transfer of a T cell clone recognizing a low molecular weight antigen released by promastigotes', *The journal of immunology*. The American Association of Immunologists, 144(3), pp. 1075–1079. doi: 10.4049/jimmunol.144.3.1075.

Seco-Hidalgo, V., Osuna, A. and Pablos, L. M. D. (2015) 'To bet or not to bet: deciphering cell to cell variation in protozoan infections', *Trends in parasitology*. Elsevier BV, 31(8), pp. 350–356. doi: 10.1016/j.pt.2015.05.004.

Séguin, O. and Descoteaux, A. (2016) 'Leishmania, the phagosome, and host responses: The journey of a parasite', *Cellular immunology*. Elsevier BV, 309, pp. 1–6. doi: 10.1016/j.cellimm.2016.08.004.

Serafim, T. D. *et al.* (2007) 'The effect of malnutrition in the immune response of BALB/c mice in Leishmania chagasi infection', in *XXIII Annual Meeting of the Brazilian Society of Protozoology/XXXIV Annual Meeting on Basic Research in Chagas' Disease*. Caxambu, MG, Brazil, pp. 129–130.

Serafim, T. D. *et al.* (2018) 'Sequential blood meals promote Leishmania replication and reverse metacyclogenesis augmenting vector infectivity', *Nature microbiology*. Springer Science and Business Media LLC, 3(5), pp. 548–555. doi: 10.1038/s41564-018-0125-7.

Seybold, H. *et al.* (2020) 'A fungal pathogen induces systemic susceptibility and systemic shifts in wheat metabolome and microbiome composition', *Nature communications*. Springer Science and Business Media LLC, 11(1), p. 1910. doi: 10.1038/s41467-020-15633-x.

Sghaier, R. M. *et al.* (2022) 'Healed lesions of human cutaneous leishmaniasis caused by Leishmania major do not shelter persistent residual parasites', *Frontiers in cellular and infection microbiology*. Frontiers Media SA, 12, p. 839216. doi: 10.3389/fcimb.2022.839216.

Shannon, C. (2021) 'A mathematical theory of communication (1948)', in *Ideas That Created the Future*. The MIT Press, pp. 121–134. doi: 10.7551/mitpress/12274.003.0014.

Shi, Q. *et al.* (2015) 'Complement C3-deficient mice fail to display age-related hippocampal decline', *The Journal of neuroscience: the official journal of the Society for Neuroscience*. Society for Neuroscience, 35(38), pp. 13029–13042. doi: 10.1523/JNEUROSCI.1698-15.2015.

Shio, M. T. *et al.* (2015) 'PKC/ROS-mediated NLRP3 inflammasome activation is attenuated by Leishmania zinc-metalloprotease during infection', *PLoS neglected tropical diseases*. Public Library of Science (PLoS), 9(6), p. e0003868. doi: 10.1371/journal.pntd.0003868.

Sidik, S. M. *et al.* (2016) 'A genome-wide CRISPR screen in Toxoplasma identifies essential Apicomplexan genes', *Cell*. Elsevier BV, 166(6), pp. 1423-1435.e12. doi: 10.1016/j.cell.2016.08.019.

Sikorski, P. M., Commodaro, A. G. and Grigg, M. E. (2019) 'Toxoplasma gondii recruits Factor H and C4b-binding protein to mediate resistance to serum killing and promote parasite persistence in vivo', *Frontiers in immunology*. Frontiers Media SA, 10, p. 3105. doi: 10.3389/fimmu.2019.03105.

Silva-Barrios, S., Charpentier, T. and Stäger, S. (2018) 'The deadly dance of B cells with trypanosomatids', *Trends in parasitology*. Elsevier BV, 34(2), pp. 155–171. doi: 10.1016/j.pt.2017.10.001.

Silva-Barrios, S. *et al.* (2016) 'Innate immune B cell activation by Leishmania donovani exacerbates disease and mediates hypergammaglobulinemia', *Cell reports*. Elsevier BV, 15(11), pp. 2427–2437. doi: 10.1016/j.celrep.2016.05.028.

Silva-Barrios, S. and Stäger, S. (2019) 'Hypergammaglobulinemia sustains the development of regulatory responses during chronic Leishmania donovani infection in mice', *European journal of immunology*. Wiley, 49(7), pp. 1082–1091. doi: 10.1002/eji.201847917.

Singh, N., Kumar, M. and Singh, R. K. (2012) 'Leishmaniasis: Current status of available drugs and new potential drug targets', *Asian Pacific journal of tropical medicine*. Medknow, 5(6), pp. 485–497. doi: 10.1016/s1995-7645(12)60084-4.

Sloan, W. T. *et al.* (2006) 'Quantifying the roles of immigration and chance in shaping prokaryote community structure', *Environmental microbiology*. Wiley, 8(4), pp. 732–740. doi: 10.1111/j.1462-2920.2005.00956.x.

Small, A. G. *et al.* (2022) 'Neutrophils require activation to express functional cell-surface Complement Receptor Immunoglobulin', *Frontiers in immunology*. Frontiers Media SA, 13, p. 840510. doi: 10.3389/fimmu.2022.840510.

Smelt, S. C. *et al.* (2000) 'B cell-deficient mice are highly resistant to Leishmania donovani infection, but develop neutrophil-mediated tissue pathology', *The journal of immunology*. The American Association of Immunologists, 164(7), pp. 3681–3688. doi: 10.4049/jimmunol.164.7.3681.

Smith, A. M. *et al.* (2009) 'Quantitative phenotyping via deep barcode sequencing', *Genome research*. Cold Spring Harbor Laboratory, 19(10), pp. 1836–1842. doi: 10.1101/gr.093955.109.

Smith, A. M. *et al.* (2010) 'Highly-multiplexed barcode sequencing: an efficient method for parallel analysis of pooled samples', *Nucleic acids research*. Oxford University Press (OUP), 38(13), p. e142. doi: 10.1093/nar/gkq368.

Soong, L. (2008) 'Modulation of dendritic cell function by Leishmania parasites', *The journal of immunology*. The American Association of Immunologists, 180(7), pp. 4355–4360. doi: 10.4049/jimmunol.180.7.4355.

Späth, G. F. *et al.* (2003) 'The role(s) of lipophosphoglycan (LPG) in the establishment of Leishmania major infections in mammalian hosts', *Proceedings of the National Academy of Sciences of the United States of America*. Proceedings of

the National Academy of Sciences, 100(16), pp. 9536–9541. doi: 10.1073/pnas.1530604100.

Spotin, A. and Parvizi, P. (2016) 'Comparative study of viscerotropic pathogenicity of Leishmania major amastigotes and promastigotes based on identification of mitochondrial and nucleus sequences', *Parasitology research*. Springer Science and Business Media LLC, 115(3), pp. 1221–1228. doi: 10.1007/s00436-015-4858-4.

Stäger, S. *et al.* (2003) 'Both interleukin-4 (IL-4) and IL-4 receptor α signaling contribute to the development of hepatic granulomas with optimal antileishmanial activity', *Infection and immunity*. American Society for Microbiology, 71(8), pp. 4804–4807. doi: 10.1128/iai.71.8.4804-4807.2003.

Stäger, S. and Rafati, S. (2012) 'CD8(+) T cells in leishmania infections: friends or foes?', *Frontiers in immunology*. Frontiers Media SA, 3, p. 5. doi: 10.3389/fimmu.2012.00005.

Stanley, A. C. and Engwerda, C. R. (2007) 'Balancing immunity and pathology in visceral leishmaniasis', *Immunology and cell biology*. Wiley, 85(2), pp. 138–147. doi: 10.1038/sj.icb7100011.

Stegen, J. C. *et al.* (2012) 'Stochastic and deterministic assembly processes in subsurface microbial communities', *The ISME journal*. Oxford University Press (OUP), 6(9), pp. 1653–1664. doi: 10.1038/ismej.2012.22.

Stegen, J. C. *et al.* (2013) 'Quantifying community assembly processes and identifying features that impose them', *The ISME journal*. Oxford University Press (OUP), 7(11), pp. 2069–2079. doi: 10.1038/ismej.2013.93.

Stegen, J. C. *et al.* (2015) 'Estimating and mapping ecological processes influencing microbial community assembly', *Frontiers in microbiology*. Frontiers Media SA, 6, p. 370. doi: 10.3389/fmicb.2015.00370.

Stögerer, T. *et al.* (2023) 'Leishmania donovani exploits tunneling nanotubes for dissemination and propagation of B cell activation', *Microbiology spectrum*, 11(4), p. e0509622. doi: 10.1128/spectrum.05096-22.

Suzuki, M. M., Satoh, N. and Nonaka, M. (2002) 'C6-like and C3-like molecules from the cephalochordate, amphioxus, suggest a cytolytic complement system in invertebrates', *Journal of molecular evolution*. Springer Science and Business Media LLC, 54(5), pp. 671–679. doi: 10.1007/s00239-001-0068-z.

Tabatabaee, P.-A. *et al.* (2011) 'Leishmania major: secreted antigens of Leishmania major promastigotes shift the immune response of the C57BL/6 mice toward Th2 in vitro', *Experimental parasitology*. Elsevier BV, 127(1), pp. 46–51. doi: 10.1016/j.exppara.2010.06.033.

Tabatabaei, M. S., Islam, R. and Ahmed, M. (2021) 'Applications of gold nanoparticles in ELISA, PCR, and immuno-PCR assays: A review', *Analytica chimica acta*. Elsevier BV, 1143, pp. 250–266. doi: 10.1016/j.aca.2020.08.030.

Takele, Y. *et al.* (2022) 'Immunological factors, but not clinical features, predict visceral leishmaniasis relapse in patients co-infected with HIV', *Cell reports. Medicine*. Elsevier BV, 3(1), p. 100487. doi: 10.1016/j.xcrm.2021.100487.

Tan, J. *et al.* (2012) 'Species phylogenetic relatedness, priority effects, and ecosystem functioning', *Ecology*. Wiley, 93(5), pp. 1164–1172. doi: 10.1890/11-1557.1.

Tanaka, S., Suzuki, T. and Nishioka, K. (1986) 'Assay of classical and alternative pathway activities of murine complement using antibody-sensitized rabbit erythrocytes', *Journal of immunological methods*. Elsevier BV, 86(2), pp. 161–170. doi: 10.1016/0022-1759(86)90448-5.

Terrazas, C. *et al.* (2017) 'Ly6Chi inflammatory monocytes promote susceptibility to Leishmania donovani infection', *Scientific reports*. Springer Science and Business Media LLC, 7(1). doi: 10.1038/s41598-017-14935-3.

The PLOS ONE Staff (2018) 'Correction: B cell: T cell interactions occur within hepatic granulomas during experimental visceral leishmaniasis', *PloS one*. Public Library of Science (PLoS), 13(6), p. e0199184. doi: 10.1371/journal.pone.0199184.

Tillack, K. *et al.* (2012) 'T lymphocyte priming by neutrophil extracellular traps links innate and adaptive immune responses', *The journal of immunology*. The American Association of Immunologists, 188(7), pp. 3150–3159. doi: 10.4049/jimmunol.1103414.

Tirado, T. C. *et al.* (2021) 'A comparative approach on the activation of the three complement system pathways in different hosts of Visceral Leishmaniasis after stimulation with Leishmania infantum', *Developmental and comparative immunology*. Elsevier BV, 120(104061), p. 104061. doi: 10.1016/j.dci.2021.104061.

Titus, R. G. *et al.* (1995) 'Development of a safe live Leishmania vaccine line by gene replacement', *Proceedings of the National Academy of Sciences of the United States of America*. Proceedings of the National Academy of Sciences, 92(22), pp. 10267–10271. doi: 10.1073/pnas.92.22.10267.

Tripathi, A. *et al.* (2021) 'Type 1 interferon mediates chronic stress-induced neuroinflammation and behavioral deficits via complement component 3-dependent pathway', *Molecular psychiatry*. Springer Science and Business Media LLC, 26(7), pp. 3043–3059. doi: 10.1038/s41380-021-01065-6.

Tucker, C. M. and Fukami, T. (2014) 'Environmental variability counteracts priority effects to facilitate species coexistence: evidence from nectar microbes', *Proceedings. Biological sciences*. The Royal Society, 281(1778), p. 20132637. doi: 10.1098/rspb.2013.2637.

Ubeda, J.-M. *et al.* (2008) 'Modulation of gene expression in drug resistant Leishmania is associated with gene amplification, gene deletion and chromosome aneuploidy', *Genome biology*. Springer Nature, 9(7), p. R115. doi: 10.1186/gb-2008-9-7-r115.

Ubeda, J.-M. *et al.* (2014) 'Genome-wide stochastic adaptive DNA amplification at direct and inverted DNA repeats in the parasite Leishmania', *PLoS biology*. Public Library of Science (PLoS), 12(5), p. e1001868. doi: 10.1371/journal.pbio.1001868.

Ueno, N. and Wilson, M. E. (2012) 'Receptor-mediated phagocytosis of Leishmania: implications for intracellular survival', *Trends in parasitology*. Elsevier BV, 28(8), pp. 335–344. doi: 10.1016/j.pt.2012.05.002.

Valero, N. N. H. and Uriarte, M. (2020) 'Environmental and socioeconomic risk factors associated with visceral and cutaneous leishmaniasis: a systematic review', *Parasitology research*. Springer Science and Business Media LLC, 119(2), pp. 365–384. doi: 10.1007/s00436-019-06575-5.

van der Torre, M.H. *et al.* (2022) 'Systematic review on Cryptococcus neoformans/Cryptococcus gattii species complex infections with recommendations for practice in health and care settings', *Clinical infection in practice*, 15(100154), p. 100154. Available at: https://doi.org/10.1016/j.clinpr.2022.100154.

Varadharajan, B. and Parani, M. (2021) 'DMSO and betaine significantly enhance the PCR amplification of ITS2 DNA barcodes from plants', *Genome*. Canadian Science Publishing, 64(3), pp. 165–171. doi: 10.1139/gen-2019-0221.

Vargas-Inchaustegui, D. A., Xin, L. and Soong, L. (2008) 'Leishmania braziliensis infection induces dendritic cell activation, ISG15 transcription, and the generation of protective immune responses', *The journal of immunology*. The American Association of Immunologists, 180(11), pp. 7537–7545. doi: 10.4049/jimmunol.180.11.7537.

Vasquez, K. S. *et al.* (2021) 'Quantifying rapid bacterial evolution and transmission within the mouse intestine', *Cell host & microbe*. Elsevier BV, 29(9), pp. 1454-1468.e4. doi: 10.1016/j.chom.2021.08.003.

Vellend, M. (2010) 'Conceptual synthesis in community ecology', *The quarterly review of biology*. University of Chicago Press, 85(2), pp. 183–206. doi: 10.1086/652373.

Verma, J. K., Rastogi, R. and Mukhopadhyay, A. (2017) 'Leishmania donovani resides in modified early endosomes by upregulating Rab5a expression via the downregulation of miR-494', *PLoS pathogens*. Public Library of Science (PLoS), 13(6), p. e1006459. doi: 10.1371/journal.ppat.1006459.

Verma, S. *et al.* (2018) 'Leishmania donovani inhibitor of Serine peptidases 2 mediated inhibition of lectin pathway and upregulation of C5aR signaling promote parasite survival inside host', *Frontiers in immunology*. Frontiers Media SA, 9. doi: 10.3389/fimmu.2018.00063.

Vogel, C.-W. *et al.* (2004) 'Recombinant cobra venom factor', *Molecular immunology*. Elsevier BV, 41(2–3), pp. 191–199. doi: 10.1016/j.molimm.2004.03.011.

Volf, P. and Volfova, V. (2011) 'Establishment and maintenance of sand fly colonies', *Journal of vector ecology: journal of the Society for Vector Ecology*. Wiley, 36 Suppl 1, pp. S1-9. doi: 10.1111/j.1948-7134.2011.00106.x.

Volpedo, G. *et al.* (2022) 'The history of live attenuated centrin gene-deleted Leishmania vaccine candidates', *Pathogens*. MDPI AG, 11(4), p. 431. doi: 10.3390/pathogens11040431.

Vu, B. V., Litvinov, D. and Willson, R. C. (2008) 'Gold nanoparticle effects in polymerase chain reaction: favoring of smaller products by polymerase

adsorption', *Analytical chemistry*. American Chemical Society (ACS), 80(14), pp. 5462–5467. doi: 10.1021/ac8000258.

Wang, Y. *et al.* (2011) 'Differential microbicidal effects of human histone proteins H2A and H2B on Leishmania promastigotes and amastigotes', *Infection and immunity*. American Society for Microbiology, 79(3), pp. 1124–1133. doi: 10.1128/IAI.00658-10.

Ward, A. I. *et al.* (2020) 'In vivo analysis of Trypanosoma cruzi persistence foci at single-cell resolution', *mBio*. American Society for Microbiology, 11(4). doi: 10.1128/mBio.01242-20.

Weaver, S. C. *et al.* (2021) 'Population bottlenecks and founder effects: implications for mosquito-borne arboviral emergence', *Nature reviews. Microbiology*. Springer Science and Business Media LLC, 19(3), pp. 184–195. doi: 10.1038/s41579-020-00482-8.

Weber, E. M. *et al.* (2022) 'Aggression in group-housed male mice: A systematic review', *Animals: an open access journal from MDPI*. MDPI AG, 13(1), p. 143. doi: 10.3390/ani13010143.

Weinheber, N. *et al.* (1998) 'Phagocytosis ofLeishmania mexicana amastigotes by macrophages leads to a sustained suppression of IL-12 production', *European journal of immunology*. Wiley, 28(8), pp. 2467–2477. doi: 10.1002/(sici)1521-4141(199808)28:08<2467::aid-immu2467>3.0.co;2-1.

Wenning, B. *et al.* (2022) 'Exploring the cultural effects of gender on perceptions of cutaneous leishmaniasis: a systematic literature review', *Global health research and policy*, 7(1), p. 34. doi: 10.1186/s41256-022-00266-y.

Wessels, M. R. *et al.* (1995) 'Studies of group B streptococcal infection in mice deficient in complement component C3 or C4 demonstrate an essential role for complement in both innate and acquired immunity', *Proceedings of the National*

Academy of Sciences of the United States of America. Proceedings of the National Academy of Sciences, 92(25), pp. 11490–11494. doi: 10.1073/pnas.92.25.11490.

Westacott, L. J. *et al.* (2021) 'Dissociable effects of complement C3 and C3aR on survival and morphology of adult born hippocampal neurons, pattern separation, and cognitive flexibility in male mice', *Brain, behavior, and immunity*. Elsevier BV, 98, pp. 136–150. doi: 10.1016/j.bbi.2021.08.215.

West, E. E. and Kemper, C. (2023) 'Complosome - the intracellular complement system', *Nature reviews. Nephrology*. Springer Science and Business Media LLC, 19(7), pp. 426–439. doi: 10.1038/s41581-023-00704-1.

WHO. *Leishmaniasis* (no date). Available at: https://www.who.int/news-room/fact-sheets/detail/leishmaniasis (Accessed: 29 September 2024).

Wilhelm, P. *et al.* (2001) 'Rapidly fatal leishmaniasis in resistant C57BL/6 mice lacking TNF', *The journal of immunology*. The American Association of Immunologists, 166(6), pp. 4012–4019. doi: 10.4049/jimmunol.166.6.4012.

Wijnant, G.-J. *et al.* (2022) 'Tackling drug resistance and other causes of treatment failure in leishmaniasis', *Frontiers in tropical diseases*. Frontiers Media SA, 3. doi: 10.3389/fitd.2022.837460.

Wincott, C. J. *et al.* (2022) 'Cellular barcoding of protozoan pathogens reveals the within-host population dynamics of Toxoplasma gondii host colonization', *Cell reports methods*. Elsevier BV, 2(8), p. 100274. doi: 10.1016/j.crmeth.2022.100274.

Windels, E. M. *et al.* (2021) 'Population bottlenecks strongly affect the evolutionary dynamics of antibiotic persistence', *Molecular biology and evolution*. Oxford University Press (OUP), 38(8), pp. 3345–3357. doi: 10.1093/molbev/msab107.

Wright, O. *et al.* (2023) 'C5aR2 regulates STING-mediated interferon beta production in human macrophages', *Cells (Basel, Switzerland)*, 12(23). doi: 10.3390/cells12232707.

Wright, S. (1950) 'Genetical structure of populations', *Nature*. Springer Science and Business Media LLC, 166(4215), pp. 247–249. doi: 10.1038/166247a0.

Wu, Y. *et al.* (2000) 'A new developmentally regulated gene family in Leishmania amastigotes encoding a homolog of amastin surface proteins', *Molecular and biochemical parasitology*. Elsevier BV, 110(2), pp. 345–357. doi: 10.1016/s0166-6851(00)00290-5.

Wu, M. *et al.* (2024) 'Gut complement induced by the microbiota combats pathogens and spares commensals', *Cell*. Elsevier BV, 187(4), pp. 897-913.e18. doi: 10.1016/j.cell.2023.12.036.

Xiao, Y. *et al.* (2017) 'Mapping the ecological networks of microbial communities', *Nature communications*. Springer Science and Business Media LLC, 8(1), p. 2042. doi: 10.1038/s41467-017-02090-2.

Xin, L. *et al.* (2010) 'Type I IFN receptor regulates neutrophil functions and innate immunity to Leishmania parasites', *The journal of immunology*. The American Association of Immunologists, 184(12), pp. 7047–7056. doi: 10.4049/jimmunol.0903273.

Xie, X. *et al.* (2020) 'Single-cell transcriptome profiling reveals neutrophil heterogeneity in homeostasis and infection', *Nature immunology*. Springer Science and Business Media LLC, 21(9), pp. 1119–1133. doi: 10.1038/s41590-020-0736-z.

Xu, L. *et al.* (2021) 'River underwater 3D vision measurement method', *Guangxue Jingmi Gongcheng (Optics and Precision Engineering)*. Changchun Institute of Optics, Fine Mechanics and Physics, Chinese Academy of Sciences, 29(10), pp. 2465–2480. doi: 10.37188/ope.20212910.2465.

Xun, W. *et al.* (2019) ‘Diversity-triggered deterministic bacterial assembly constrains community functions’, *Nature communications*. Springer Science and Business Media LLC, 10(1), p. 3833. doi: 10.1038/s41467-019-11787-5.

Yadav, M. *et al.* (2023) ‘Leishmania donovani attenuates dendritic cell trafficking to lymph nodes by inhibiting C-type lectin receptor 2 expression via transforming growth factor-β’, *Microbiology spectrum*. American Society for Microbiology, 11(3), p. e0412222. doi: 10.1128/spectrum.04122-22.

Ye, Y. *et al.* (2020) ‘Plasmacytoid dendritic cell biology and its role in immune-mediated diseases’, *Clinical & translational immunology*. Wiley, 9(5), p. e1139. doi: 10.1002/cti2.1139.

Younis, B. M. *et al.* (2021) ‘Safety and immunogenicity of ChAd63-KH vaccine in post-kala-azar dermal leishmaniasis patients in Sudan’, *Molecular therapy: the journal of the American Society of Gene Therapy*. Elsevier BV, 29(7), pp. 2366–2377. doi: 10.1016/j.ymthe.2021.03.020.

Younis, B. M. *et al.* (2024) ‘A randomized, double-blind phase 2b trial to evaluate efficacy of ChAd63-KH for treatment of post kala-azar dermal leishmaniasis’, *Molecular therapy. Methods & clinical development*. Elsevier BV, 32(3), p. 101310. doi: 10.1016/j.omtm.2024.101310.

Yuce, M. *et al.* (2014) ‘Employment of nanomaterials in polymerase chain reaction: insight into the impacts and putative operating mechanisms of nano-additives in PCR’, *RSC advances*. Royal Society of Chemistry (RSC), 4(69), pp. 36800–36814. doi: 10.1039/c4ra06144f.

Yuste, E. *et al.* (1999) ‘Drastic fitness loss in human immunodeficiency virus type 1 upon serial bottleneck events’, *Journal of virology*. American Society for Microbiology, 73(4), pp. 2745–2751. doi: 10.1128/JVI.73.4.2745-2751.1999.

Zamboni, D. S. and Sacks, D. L. (2019) 'Inflammasomes and Leishmania: in good times or bad, in sickness or in health', *Current opinion in microbiology*. Elsevier BV, 52, pp. 70–76. doi: 10.1016/j.mib.2019.05.005.

Zhang, T. *et al.* (2017) 'Deciphering the landscape of host barriers to*Listeria monocytogenes*infection', *Proceedings of the National Academy of Sciences of the United States of America*. Proceedings of the National Academy of Sciences, 114(24), pp. 6334–6339. doi: 10.1073/pnas.1702077114.

Zhang, WW. and Matlashewski, G. (2015) 'CRISPR-Cas9-mediated genome editing in Leishmania donovani', 6, pp. e00861-876.

Zhang, W. W. and Matlashewski, G. (1997) 'Loss of virulence in Leishmania donovani deficient in an amastigote-specific protein, A2', *Proceedings of the National Academy of Sciences of the United States of America*. Proceedings of the National Academy of Sciences, 94(16), pp. 8807–8811. doi: 10.1073/pnas.94.16.8807.

Zhang, W. W. *et al.* (2014) 'Genetic analysis of Leishmania donovani tropism using a naturally attenuated cutaneous strain', *PLoS pathogens*. Public Library of Science (PLoS), 10(7), p. e1004244. doi: 10.1371/journal.ppat.1004244.

Zhang, W.-W., Lypaczewski, P. and Matlashewski, G. (2017) 'Optimized CRISPR-Cas9 genome editing for Leishmania and its use to target a multigene family, induce chromosomal translocation, and study DNA break repair mechanisms', *mSphere*. American Society for Microbiology, 2(1). doi: 10.1128/mSphere.00340-16.

Zhang, W.-W., Lypaczewski, P. and Matlashewski, G. (2020) 'Application of CRISPR/Cas9-mediated genome editing in Leishmania', in *Methods in Molecular Biology*. New York, NY: Springer US (Methods in molecular biology (Clifton, N.J.)), pp. 199–224. doi: 10.1007/978-1-0716-0294-2_14.

Zhao, Z. *et al.* (2022) ‘Mosquito brains encode unique features of human odour to drive host seeking’, *Nature*. Springer Science and Business Media LLC, 605(7911), pp. 706–712. doi: 10.1038/s41586-022-04675-4.

Zhou, J. and Ning, D. (2017) ‘Stochastic community assembly: Does it matter in microbial ecology?’, *Microbiology and molecular biology reviews: MMBR*. American Society for Microbiology, 81(4). doi: 10.1128/MMBR.00002-17.

Zhu, Y. *et al.* (2005) ‘The ancient origin of the complement system’, *The EMBO journal*. Wiley, 24(2), pp. 382–394. doi: 10.1038/sj.emboj.7600533.

Zijlstra, E. E. (2014) ‘PKDL and other dermal lesions in HIV co-infected patients with Leishmaniasis: review of clinical presentation in relation to immune responses’, *PLoS neglected tropical diseases*. Public Library of Science (PLoS), 8(11), p. e3258. doi: 10.1371/journal.pntd.0003258.

Zijlstra, E. E. (2016) ‘The immunology of post-kala-azar dermal leishmaniasis (PKDL)’, *Parasites & vectors*. Springer Science and Business Media LLC, 9(1). doi: 10.1186/s13071-016-1721-0.

Appendices:

Appendix Table 1: Barcodes and primers used for library generation.

Barcode ID	Barcode	Primer ID	Primer Sequence
B1	cgtggattggaacagcttct	OL12845	5’ CTCGGACGAGTGCTGGGGCGTCG GTTTCCAAATTAACCCTCACTAAAG GGAGGAGAAGCTGTTCCAATCCAC GcttccagttttatagtgagtcgtattaCCA ATTTGAGAGACCTGTGC 3’
B2	ttctaccgagtcctcgtaga	OL12881	5’ CTCGGACGAGTGCTGGGGCGTCG GTTTCCAAATTAACCCTCACTAAAG GGAGGtctacgaggactcggtagaacttc cagttttatagtgagtcgtattaCCAATTTG AGAGACCTGTGC 3’
B3	gggcttggacttagtgaaag	OL12888	5’ CTCGGACGAGTGCTGGGGCGTCG GTTTCCAAATTAACCCTCACTAAAG GGAGGctttcactaagtccaagccccttcc agttttatagtgagtcgtattaCCAATTTG AGAGACCTGTGC 3’
B4	aacgctgaccacaagttcga	OL12889	5’ CTCGGACGAGTGCTGGGGCGTCG GTTTCCAAATTAACCCTCACTAAAG GGAGGtcgaacttgtggtcagcgttcttcc agttttatagtgagtcgtattaCCAATTTG AGAGACCTGTGC 3’
B5	gacccaacaaagtgttcacc	OL12892	5’ CTCGGACGAGTGCTGGGGCGTCG GTTTCCAAATTAACCCTCACTAAAG GGAGGgtgaacactttgttgggtccttcc agttttatagtgagtcgtattaCCAATTTG AGAGACCTGTGC 3’
B6	gtggtacactatacggttgc	OL12893	5’ CTCGGACGAGTGCTGGGGCGTCG GTTTCCAAATTAACCCTCACTAAAG GGAGGgcaaccgtatagtgtaccaccttcc agttttatagtgagtcgtattaCCAATTTG AGAGACCTGTGC 3’
B7	aacgacacatggacgtacca	OL12894	5’ CTCGGACGAGTGCTGGGGCGTCG GTTTCCAAATTAACCCTCACTAAAG GGAGGtggtacgtccatgtgtcgttcttcca gttttatagtgagtcgtattaCCAATTTGA GAGACCTGTGC 3’
B8	ttagcgtctccctacacctt	OL12895	5’ CTCGGACGAGTGCTGGGGCGTCG GTTTCCAAATTAACCCTCACTAAAG GGAGGaaggtgtagggagacgctaacttc cagttttatagtgagtcgtattaCCAATTTG AGAGACCTGTGC 3’
B9	tagcggcttccctaaccata	OL12896	5’ CTCGGACGAGTGCTGGGGCGTCG GTTTCCAAATTAACCCTCACTAAAG GGAGGtatggttagggaagccgctacttcc

			agttttatagtgagtcgtattaCCAATTTG AGAGACCTGTGC 3’
B10	tcgtgtcaacagtgcgacta	OL12897	5’ CTCGGACGAGTGCTGGGGCGTCG GTTTCCAAATTAACCCTCACTAAAG GGAGGtagtcgcactgttgacacgacttcc agttttatagtgagtcgtattaCCAATTTG AGAGACCTGTGC 3’
B11	CCGTCACGACCGGCTG AAAT	OL13279	5’ CTCGGACGAGTGCTGGGGCGTCG GTTTCCAAATTAACCCTCACTAAAG GGAGGATTTCAGCCGGTCGTGACG GcttccagttttatagtgagtcgtattaCCA ATTTGAGAGACCTGTGC 3’
B12	AGGTTTATATCATCAGC GCT	OL13280	5’ CTCGGACGAGTGCTGGGGCGTCG GTTTCCAAATTAACCCTCACTAAAG GGAGGAGCGCTGATGATATAAACC TcttccagttttatagtgagtcgtattaCCAA TTTGAGAGACCTGTGC 3’
B13	GTGCTGACCAATCGACC AAG	OL13281	5’ CTCGGACGAGTGCTGGGGCGTCG GTTTCCAAATTAACCCTCACTAAAG GGAGGCTTGGTCGATTGGTCAGCA CcttccagttttatagtgagtcgtattaCCAA TTTGAGAGACCTGTGC 3’
B14	CTCATCCTGCGAACCAA GTG	OL13282	5’ CTCGGACGAGTGCTGGGGCGTCG GTTTCCAAATTAACCCTCACTAAAG GGAGGCACTTGGTTCGCAGGATG AGcttccagttttatagtgagtcgtattaCC AATTTGAGAGACCTGTGC 3’
B15	TGAAGGTAGACGGAGG ATTT	OL13283	5’ CTCGGACGAGTGCTGGGGCGTCG GTTTCCAAATTAACCCTCACTAAAG GGAGGAAATCCTCCGTCTACCTTC ActtccagttttatagtgagtcgtattaCCAA TTTGAGAGACCTGTGC 3’
B16	ACCACCCGGACATGATC GTC	OL13284	5’ CTCGGACGAGTGCTGGGGCGTCG GTTTCCAAATTAACCCTCACTAAAG GGAGGGACGATCATGTCCGGGTG GTcttccagttttatagtgagtcgtattaCCA ATTTGAGAGACCTGTGC 3’
B17	CGTGATAGACTAGCGC AAGG	OL13285	5’ CTCGGACGAGTGCTGGGGCGTCG GTTTCCAAATTAACCCTCACTAAAG GGAGGCCTTGCGCTAGTCTATCAC GcttccagttttatagtgagtcgtattaCCA ATTTGAGAGACCTGTGC 3’
B18	ATTCCCGACGTTAACAG TAC	OL13286	5’ CTCGGACGAGTGCTGGGGCGTCG GTTTCCAAATTAACCCTCACTAAAG GGAGGGTACTGTTAACGTCGGGA ATcttccagttttatagtgagtcgtattaCCA ATTTGAGAGACCTGTGC 3’
B19	TGGCACCGGACCGTCT GAAA	OL13287	5’ CTCGGACGAGTGCTGGGGCGTCG GTTTCCAAATTAACCCTCACTAAAG

			GGAGGTTTCAGACGGTCCGGTGCC ActtccagttttatagtgagtcgtattaCCAA TTTGAGAGACCTGTGC 3’
B20	CTAGACGCTGTAGTTCT GTA	OL13288	5’ CTCGGACGAGTGCTGGGGCGTCG GTTTCCAAATTAACCCTCACTAAAG GGAGGTACAGAACTACAGCGTCTA GcttccagttttatagtgagtcgtattaCCA ATTTGAGAGACCTGTGC 3’
B21	CTGTCCATGCTTGCCTG TGA	OL13289	5’ CTCGGACGAGTGCTGGGGCGTCG GTTTCCAAATTAACCCTCACTAAAG GGAGGTCACAGGCAAGCATGGAC AGcttccagttttatagtgagtcgtattaCC AATTTGAGAGACCTGTGC 3’
B22	CTAGTAAATGGAGCGG GCCA	OL13290	5’ CTCGGACGAGTGCTGGGGCGTCG GTTTCCAAATTAACCCTCACTAAAG GGAGGTGGCCCGCTCCATTTACTA GcttccagttttatagtgagtcgtattaCCA ATTTGAGAGACCTGTGC 3’
B23	CATAACTTGGATAGAA GTGC	OL13291	5’ CTCGGACGAGTGCTGGGGCGTCG GTTTCCAAATTAACCCTCACTAAAG GGAGGGCACTTCTATCCAAGTTAT GcttccagttttatagtgagtcgtattaCCA ATTTGAGAGACCTGTGC 3’
B24	GGTAGTTCGACGTGCC CTGA	OL13292	5’ CTCGGACGAGTGCTGGGGCGTCG GTTTCCAAATTAACCCTCACTAAAG GGAGGTCAGGGCACGTCGAACTA CCcttccagttttatagtgagtcgtattaCCA ATTTGAGAGACCTGTGC 3’
B25	TAGTATTGGCGGGACT CATC	OL13293	5’ CTCGGACGAGTGCTGGGGCGTCG GTTTCCAAATTAACCCTCACTAAAG GGAGGGATGAGTCCCGCCAATACT ActtccagttttatagtgagtcgtattaCCAA TTTGAGAGACCTGTGC 3’
B26	TAATTGTACCTCTTCCC GGA	OL13294	5’ CTCGGACGAGTGCTGGGGCGTCG GTTTCCAAATTAACCCTCACTAAAG GGAGGTCCGGGAAGAGGTACAAT TActtccagttttatagtgagtcgtattaCCA ATTTGAGAGACCTGTGC 3’
B27	CACTGCATTTGCTTCCC TGC	OL13295	5’ CTCGGACGAGTGCTGGGGCGTCG GTTTCCAAATTAACCCTCACTAAAG GGAGGGCAGGGAAGCAAATGCAG TGcttccagttttatagtgagtcgtattaCCA ATTTGAGAGACCTGTGC 3’
B28	CAAGACCTTAGTTTGAT CTG	OL13296	5’ CTCGGACGAGTGCTGGGGCGTCG GTTTCCAAATTAACCCTCACTAAAG GGAGGCAGATCAAACTAAGGTCTT GcttccagttttatagtgagtcgtattaCCA ATTTGAGAGACCTGTGC 3’
B29	GCTGCAGGTTAAGATG TTCA	OL13297	5’ CTCGGACGAGTGCTGGGGCGTCG

			GTTTCCAAATTAACCCTCACTAAAG GGAGGTGAACATCTTAACCTGCAG CcttccagttttatagtgagtcgtattaCCAA TTTGAGAGACCTGTGC 3’
B30	CTAACACGGTACTACTC ATC	OL13298	5’ CTCGGACGAGTGCTGGGGCGTCG GTTTCCAAATTAACCCTCACTAAAG GGAGGGATGAGTAGTACCGTGTT AGcttccagttttatagtgagtcgtattaCC AATTTGAGAGACCTGTGC 3’
B31	CCTCCCGTGAAAGTACT AAC	OL13411	5’ CTCGGACGAGTGCTGGGGCGTCG GTTTCCAAATTAACCCTCACTAAAG GGAGGGTTAGTACTTTCACGGGAG GcttccagttttatagtgagtcgtattaCCA ATTTGAGAGACCTGTGC 3’
B32	GTTAGTACTTTCACGGG AGG	OL13412	5’ CTCGGACGAGTGCTGGGGCGTCG GTTTCCAAATTAACCCTCACTAAAG GGAGGCGTGTAGGCTCTAAACTAG ActtccagttttatagtgagtcgtattaCCAA TTTGAGAGACCTGTGC 3’
B33	TCTGGTGGTCTCCTATC TAG	OL13413	5’ CTCGGACGAGTGCTGGGGCGTCG GTTTCCAAATTAACCCTCACTAAAG GGAGGCTAGATAGGAGACCACCA GActtccagttttatagtgagtcgtattaCC AATTTGAGAGACCTGTGC 3’
B34	CTTACGCCAGCGTTTGC AAT	OL13414	5’ CTCGGACGAGTGCTGGGGCGTCG GTTTCCAAATTAACCCTCACTAAAG GGAGGATTGCAAACGCTGGCGTA AGcttccagttttatagtgagtcgtattaCC AATTTGAGAGACCTGTGC 3’
B35	CGAACGGCGCAGTTAG TCAA	OL13415	5’ CTCGGACGAGTGCTGGGGCGTCG GTTTCCAAATTAACCCTCACTAAAG GGAGGTTGACTAACTGCGCCGTTC GcttccagttttatagtgagtcgtattaCCA ATTTGAGAGACCTGTGC 3’
B36	GATTTAAACACCCAACG CTA	OL13416	5’ CTCGGACGAGTGCTGGGGCGTCG GTTTCCAAATTAACCCTCACTAAAG GGAGGTAGCGTTGGGTGTTTAAAT CcttccagttttatagtgagtcgtattaCCAA TTTGAGAGACCTGTGC 3’
B37	GGTACTCTTGAGTTGCA GAC	OL13417	5’ CTCGGACGAGTGCTGGGGCGTCG GTTTCCAAATTAACCCTCACTAAAG GGAGGGTCTGCAACTCAAGAGTAC CcttccagttttatagtgagtcgtattaCCAA TTTGAGAGACCTGTGC 3’
B38	ATGGGTAAGTAGCCTG AAGT	OL13418	5’ CTCGGACGAGTGCTGGGGCGTCG GTTTCCAAATTAACCCTCACTAAAG GGAGGACTTCAGGCTACTTACCCA TcttccagttttatagtgagtcgtattaCCAA TTTGAGAGACCTGTGC 3’

B39	ACTGCAACCGTCAATAGGAC	OL13419	5’ CTCGGACGAGTGCTGGGGCGTCG GTTTCCAAATTAACCCTCACTAAAG GGAGGGTCCTATTGACGGTTGCAG TcttccagttttatagtgagtcgtattaCCAA TTTGAGAGACCTGTGC 3’
B40	GTACTCAGCTCTGGTAGTGA	OL13420	5’ CTCGGACGAGTGCTGGGGCGTCG GTTTCCAAATTAACCCTCACTAAAG GGAGGTCACTACCAGAGCTGAGTA CcttccagttttatagtgagtcgtattaCCAA TTTGAGAGACCTGTGC 3’
B41	GGATCGGCGAATGAGCGTAC	OL13421	5’ CTCGGACGAGTGCTGGGGCGTCG GTTTCCAAATTAACCCTCACTAAAG GGAGGTACGCTCATTCGCCGATCC GcttccagttttatagtgagtcgtattaCCA ATTTGAGAGACCTGTGC 3’
B42	CCTCGGGCAACCACTTGACA	OL13422	5’ CTCGGACGAGTGCTGGGGCGTCG GTTTCCAAATTAACCCTCACTAAAG GGAGGTGTCAAGTGGTTGCCCGA GGcttccagttttatagtgagtcgtattaCC AATTTGAGAGACCTGTGC 3’
B43	TTAACGCATCCGCAGCCGTG	OL13423	5’ CTCGGACGAGTGCTGGGGCGTCG GTTTCCAAATTAACCCTCACTAAAG GGAGGCACGGCTGCGGATGCGTT AActtccagttttatagtgagtcgtattaCCA ATTTGAGAGACCTGTGC 3’
B44	GAGTCGGTGAACCAGTTATT	OL13424	5’ CTCGGACGAGTGCTGGGGCGTCG GTTTCCAAATTAACCCTCACTAAAG GGAGGAATAACTGGTTCACCGACT CcttccagttttatagtgagtcgtattaCCAA TTTGAGAGACCTGTGC 3’
B45	TGATTCGTGGAAATAATGCT	OL13425	5’ CTCGGACGAGTGCTGGGGCGTCG GTTTCCAAATTAACCCTCACTAAAG GGAGGAGCATTATTTCCACGAATC ActtccagttttatagtgagtcgtattaCCAA TTTGAGAGACCTGTGC 3’
B46	ATATCGCGCAAGTGGATGAG	OL13426	5’ CTCGGACGAGTGCTGGGGCGTCG GTTTCCAAATTAACCCTCACTAAAG GGAGGCTCATCCACTTGCGCGATA TcttccagttttatagtgagtcgtattaCCAA TTTGAGAGACCTGTGC 3’
B47	TACTACGAGCGCACAGTTAA	OL13427	5’ CTCGGACGAGTGCTGGGGCGTCG GTTTCCAAATTAACCCTCACTAAAG GGAGGTTAACTGTGCGCTCGTAGT ActtccagttttatagtgagtcgtattaCCAA TTTGAGAGACCTGTGC 3’
B48	GGGTGATCCTGATTCCAAAC	OL13428	5’ CTCGGACGAGTGCTGGGGCGTCG GTTTCCAAATTAACCCTCACTAAAG GGAGGGTTTGGAATCAGGATCACC

			CcttccagttttatagtgagtcgtattaCCAA TTTGAGAGACCTGTGC 3’
B49	TCAGTAAGAGCCGGTT CGAA	OL13429	5’ CTCGGACGAGTGCTGGGGCGTCG GTTTCCAAATTAACCCTCACTAAAG GGAGGTTCGAACCGGCTCTTACTG ActtccagttttatagtgagtcgtattaCCAA TTTGAGAGACCTGTGC 3’
B50	CCGCTAATGAGGAATG TAGG	OL13430	5’ CTCGGACGAGTGCTGGGGCGTCG GTTTCCAAATTAACCCTCACTAAAG GGAGGCCTACATTCCTCATTAGCG GcttccagttttatagtgagtcgtattaCCA ATTTGAGAGACCTGTGC 3’
B51	GGTCTACACAAATCCAC TGG	OL13431	5’ CTCGGACGAGTGCTGGGGCGTCG GTTTCCAAATTAACCCTCACTAAAG GGAGGGGTCTACACAAATCCACTG GcttccagttttatagtgagtcgtattaCCA ATTTGAGAGACCTGTGC 3’
B52	TCCTCTTTGACCTCCCTA AA	OL13432	5’ CTCGGACGAGTGCTGGGGCGTCG GTTTCCAAATTAACCCTCACTAAAG GGAGGTTTAGGGAGGTCAAAGAG GActtccagttttatagtgagtcgtattaCC AATTTGAGAGACCTGTGC 3’
B53	ATTCGCAGTCTCGTAAC CTC	OL13433	5’ CTCGGACGAGTGCTGGGGCGTCG GTTTCCAAATTAACCCTCACTAAAG GGAGGGAGGTTACGAGACTGCGA ATcttccagttttatagtgagtcgtattaCCA ATTTGAGAGACCTGTGC 3’
B54	AGCAATACGAATGGAG CATG	OL13434	5’ CTCGGACGAGTGCTGGGGCGTCG GTTTCCAAATTAACCCTCACTAAAG GGAGGCATGCTCCATTCGTATTGC TcttccagttttatagtgagtcgtattaCCAA TTTGAGAGACCTGTGC 3’
B55	ACATTACTCCGCTCCCG TTC	OL13435	5’ CTCGGACGAGTGCTGGGGCGTCG GTTTCCAAATTAACCCTCACTAAAG GGAGGGAACGGGAGCGGAGTAAT GTcttccagttttatagtgagtcgtattaCCA ATTTGAGAGACCTGTGC 3’
B56	CAGTATGCCTCTGTATA GGT	OL13436	5’ CTCGGACGAGTGCTGGGGCGTCG GTTTCCAAATTAACCCTCACTAAAG GGAGGACCTATACAGAGGCATACT GcttccagttttatagtgagtcgtattaCCA ATTTGAGAGACCTGTGC 3’
B57	ACGTGCTTGGATGGTCC GTT	OL13437	5’ CTCGGACGAGTGCTGGGGCGTCG GTTTCCAAATTAACCCTCACTAAAG GGAGGAACGGACCATCCAAGCAC GTcttccagttttatagtgagtcgtattaCCA ATTTGAGAGACCTGTGC 3’
B58	TAATCCCAACCGTTTGT CTT	OL13438	5’ CTCGGACGAGTGCTGGGGCGTCG GTTTCCAAATTAACCCTCACTAAAG

			GGAGGAAGACAAACGGTTGGGAT TActtccagttttatagtgagtcgtattaCCA ATTTGAGAGACCTGTGC 3’
B59	TGGATCTTAGGACGCA ACAT	OL13439	5’ CTCGGACGAGTGCTGGGGCGTCG GTTTCCAAATTAACCCTCACTAAAG GGAGGATGTTGCGTCCTAAGATCC ActtccagttttatagtgagtcgtattaCCAA TTTGAGAGACCTGTGC 3’
B60	CTTTGTGCACCCTTATA GTA	OL13440	5’ CTCGGACGAGTGCTGGGGCGTCG GTTTCCAAATTAACCCTCACTAAAG GGAGGTACTATAAGGGTGCACAA AGcttccagttttatagtgagtcgtattaCC AATTTGAGAGACCTGTGC 3’
B61	GATCGTGCTAGATTGAC CTA	OL13441	5’ CTCGGACGAGTGCTGGGGCGTCG GTTTCCAAATTAACCCTCACTAAAG GGAGGTAGGTCAATCTAGCACGAT CcttccagttttatagtgagtcgtattaCCAA TTTGAGAGACCTGTGC 3’
B62	TACCGCTTATTATTACG TGG	OL13442	5’ CTCGGACGAGTGCTGGGGCGTCG GTTTCCAAATTAACCCTCACTAAAG GGAGGCCACGTAATAATAAGCGGT ActtccagttttatagtgagtcgtattaCCAA TTTGAGAGACCTGTGC 3’
B63	CTTCGCAAGATGAGTCT GAG	OL13443	5’ CTCGGACGAGTGCTGGGGCGTCG GTTTCCAAATTAACCCTCACTAAAG GGAGGCTCAGACTCATCTTGCGAA GcttccagttttatagtgagtcgtattaCCA ATTTGAGAGACCTGTGC 3’
B64	TCGCTGAAGTCCTGGG ATCA	OL13444	5’ CTCGGACGAGTGCTGGGGCGTCG GTTTCCAAATTAACCCTCACTAAAG GGAGGTGATCCCAGGACTTCAGCG ActtccagttttatagtgagtcgtattaCCAA TTTGAGAGACCTGTGC 3’
B65	CCAAGTACTAATGGTCG GAT	OL13445	5’ CTCGGACGAGTGCTGGGGCGTCG GTTTCCAAATTAACCCTCACTAAAG GGAGGATCCGACCATTAGTACTTG GcttccagttttatagtgagtcgtattaCCA ATTTGAGAGACCTGTGC 3’
B66	CCTGCTATCTTATACCTC GG	OL13446	5’ CTCGGACGAGTGCTGGGGCGTCG GTTTCCAAATTAACCCTCACTAAAG GGAGGCCGAGGTATAAGATAGCA GGcttccagttttatagtgagtcgtattaCC AATTTGAGAGACCTGTGC 3’
B67	CTCATCAGTGCACTATC CGC	OL13447	5’ CTCGGACGAGTGCTGGGGCGTCG GTTTCCAAATTAACCCTCACTAAAG GGAGGGCGGATAGTGCACTGATG AGcttccagttttatagtgagtcgtattaCC AATTTGAGAGACCTGTGC 3’
B68	AACTCTTTGCACGCTCG TCC	OL13448	5’ CTCGGACGAGTGCTGGGGCGTCG

			GTTTCCAAATTAACCCTCACTAAAG GGAGGGGACGAGCGTGCAAAGAG TTcttccagttttatagtgagtcgtattaCCA ATTTGAGAGACCTGTGC 3’
B69	TTCTCTTCCCACTGGTTT CT	OL13449	5’ CTCGGACGAGTGCTGGGGCGTCG GTTTCCAAATTAACCCTCACTAAAG GGAGGAGAAACCAGTGGGAAGAG AActtccagttttatagtgagtcgtattaCCA ATTTGAGAGACCTGTGC 3’
B70	TTAGCCGACATACACAG AAA	OL13450	5’ CTCGGACGAGTGCTGGGGCGTCG GTTTCCAAATTAACCCTCACTAAAG GGAGGTTTCTGTGTATGTCGGCTA ActtccagttttatagtgagtcgtattaCCAA TTTGAGAGACCTGTGC 3’
B71	CGTCTAGAATTGCTTGG AAC	OL13451	5’ CTCGGACGAGTGCTGGGGCGTCG GTTTCCAAATTAACCCTCACTAAAG GGAGGGTTCCAAGCAATTCTAGAC GcttccagttttatagtgagtcgtattaCCA ATTTGAGAGACCTGTGC 3’
B72	ATCCTGCTTATCGATCC TTC	OL13452	5’ CTCGGACGAGTGCTGGGGCGTCG GTTTCCAAATTAACCCTCACTAAAG GGAGGGAAGGATCGATAAGCAGG ATcttccagttttatagtgagtcgtattaCCA ATTTGAGAGACCTGTGC 3’
B73	TGGTCGTCCGATGAGG TCTG	OL13453	5’ CTCGGACGAGTGCTGGGGCGTCG GTTTCCAAATTAACCCTCACTAAAG GGAGGCAGACCTCATCGGACGACC ActtccagttttatagtgagtcgtattaCCAA TTTGAGAGACCTGTGC 3’
B74	AGTTTCCGTCTCGTTAG GCC	OL13454	5’ CTCGGACGAGTGCTGGGGCGTCG GTTTCCAAATTAACCCTCACTAAAG GGAGGGGCCTAACGAGACGGAAA CTcttccagttttatagtgagtcgtattaCCA ATTTGAGAGACCTGTGC 3’
B75	TAGAGCGTAGCAGACT GAAG	OL13455	5’ CTCGGACGAGTGCTGGGGCGTCG GTTTCCAAATTAACCCTCACTAAAG GGAGGCTTCAGTCTGCTACGCTCT ActtccagttttatagtgagtcgtattaCCAA TTTGAGAGACCTGTGC 3’
B76	AAGCGAGTCCCATTACG AAC	OL13456	5’ CTCGGACGAGTGCTGGGGCGTCG GTTTCCAAATTAACCCTCACTAAAG GGAGGGTTCGTAATGGGACTCGCT TcttccagttttatagtgagtcgtattaCCAA TTTGAGAGACCTGTGC 3’
B77	CTAAATTGACCATCAGC ACC	OL13457	5’ CTCGGACGAGTGCTGGGGCGTCG GTTTCCAAATTAACCCTCACTAAAG GGAGGGGTGCTGATGGTCAATTTA GcttccagttttatagtgagtcgtattaCCA ATTTGAGAGACCTGTGC 3’

B78	CGATTATACTAATGCGTAGC	OL13458	5’ CTCGGACGAGTGCTGGGGCGTCG GTTTCCAAATTAACCCTCACTAAAG GGAGGGCTACGCATTAGTATAATC GcttccagttttatagtgagtcgtattaCCA ATTTGAGAGACCTGTGC 3’
B79	GACCGAGAGAGTTAGATATG	OL13459	5’ CTCGGACGAGTGCTGGGGCGTCG GTTTCCAAATTAACCCTCACTAAAG GGAGGCATATCTAACTCTCTCGGT CcttccagttttatagtgagtcgtattaCCAA TTTGAGAGACCTGTGC 3’
B80	CAGCTGTACGGCATGTCGCA	OL13460	5’ CTCGGACGAGTGCTGGGGCGTCG GTTTCCAAATTAACCCTCACTAAAG GGAGGTGCGACATGCCGTACAGCT GcttccagttttatagtgagtcgtattaCCA ATTTGAGAGACCTGTGC 3’
B81	GTGAGCGGAACGTAAGCATG	OL13461	5’ CTCGGACGAGTGCTGGGGCGTCG GTTTCCAAATTAACCCTCACTAAAG GGAGGCATGCTTACGTTCCGCTCA CcttccagttttatagtgagtcgtattaCCAA TTTGAGAGACCTGTGC 3’
B82	CTTCTATGTACAAGTGGCCC	OL13462	5’ CTCGGACGAGTGCTGGGGCGTCG GTTTCCAAATTAACCCTCACTAAAG GGAGGGGGCCACTTGTACATAGA AGcttccagttttatagtgagtcgtattaCC AATTTGAGAGACCTGTGC 3’
B83	GCCTGTTTCGCAATATGTAA	OL13463	5’ CTCGGACGAGTGCTGGGGCGTCG GTTTCCAAATTAACCCTCACTAAAG GGAGGTTACATATTGCGAAACAGG CcttccagttttatagtgagtcgtattaCCAA TTTGAGAGACCTGTGC 3’
B84	AGCTAGACGATCCAAGCATC	OL13464	5’ CTCGGACGAGTGCTGGGGCGTCG GTTTCCAAATTAACCCTCACTAAAG GGAGGGATGCTTGGATCGTCTAGC TcttccagttttatagtgagtcgtattaCCAA TTTGAGAGACCTGTGC 3’
B85	TCAGTTTACGGAGATGGATC	OL13465	5’ CTCGGACGAGTGCTGGGGCGTCG GTTTCCAAATTAACCCTCACTAAAG GGAGGGATCCATCTCCGTAAACTG ActtccagttttatagtgagtcgtattaCCAA TTTGAGAGACCTGTGC 3’
B86	AATCTACGTCAACCCAGTCG	OL13466	5’ CTCGGACGAGTGCTGGGGCGTCG GTTTCCAAATTAACCCTCACTAAAG GGAGGCGACTGGGTTGACGTAGA TTcttccagttttatagtgagtcgtattaCCA ATTTGAGAGACCTGTGC 3’
B87	CAGCTGGTGCATAAACAGGG	OL13467	5’ CTCGGACGAGTGCTGGGGCGTCG GTTTCCAAATTAACCCTCACTAAAG GGAGGCCCTGTTTATGCACCAGCT

			GcttccagttttatagtgagtcgtattaCCA ATTTGAGAGACCTGTGC 3’
B88	CACGGTCCCTCAACATC TTG	OL13468	5’ CTCGGACGAGTGCTGGGGCGTCG GTTTCCAAATTAACCCTCACTAAAG GGAGGCAAGATGTTGAGGGACCG TGcttccagttttatagtgagtcgtattaCCA ATTTGAGAGACCTGTGC 3’
B89	TTTACAACTGGCCAACG ACT	OL13469	5’ CTCGGACGAGTGCTGGGGCGTCG GTTTCCAAATTAACCCTCACTAAAG GGAGGAGTCGTTGGCCAGTTGTAA ActtccagttttatagtgagtcgtattaCCAA TTTGAGAGACCTGTGC 3’
B90	GATCCTCACGTCCCTAT TAC	OL13470	5’ CTCGGACGAGTGCTGGGGCGTCG GTTTCCAAATTAACCCTCACTAAAG GGAGGGTAATAGGGACGTGAGGA TCcttccagttttatagtgagtcgtattaCCA ATTTGAGAGACCTGTGC 3’
B91	AGAGGACGTGGTCAGA CAGA	OL13471	5’ CTCGGACGAGTGCTGGGGCGTCG GTTTCCAAATTAACCCTCACTAAAG GGAGGTCTGTCTGACCACGTCCTC TcttccagttttatagtgagtcgtattaCCAA TTTGAGAGACCTGTGC 3’
B92	CTTTGTCCAGGAAAGC GGCT	OL13472	5’ CTCGGACGAGTGCTGGGGCGTCG GTTTCCAAATTAACCCTCACTAAAG GGAGGAGCCGCTTTCCTGGACAAA GcttccagttttatagtgagtcgtattaCCA ATTTGAGAGACCTGTGC 3’
B93	GAATTGCACTGACAACC CTC	OL13473	5’ CTCGGACGAGTGCTGGGGCGTCG GTTTCCAAATTAACCCTCACTAAAG GGAGGGAGGGTTGTCAGTGCAAT TCcttccagttttatagtgagtcgtattaCCA ATTTGAGAGACCTGTGC 3’
B94	GCTCCCTACCTTGCACT ATG	OL13474	5’ CTCGGACGAGTGCTGGGGCGTCG GTTTCCAAATTAACCCTCACTAAAG GGAGGGCTCCCTACCTTGCACTAT GcttccagttttatagtgagtcgtattaCCA ATTTGAGAGACCTGTGC 3’
B95	TCCTATAACGCAGGGAC CTT	OL13475	5’ CTCGGACGAGTGCTGGGGCGTCG GTTTCCAAATTAACCCTCACTAAAG GGAGGAAGGTCCCTGCGTTATAG GActtccagttttatagtgagtcgtattaCC AATTTGAGAGACCTGTGC 3’
B96	AAGTCGCGGTGAATCTT TCC	OL13476	5’ CTCGGACGAGTGCTGGGGCGTCG GTTTCCAAATTAACCCTCACTAAAG GGAGGGGAAAGATTCACCGCGAC TTcttccagttttatagtgagtcgtattaCCA ATTTGAGAGACCTGTGC 3’
B97	CATCGATGTCGAATACC TTC	OL13477	5’ CTCGGACGAGTGCTGGGGCGTCG GTTTCCAAATTAACCCTCACTAAAG

			GGAGGGAAGGTATTCGACATCGAT GcttccagttttatagtgagtcgtattaCCA ATTTGAGAGACCTGTGC 3’
B98	GATATGGGTTAGCGCC GAAG	OL13478	5’ CTCGGACGAGTGCTGGGGCGTCG GTTTCCAAATTAACCCTCACTAAAG GGAGGCTTCGGCGCTAACCCATAT CcttccagttttatagtgagtcgtattaCCAA TTTGAGAGACCTGTGC 3’
B99	GGAACAGTTCGACATCT GGT	OL13479	5’ CTCGGACGAGTGCTGGGGCGTCG GTTTCCAAATTAACCCTCACTAAAG GGAGGACCAGATGTCGAACTGTTC CcttccagttttatagtgagtcgtattaCCAA TTTGAGAGACCTGTGC 3’
B100	GGCGCGATTGGTTGAA AGGG	OL13480	5’ CTCGGACGAGTGCTGGGGCGTCG GTTTCCAAATTAACCCTCACTAAAG GGAGGCCCTTTCAACCAATCGCGC CcttccagttttatagtgagtcgtattaCCAA TTTGAGAGACCTGTGC 3’
B101	CAAATGCAAGGGTCGC GCAG	OL14218	5’ CTCGGACGAGTGCTGGGGCGTCG GTTTCCAAATTAACCCTCACTAAAG GGAGGCTGCGCGACCCTTGCATTT GcttccagttttatagtgagtcgtattaCCA ATTTGAGAGACCTGTGC 3’
B102	CCCACGTCTAGTTGCTC CCT	OL14219	5’ CTCGGACGAGTGCTGGGGCGTCG GTTTCCAAATTAACCCTCACTAAAG GGAGGAGGGAGCAACTAGACGTG GGcttccagttttatagtgagtcgtattaCC AATTTGAGAGACCTGTGC 3’
B103	TTAGCAATTCCCTGTTT CAC	OL14220	5’ CTCGGACGAGTGCTGGGGCGTCG GTTTCCAAATTAACCCTCACTAAAG GGAGGGTGAAACAGGGAATTGCT AActtccagttttatagtgagtcgtattaCCA ATTTGAGAGACCTGTGC 3’
B104	TCTGCTGTAGCTTATCT TAC	OL14221	5’ CTCGGACGAGTGCTGGGGCGTCG GTTTCCAAATTAACCCTCACTAAAG GGAGGGTAAGATAAGCTACAGCA GActtccagttttatagtgagtcgtattaCC AATTTGAGAGACCTGTGC 3’
B105	CCCTTGAAGATAGACGT AAT	OL14222	5’ CTCGGACGAGTGCTGGGGCGTCG GTTTCCAAATTAACCCTCACTAAAG GGAGGATTACGTCTATCTTCAAGG GcttccagttttatagtgagtcgtattaCCA ATTTGAGAGACCTGTGC 3’
B106	CACAGATAGGAGACCG GTAC	OL14223	5’ CTCGGACGAGTGCTGGGGCGTCG GTTTCCAAATTAACCCTCACTAAAG GGAGGGTACCGGTCTCCTATCTGT GcttccagttttatagtgagtcgtattaCCA ATTTGAGAGACCTGTGC 3’

Appendix Script 1: Barcode extraction script (Rust).

```
#Notes: set m (merge count) as 500, t (threshold distance) as 5#

use clap::Parser;
use flate2::read::MultiGzDecoder;
use levenshtein::levenshtein;
use log::*;
use rand::seq::SliceRandom;
use rand::thread_rng;
use regex::Regex;
use simple_eyre::eyre::Report;
use std::collections::{HashMap, HashSet};
use std::fs::File;
use std::io::{BufRead, BufReader, BufWriter, Write};
use std::path::PathBuf;

#[derive(Parser)]
#[command(version)]
struct Args {
   /// Show log messages. Multiple -v options increase the verbosity
   #[clap(short='v', long="verbose", action=clap::ArgAction::Count)]
   verbose: u8,
   /// Write non-barcoded sequences to file
   #[clap(short = 'n', long = "unmatched", value_name = "FILE")]
   unmatched_path: Option<PathBuf>,
   /// Search expresion
   #[clap(value_name = "REGEX")]
   barcode_expression: String,
   /// Replacement expression
   #[clap(
      short = 'r',
      long = "replacement",
      value_name = "EXPR",
      default_value = "${1}"
   )]
   barcode_replacement: String,
   /// Threshold count for merging
   #[clap(
      short = 'm',
      long = "merge-count",
      value_name = "N",
      default_value = "0"
   )]
   threshold_count: u64,
   /// Threshold edit distance for merging
   #[clap(
      short = 't',
      long = "threshold-distance",
      value_name = "D",
      default_value = "1"
   )]
   threshold_distance: usize,
   /// Input fastq.gz file
```

```
    #[clap(value_name = "FILE")]
    file_path: PathBuf,
}

fn sort_barcodes(s: &mut [(String, u64)]) {
    s.sort_by(|a, b| b.1.cmp(&a.1));
}

fn count_barcodes(m: &HashMap<String, u64>) -> u64 {
    m.iter().map(|(_, i)| *i).sum()
}

fn main() -> Result<(), Report> {
    // Register the Eyre handler:
    simple_eyre::install()?;
    // Parse the CLI arguments:
    let args = Args::parse();
    // Build the logger:
    stderrlog::new()
        .module(module_path!())
        .verbosity(args.verbose as usize)
        .timestamp(stderrlog::Timestamp::Millisecond)
        .init()?;
    // Set up the RNG:
    let mut rng = thread_rng();
    // Load the input file:
    info!("parsing reads from {}", args.file_path.to_string_lossy());
    let input_file = BufReader::new(File::open(&args.file_path)?);
    let input_buffer = BufReader::new(MultiGzDecoder::new(input_file));
    // Build the regular expression:
    debug!("building barcode regular expression");
    trace!("barcode regular expression is {}", args.barcode_expression);
    let barcode_re = Regex::new(&args.barcode_expression)?;
    // Define the barcode counts:
    let mut barcode_label = String::new();
    let mut barcodes: HashMap<String, u64> = HashMap::new();
    let mut total_reads: u64 = 0_u64;
    let mut no_barcode: u64 = 0_u64;
    // If requested, open the non-barcode sequence file:
    let mut unmatched_buffer = match args.unmatched_path {
        Some(unmatched_path) => {
            info!(
                "writing non-barcoded sequences to {}",
                unmatched_path.to_string_lossy()
            );
            Some(BufWriter::new(File::create(unmatched_path)?))
        }
        None => None,
    };
    // Loop over all reads:
    debug!("processing reads");
    for read in input_buffer.lines().skip(1).step_by(4) {
        let read = read?;
        match barcode_re.captures(&read) {
```

```
            Some(c) => {
                c.expand(&args.barcode_replacement, &mut barcode_label);
                trace!("read {} barcode label is {}", read, barcode_label);
                let i = barcodes.entry(barcode_label.to_owned()).or_insert(0_u64);
                *i += 1_u64;
                barcode_label.clear();
            }
            None => {
                trace!("no barcode detected in read {}", read);
                if let Some(ref mut buffer) = unmatched_buffer {
                    writeln!(buffer, "{read}")?;
                }
                no_barcode += 1_u64;
            }
        }
        total_reads += 1_u64;
    }
    info!("processed {total_reads} reads");
    info!(
        "{}/{} ({:0.2}%) reads did not match barcode",
        no_barcode,
        total_reads,
        (no_barcode as f32 / total_reads as f32) * 100_f32
    );
    info!("{} barcodes detected", barcodes.len());

    // Now we have all the barcodes we can extract a list of the "endpoint" barcodes, i.e.
    // those that can accept merged barcodes. A barcode is an endpoint if it currently has
    // at least args.threshold_count reads associated to it.
    let endpoint_barcodes: HashSet<String> = barcodes
        .iter()
        .filter_map(|(barcode, count)| match count > &args.threshold_count {
            false => None,
            true => Some(barcode.to_owned()),
        })
        .collect();
    debug!("{} barcodes pass threshold count", endpoint_barcodes.len());

    if !endpoint_barcodes.is_empty() {
        // Get a list of the non-endpoint barcodes sorted by their count (lowest first):
        let mut non_endpoint_barcodes: Vec<(String, u64)> = barcodes
            .iter()
            .filter_map(|(barcode, count)| {
                let barcode = barcode.to_owned();
                match endpoint_barcodes.contains(&barcode) {
                    false => Some((barcode, *count)),
                    true => None,
                }
            })
            .collect();
        non_endpoint_barcodes.sort_by(|a, b| a.1.cmp(&b.1));

        // Iterate through each of the non-endpoint barcodes, and attempt to mege it into a
single on of the endpoints.
```

```
    for (barcode, count) in non_endpoint_barcodes.iter() {
      let barcode = barcode.to_owned();
      debug!(
        "barcode {barcode} count {count} <= {}; attempting to merge",
        &args.threshold_count
      );
      // Get the edit distances between this barcode and all endpoints:
      let mut end_point_distances: HashMap<String, usize> =
        HashMap::with_capacity(endpoint_barcodes.len());
      for endpoint_barcode in endpoint_barcodes.iter() {
        end_point_distances.insert(
          endpoint_barcode.to_owned(),
          levenshtein(&barcode, endpoint_barcode),
        );
      }
      // Find the minimum edit distance:
      let min_endpoint_distance: usize = end_point_distances
        .values()
        .copied()
        .min()
        .unwrap_or(0_usize);
      // Only continue if the minimum distance is suitably low:
      if min_endpoint_distance <= args.threshold_distance {
        // Get a set of all the endpoints with the minimum distance:
        let min_distance_endpoint_barcodes: Vec<String> = end_point_distances
          .iter()
          .filter_map(
            |(barcode, distance)| match distance == &min_endpoint_distance {
              false => None,
              true => Some(barcode.to_owned()),
            },
          )
          .collect();
        // Select a single endpoint from the available options:
        if let Some(selected_endpoint) = min_distance_endpoint_barcodes.choose(&mut
rng) {
          let selected_endpoint = selected_endpoint.to_owned();
          debug!("merging barcode {barcode} (count={count}) into {selected_endpoint}
(distance is {min_endpoint_distance})");
          // Move across the merged counts to the endpoint:
          *barcodes.entry(selected_endpoint).or_insert(0) += count;
          // Delete this now-merged barcode:
          barcodes.remove(&barcode);
        }
      } else {
        debug!("barcode {barcode} minumum edit distance ({min_endpoint_distance}) is
too great; not merging");
      }
    }
  } else {
    info!(
      "no barcodes have counts > {}; merging not performed",
      args.threshold_count
    );
```

```
    }

    // Print out the results:
    info!("{} reads assigned a barcode", count_barcodes(&barcodes));
    info!("{} barcodes remain after merging", barcodes.len());
    let mut barcodes: Vec<(String, u64)> =
        barcodes.iter().map(|(s, i)| (s.to_owned(), *i)).collect();
    sort_barcodes(&mut barcodes);
    for (barcode, count) in barcodes {
        println!("{}\t{}", count, barcode);
    }
    println!("{no_barcode}\tno_barcode");
    Ok(())
}
```

Appendix Script 2: Barcode counts script (R).

```
#Notes: counts barcodes, matches to canonical barcodes and unmatched#
dataset_label <- "2023-JCM-002"
io <- list(
    "all_barcodes" = file.path("..", "metadata", "all-barcodes.txt"),
    "dataset_dir" = file.path("..", "data", dataset_label)
)
# Load all the barcodes:
all_barcodes <- readLines(io$all_barcodes)
# Find all of the sample input files:
sample_names <- gsub("-barcodes.txt", "",
list.files(file.path(io$dataset_dir, "barcodes"), pattern = "*-
barcodes.txt"), fixed = TRUE)
# Load all the data:
message("loading barcode counts...")
d_raw <- sapply(sample_names, function(sample_name){
    barcode_path <- file.path(io$dataset_dir, "barcodes", sprintf("%s-
barcodes.txt", sample_name))
    res <- read.table(barcode_path, sep = "\t", header = FALSE,
colClasses=c("numeric", "character"), col.names=c("count", "barcode"),
row.names=2)
    return(res)
}, simplify=FALSE)
# Count the number of unmatched reads and save these to file:
message("loading unmatched counts...")
n_total <- sapply(d_raw, function(i){sum(i[,1])})
n_unmatched <- sapply(d_raw, function(i){i["no_barcode", 1]})
counts <- data.frame(
    matched = n_total - n_unmatched,
    unmatched = n_unmatched,
    row.names = names(d_raw)
)

# Build a matrix of all our data:
message("merging barcode data...")
barcode_counts <- data.frame(
    row.names = all_barcodes
)
for(i in names(d_raw)){
    x <- d_raw[[i]]
    barcode_counts[[i]] <- d_raw[[i]][rownames(barcode_counts), ]
```

```
}
barcode_counts[is.na(barcode_counts)] <- 0
counts$canonical <- colSums(barcode_counts) # NEW: Add the number of
"canonical barcodes"
# Write the output counts to file:
write.csv(barcode_counts, file.path(io$dataset_dir, sprintf("%s-
barcodes.csv", dataset_label)))

write.csv(counts, file=file.path(io$dataset_dir, sprintf("%s-
counts.csv", dataset_label)))
```

Appendix Script 3: Population dynamics exemplar function script (R).

```
#Note: exemplar scripts for AMA mice#

# Install and load vegan package
install.packages("vegan")
library(vegan)
#other packages#
install.packages("tidyverse")
library(tidyverse)
install.packages("ggsci")
library(ggsci)
install.packages("ggpubr")
library(ggpubr)


#shannon index#

#load data#

barcode_counts_ams<- read.csv("R:/rsrch/pmk2/lab/Ciara/Experimental
work/Barseq/Population dynamics analysis/2024-JCM-003-barcodes.csv")

#format data#
barcode_counts_ams <- subset(barcode_counts_ams, X != "TGGATCTTAGGACGCAACAT")

barcode_ams2 <- subset(barcode_counts_ams, subset = INPUT1B>0)
barcode_ams3 <- subset(barcode_ams2, subset = INPUT1B>500)


#samples list#
ttt <- names(barcode_ams3[-1])
#transform data#
barcode_ams_trans3 <- t(barcode_ams3[,-1])

#dataframe for amastigote samples#
shannon_data_ams <- data.frame(Sample = ttt,
                      Shannon =  diversity(x = barcode_ams_trans3,
                                           index = 'shannon'
                      )
)
```

```
#load data#
#amastigotes#
barcode_counts_ams <- read.csv("R:/rsrch/pmk2/lab/Ciara/Experimental
work/Barseq/Population dynamics analysis/2024-JCM-003-barcodes.csv")
barcode_counts2_ams<-subset(barcode_counts_ams, subset = INPUT1B>0)
barcode_counts3_ams<-subset(barcode_counts2_ams, subset = INPUT1B>500)
#remove odd barcode#
barcode_counts3_ams <- subset(barcode_counts3_ams, subset = X !=
"TGGATCTTAGGACGCAACAT")


#founder pop function#
#fhat is (1/total barcodes [10 here])*sum[(output freq - input freq)squared / input freq*(1-
input freq)]#
#founder pop is 1/(fhat - 1/seq reads in - 1/ seq reads out)#
#g set as 1#

founder_pop <- function(output_count, input_count){
  output_count = as.vector(output_count)
  input_count = as.vector(input_count)
  output_freq = output_count/sum(output_count)
  input_freq = input_count/sum(input_count)
  diffsq = sum(((output_freq - input_freq)^2)/(input_freq*(1-input_freq)))
  fhat = (1/(length(input_count)))*diffsq
  input_seq_total = sum(input_count)
  output_seq_total = sum(output_count)
  founderpop = 1/(fhat - (1/input_seq_total) - (1/output_seq_total))
  founderpop
}


#calculate FPs across all columns#
barcode_counts_samples_ams <- barcode_counts3_ams[-c(2:7)]
barcode_counts_samples_ams <- barcode_counts_samples_ams[-c(1)]
barcode_counts_samples_pros <- barcode_counts2_pros[-c(98:106)]
barcode_counts_samples_pros <- barcode_counts_samples_pros[-c(1)]
barcode_counts_samples_comp <- barcode_counts2_comp[-c(2:12)]
barcode_counts_samples_comp <- barcode_counts_samples_comp[-c(1)]

barcode_counts_samples_reinf <- barcode_counts2_reinf[-c(2:7)]
barcode_counts_samples_reinf <- barcode_counts_samples_reinf[-c(1)]


#FP_data_ams <- vector()
samples_ams <- colnames(barcode_counts_samples_ams)
samples_pros <- colnames(barcode_counts_samples_pros)
samples_comp <- colnames(barcode_counts_samples_comp)
samples_reinf <- colnames(barcode_counts_samples_reinf)
FP_ams <- apply(barcode_counts_samples_ams, 2, function(col_in) {
```

```
  return(founder_pop(col_in, INPUT_MEAN))
})

#data frame#
ams_FP_data <- data.frame(samples_ams, FP_ams)
shannon_data_timeline$tissue[shannon_data_timeline$tissue %in% "INPUT"] <- "INVITRO"
shannon_data_comp$tissue[shannon_data_comp$tissue %in% "INPUT"] <- "INVITRO"
ams_dataframe <- data.frame(ams_FP_data, subset(shannon_data_ams, tissue != "INPUT")
write.csv(ams_dataframe, "ams_dataframe.csv")




#GD calculations#



library(tidyverse)
library(ggsci)
library(ggpubr)

#load data#
#filter#
barcode_counts_ams <- read.csv("R:/rsrch/pmk2/lab/Ciara/Experimental
work/Barseq/Population dynamics analysis/2024-JCM-003-barcodes.csv")
barcode_counts2_ams<-subset(barcode_counts_ams, subset = INPUT1B>0)
barcode_counts3_ams<-subset(barcode_counts2_ams, subset = INPUT1B>500)
#remove odd barcode#
barcode_counts3_ams <- subset(barcode_counts3_ams, subset = X !=
"TGGATCTTAGGACGCAACAT")



#function#
#chord distance#
chord_dist = function(count1, count2){
  freq1 = count1/sum(unlist(count1))
  freq2 = count2/sum(unlist(count2))
  costheta = sum(sqrt(freq1*freq2))
  chord_dist = (2*(sqrt(2))/pi)*(sqrt(1-costheta))
  chord_dist
}



#amastigotes GD
There are no sources in the current document.#


barcode_ams3_long <- pivot_longer(
  barcode_counts3_ams,
  cols = INPUT1A:M9S,
```

```
  names_to = "Sample",
  values_to = "barcode_count"
)

#format data for mouse properties#

barcode_ams3_long$mouse_ID[barcode_ams3_long$Sample %in% grep("^M[0-9]",
                                        barcode_ams3_long$Sample, value = T)]<-
parse_number(substr(grep("^M[0-9]",

barcode_ams3_long$Sample, value = T), 2, 4))

Tissue3 <- sub('^M[0-9]+', '', barcode_ams3_long$Sample)
barcode_ams3_long <- mutate(barcode_ams3_long, "Tissue" = Tissue3)


#find the mice in current study
mice_in_study <- unique(((barcode_ams3_long$mouse_ID)))
#remove all nas
mice_in_study <- sort(mice_in_study[!is.na(mice_in_study)])
#create wide format for each mouse here
#create list that will store all dfs
list_DFs_per_mouse <- list()
#run loop for all mice
for (mouse in mice_in_study){
  #back to wide format for each mouse#
  list_DFs_per_mouse[[mouse]] <- pivot_wider(subset(barcode_ams3_long, mouse_ID ==
mouse),
                                id_cols = c("X"),
                                names_from = Tissue,
                                values_from = barcode_count)
  print(paste0("For mouse: ", mouse))
  print(head(list_DFs_per_mouse[[mouse]]))

  #now run code on subdivided dataframe for each mouse individually#
  #for GD calculation#

  df <- list_DFs_per_mouse[[mouse]][-c(1)]
  paired <- combn(colnames(df), 2)
  chord_df <- matrix(ncol = 8, nrow = 8)
  chord_df[,] <- 0
  colnames(chord_df) <- colnames(df)
  rownames(chord_df) <- colnames(df)

  # Loop through all the pairs of columns, calculate the chord distance, and
  # add the resulting value to the dataframe, using the pair of column names to
  # put it in the right place
  for (p in seq(1, ncol(paired))) {
    colA <- paired[1,p]
    colB <- paired[2,p]
    cd <- chord_dist(df[colA], df[colB])
    chord_df[colA,colB] <- cd
    chord_df[colB,colA] <- cd
  }
```

```
 chord_df_data <- data.frame(chord_df)
 chord_df_data <- mutate(chord_df_data, "comparison" = rownames(chord_df_data))

 df_long <- pivot_longer(
  chord_df_data,
  cols = BM1:S,
  names_to = "Sample",
  values_to = "GD"
 )

 #heatmap#
 ggplot(df_long, aes(x = Sample, y= comparison, fill = GD)) + geom_tile() +
  labs(fill = "Genetic Distance\n", x= "\nTissue", y = "Tissue\n") +
  scale_fill_viridis_c(option = "B", limits = c(0, 1)) +
  theme_bw() + theme(legend.title = element_text(size = 16),
                plot.title = element_text(hjust = 0.5, size = 24),
                panel.border = element_blank(),
                panel.grid.major = element_blank(),
                panel.grid.minor = element_blank(),
                axis.title = element_text(size=20),
                axis.line =
                  element_line(colour = "black"),
                axis.text.x = element_text(size = 16, colour = "black", angle = 90),
                axis.text.y = element_text(size = 16, colour = "black"),
                legend.text = element_text(size = 12, colour = "black"))
 ggsave(paste0(file = "GD_heatmap_ams/GD_dist_heatmap_",paste0("mouse_",
mouse),".png"), height = 10, width=10)

 #save data#
 write.csv(chord_df_data, file = paste0("GD_heatmap_ams/GD_dist_","mouse_",
mouse,".csv"))
```

Appendix Script 4: Simulation scripts (MATLAB).

#Notes: for original simulation#

```
% parasites and bottlenecks

%create output dataframe
sim_output_data_not_pooled_rerun = [];

for p = [0.0000001 0.0000005 0.000001 0.000005 0.00001 0.00005
0.0001 0.0005 0.001 0.005 0.01 0.05 0.1 0.5]; % individual's
probability of getting through bottleneck
% alter this as a loop for 0.000001, going up by 5x each run, up to
1
%simulating parasites traversing bottleneck (or not) with prob p of
each
%passing through. 70 barcodes total, 7 pools of 10 barcodes, then
try
```

```
% what is the 'real' bottleneck pop size (simulation) vs what is the
STAMP
% calculated pop size at different p values

% The aim here is to generate a vector of parasites passing through
the
% bottleneck
for nbar = [140 210 350]  % total number of barcodes used (test eqv
of 7 pools of 10, 15, 20, 30 and 50)
    % so we will need nbar barcodes in the experiement,
    % in total
inputmatrix = zeros(1, nbar);
outputmatrix = zeros(1, nbar);
foundervec = [];
for i=1:100 %100 iterations as repeat experiments, to compare to
pools version

 for looppower = 1
    % input vector of barcodes, call it initialvec
    % need to normalise number of parasites by using 10^7 x barcode
total
    % (7x pool sizes tested elsewhere)/7
    nindividperbarcode = floor(((10^7)*(nbar/7))/nbar);
    nindividperbarcode_int = int32(((10^7)*(nbar/7))/nbar) % number
of individiduals with a given barcode always
    initialvec = ones(1,nbar)*nindividperbarcode;

    % see who gets through the BN, call it foundervec
  for count = 1:nbar
        foundervec(count) = binornd(initialvec(count), p);
  end
    %inputs
    inputmatrix(looppower,:) = initialvec; %input for the experiment
(parasites for each barcode)
    % outputs
    outputmatrix(looppower,:) = foundervec; %This is the output from
the experiment
%     looppower
%     initialvec
%     foundervec
%     sum(foundervec>0) % number of barcodes through BN
%     sum(foundervec) % NB i.e. founder population
 end
inputmatrix
outputmatrix

% next task: use outputmatrix in MATLAB to calculate STAMP numbers
% which is a vector, size 1*looppower
% calculate input and output frequencies by dividing by sum totals
for each
% loop
% sum rows in outputmatrix
outputsum = sum(outputmatrix,2); %summing rows in matrix
output_freq = outputmatrix.*((outputsum).^-1);
output_freq %working to here
inputsum = sum(inputmatrix,2);
input_freq = inputmatrix.*((inputsum).^-1);
input_freq

% output - input
diff_matrix = output_freq - input_freq;
diff_square = diff_matrix.*diff_matrix;
%need matrix of 1*10 with all 1s to do 1-input_freq
one_matrix = ones(1,nbar);
```

```
to_sum = diff_square.*(input_freq.*(one_matrix - input_freq).^-1);
% square result
% divide by (input freq * (1- input freq)
denomfhat = ones(1,1).*nbar;
f_hat = sum(to_sum,2).*denomfhat; %for each loop, should have a row
each in final vector
% sum for each loop (matrix 1*looppower)
% divide by nbar and take inverse - STAMP bottleneck population
nombottleneckpop = ones (1,1);
bottleneck_pop = nombottleneckpop.*(f_hat.^-1);
% actual bottleneck population (counting individuals)
true_pop = sum(outputmatrix, 2);
true_pop %true bottleneck population from counting total
bottleneck_pop %STAMP calculated population values

% combine values to get probability
true_bottle_prob_matrix = true_pop;
true_bottle_prob = mean(true_bottle_prob_matrix);
STAMP_bottle_prob_matrix = bottleneck_pop; % mean of all pools
STAMP_bottle_prob = nanmean(STAMP_bottle_prob_matrix); %mean of all
pools
%matrix of probability for output
bottle_probability = ones(1, 1)*p
%matrix of barcode total for output
barcode_number = ones(1, 1)*nbar
%insert results into table - true bottleneck, bottleneck STAMP,
probability

% add data from sims to table
new_sim_data_not_pooled_rerun = [true_bottle_prob_matrix
STAMP_bottle_prob_matrix inputsum barcode_number
bottle_probability];
sim_output_data_not_pooled_rerun =[sim_output_data_not_pooled_rerun;
new_sim_data_not_pooled_rerun];
end
end
end
% simulate more probability values
% repeat for 15, 20, 30, 50 barcodes per pool
%repeat for 100 iterations
%output the simulation data table
sim_output_data_not_pooled_rerun
%export data as csv
writematrix(sim_output_data_not_pooled_rerun,
'sim_output_data_not_pooled_rerun_140bar.xls')



% parasites and bottlenecks

%create output dataframe
sim_output_data_not_pooled_rerun = [];

for p = [0.0000001 0.0000005 0.000001 0.000005 0.00001 0.00005
0.0001 0.0005 0.001 0.005 0.01 0.05 0.1 0.5]; % individual's
probability of getting through bottleneck
% alter this as a loop for 0.000001, going up by 5x each run, up to
1
%simulating parasites traversing bottleneck (or not) with prob p of
each
%passing through. 70 barcodes total, 7 pools of 10 barcodes, then
try
```

```
% what is the 'real' bottleneck pop size (simulation) vs what is the
STAMP
% calculated pop size at different p values

% The aim here is to generate a vector of parasites passing through
the
% bottleneck
for nbar = [70 100 105]  % total number of barcodes used (test eqv
of 7 pools of 10, 15, 20, 30 and 50)
    % so we will need nbar barcodes in the experiement,
    % in total
inputmatrix = zeros(1, nbar);
outputmatrix = zeros(1, nbar);
foundervec = [];
for i=1:100 %100 iterations as repeat experiments, to compare to
pools version

 for looppower = 1
    % input vector of barcodes, call it initialvec
    % need to normalise number of parasites by using 10^7 x barcode
total
    % (7x pool sizes tested elsewhere)/7
    nindividperbarcode = floor(((10^7)*(nbar/7))/nbar);
    nindividperbarcode_int = int32(((10^7)*(nbar/7))/nbar) % number
of individiduals with a given barcode always
    initialvec = ones(1,nbar)*nindividperbarcode;

    % see who gets through the BN, call it foundervec
  for count = 1:nbar
        foundervec(count) = binornd(initialvec(count), p);
  end
    %inputs
    inputmatrix(looppower,:) = initialvec; %input for the experiment
(parasites for each barcode)
    % outputs
    outputmatrix(looppower,:) = foundervec; %This is the output from
the experiment
%     looppower
%     initialvec
%     foundervec
%     sum(foundervec>0) % number of barcodes through BN
%     sum(foundervec) % NB i.e. founder population
 end
inputmatrix
outputmatrix

% next task: use outputmatrix in MATLAB to calculate STAMP numbers
% which is a vector, size 1*looppower
% calculate input and output frequencies by dividing by sum totals
for each
% loop
% sum rows in outputmatrix
outputsum = sum(outputmatrix,2); %summing rows in matrix
output_freq = outputmatrix.*((outputsum).^-1);
output_freq %working to here
inputsum = sum(inputmatrix,2);
input_freq = inputmatrix.*((inputsum).^-1);
input_freq

% output - input
diff_matrix = output_freq - input_freq;
diff_square = diff_matrix.*diff_matrix;
%need matrix of 1*10 with all 1s to do 1-input_freq
one_matrix = ones(1,nbar);
```

```
to_sum = diff_square.*(input_freq.*(one_matrix - input_freq).^-1);
% square result
% divide by (input freq * (1- input freq)
denomfhat = ones(1,1).*nbar;
f_hat = sum(to_sum,2).*denomfhat; %for each loop, should have a row
each in final vector
% sum for each loop (matrix 1*looppower)
% divide by nbar and take inverse - STAMP bottleneck population
nombottleneckpop = ones (1,1);
bottleneck_pop = nombottleneckpop.*(f_hat.^-1);
% actual bottleneck population (counting individuals)
true_pop = sum(outputmatrix, 2);
true_pop %true bottleneck population from counting total
bottleneck_pop %STAMP calculated population values

% combine values to get probability
true_bottle_prob_matrix = true_pop;
true_bottle_prob = mean(true_bottle_prob_matrix);
STAMP_bottle_prob_matrix = bottleneck_pop; % mean of all pools
STAMP_bottle_prob = nanmean(STAMP_bottle_prob_matrix); %mean of all
pools
%matrix of probability for output
bottle_probability = ones(1, 1)*p
%matrix of barcode total for output
barcode_number = ones(1, 1)*nbar
%insert results into table - true bottleneck, bottleneck STAMP,
probability

% add data from sims to table
new_sim_data_not_pooled_rerun = [true_bottle_prob_matrix
STAMP_bottle_prob_matrix inputsum barcode_number
bottle_probability];
sim_output_data_not_pooled_rerun =[sim_output_data_not_pooled_rerun;
new_sim_data_not_pooled_rerun];
end
end
end
% simulate more probability values
% repeat for 15, 20, 30, 50 barcodes per pool
%repeat for 100 iterations
%output the simulation data table
sim_output_data_not_pooled_rerun
%export data as csv
writematrix(sim_output_data_not_pooled_rerun,
'sim_output_data_not_pooled_rerun_100bar.xls')
```

#Notes: For simulation with growth rate variance#

```
% parasites and bottlenecks

%create output dataframe
sim_output_data_not_pooled_rerun = [];

for p = [0.0000001 0.0000005 0.000001 0.000005 0.00001 0.00005
0.0001 0.0005 0.001 0.005 0.01 0.05 0.1 0.5]; % individual's
probability of getting through bottleneck
% alter this as a loop for 0.000001, going up by 5x each run, up to
1
%simulating parasites traversing bottleneck (or not) with prob p of
each
```

```
%passing through. 70 barcodes total, 7 pools of 10 barcodes, then
try
% what is the 'real' bottleneck pop size (simulation) vs what is the
STAMP
% calculated pop size at different p values

% The aim here is to generate a vector of parasites passing through
the
% bottleneck
for nbar = [70 105 140 210 350]  % total number of barcodes used
(test eqv of 7 pools of 10, 15, 20, 30 and 50)
    % so we will need nbar barcodes in the experiement,
    % in total
inputmatrix = zeros(1, nbar);
outputmatrix = zeros(1, nbar);
foundervec = [];
for i=1:100 %100 iterations as repeat experiments, to compare to
pools version

 for looppower = 1
    % input vector of barcodes, call it initialvec
    % need to normalise number of parasites by using 10^7 x barcode
total
    % (7x pool sizes tested elsewhere)/7
    nindividperbarcode = floor(((10^7)*(nbar/7))/nbar);
    nindividperbarcode_int = int32(((10^7)*(nbar/7))/nbar) % number
of individiduals with a given barcode always
    initialvec = ones(1,nbar)*nindividperbarcode;

    % see who gets through the BN, call it foundervec
  for count = 1:nbar
        foundervec(count) = binornd(initialvec(count), p);
  end

  growth_variance = 0.95 + (1.05-0.95).*rand(1, nbar);
  new_foundervec = foundervec.*growth_variance

    %inputs
    inputmatrix(looppower,:) = initialvec; %input for the experiment
(parasites for each barcode)
    % outputs
    outputmatrix(looppower,:) = new_foundervec; %This is the output
from the experiment
%     looppower
%     initialvec
%     foundervec
%     sum(foundervec>0) % number of barcodes through BN
%     sum(foundervec) % NB i.e. founder population
 end
inputmatrix
outputmatrix

% next task: use outputmatrix in MATLAB to calculate STAMP numbers
% which is a vector, size 1*looppower
% calculate input and output frequencies by dividing by sum totals
for each
% loop
% sum rows in outputmatrix
outputsum = sum(outputmatrix,2); %summing rows in matrix
output_freq = outputmatrix.*((outputsum).^-1);
output_freq %working to here
inputsum = sum(inputmatrix,2);
input_freq = inputmatrix.*((inputsum).^-1);
input_freq
```

```
% output - input
diff_matrix = output_freq - input_freq;
diff_square = diff_matrix.*diff_matrix;
%need matrix of 1*10 with all 1s to do 1-input_freq
one_matrix = ones(1,nbar);
to_sum = diff_square.*(input_freq.*(one_matrix - input_freq).^-1);
% square result
% divide by (input freq * (1- input freq)
denomfhat = ones(1,1).*nbar;
f_hat = sum(to_sum,2).*denomfhat; %for each loop, should have a row
each in final vector
% sum for each loop (matrix 1*looppower)
% divide by nbar and take inverse - STAMP bottleneck population
nombottleneckpop = ones (1,1);
bottleneck_pop = nombottleneckpop.*(f_hat.^-1);
% actual bottleneck population (counting individuals)
true_pop = sum(outputmatrix, 2);
true_pop %true bottleneck population from counting total
bottleneck_pop %STAMP calculated population values

% combine values to get probability
true_bottle_prob_matrix = true_pop;
true_bottle_prob = mean(true_bottle_prob_matrix);
STAMP_bottle_prob_matrix = bottleneck_pop; % mean of all pools
STAMP_bottle_prob = nanmean(STAMP_bottle_prob_matrix); %mean of all
pools
%matrix of probability for output
bottle_probability = ones(1, 1)*p
%matrix of barcode total for output
barcode_number = ones(1, 1)*nbar
%insert results into table - true bottleneck, bottleneck STAMP,
probability

% add data from sims to table
new_sim_data_not_pooled_rerun = [true_bottle_prob_matrix
STAMP_bottle_prob_matrix inputsum barcode_number
bottle_probability];
sim_output_data_not_pooled_rerun =[sim_output_data_not_pooled_rerun;
new_sim_data_not_pooled_rerun];
end
end
end
% simulate more probability values
% repeat for 15, 20, 30, 50 barcodes per pool
%repeat for 100 iterations
%output the simulation data table
sim_output_data_not_pooled_rerun
%export data as csv
writematrix(sim_output_data_not_pooled_rerun,
'sim_output_data_not_pooled_variance_growth_rate_corrected.xls')
```

#Notes: for simulation of clonal expansion#

```
% parasites and bottlenecks

%create output dataframe
sim_output_data_not_pooled_rerun = [];

for p = [0.0000001 0.0000005 0.000001 0.000005 0.00001 0.00005
0.0001 0.0005 0.001 0.005 0.01 0.05 0.1 0.5]; % individual's
probability of getting through bottleneck
```

```
% alter this as a loop for 0.000001, going up by 5x each run, up to
1
%simulating parasites traversing bottleneck (or not) with prob p of
each
%passing through. 70 barcodes total, 7 pools of 10 barcodes, then
try
% what is the 'real' bottleneck pop size (simulation) vs what is the
STAMP
% calculated pop size at different p values

% The aim here is to generate a vector of parasites passing through
the
% bottleneck
for nbar = [70 105 140 210 350]  % total number of barcodes used
(test eqv of 7 pools of 10, 15, 20, 30 and 50)
    % so we will need nbar barcodes in the experiement,
    % in total
inputmatrix = zeros(1, nbar);
outputmatrix = zeros(1, nbar);
foundervec = [];
for i=1:100 %100 iterations as repeat experiments, to compare to
pools version

 for looppower = 1
    % input vector of barcodes, call it initialvec
    % need to normalise number of parasites by using 10^7 x barcode
total
    % (7x pool sizes tested elsewhere)/7
    nindividperbarcode = floor(((10^7)*(nbar/7))/nbar);
    nindividperbarcode_int = int32(((10^7)*(nbar/7))/nbar) % number
of individiduals with a given barcode always
    initialvec = ones(1,nbar)*nindividperbarcode;


    q = [10] %100 1000 %inverse of prob of clonal expansion
occurring for individual
    clonal_rate = [10] % multiplier for rate of growth if clonal
expansion individual
    % see who gets through the BN, call it foundervec
  for count = 1:nbar
        foundervec(count) = binornd(initialvec(count), p);
        for value = floor(1 + (q - 1).*rand())
            if value == 2
                foundervec(1, count) = clonal_rate.*foundervec(1,
count);
            end
        end
  end

  growth_variance = 0.95 + (1.05-0.95).*rand(1, nbar);
  new_foundervec = foundervec.*growth_variance

    %inputs
    inputmatrix(looppower,:) = initialvec; %input for the experiment
(parasites for each barcode)
    % outputs
    outputmatrix(looppower,:) = new_foundervec; %This is the output
from the experiment
%     looppower
%     initialvec
%     foundervec
%     sum(foundervec>0) % number of barcodes through BN
%     sum(foundervec) % NB i.e. founder population
 end
```

```
inputmatrix
outputmatrix

% next task: use outputmatrix in MATLAB to calculate STAMP numbers
% which is a vector, size 1*looppower
% calculate input and output frequencies by dividing by sum totals
for each
% loop
% sum rows in outputmatrix
outputsum = sum(outputmatrix,2); %summing rows in matrix
output_freq = outputmatrix.*((outputsum).^-1);
output_freq %working to here
inputsum = sum(inputmatrix,2);
input_freq = inputmatrix.*((inputsum).^-1);
input_freq

% output - input
diff_matrix = output_freq - input_freq;
diff_square = diff_matrix.*diff_matrix;
%need matrix of 1*10 with all 1s to do 1-input_freq
one_matrix = ones(1,nbar);
to_sum = diff_square.*(input_freq.*(one_matrix - input_freq).^-1);
% square result
% divide by (input freq * (1- input freq)
denomfhat = ones(1,1).*nbar;
f_hat = sum(to_sum,2).*denomfhat; %for each loop, should have a row
each in final vector
% sum for each loop (matrix 1*looppower)
% divide by nbar and take inverse - STAMP bottleneck population
nombottleneckpop = ones (1,1);
bottleneck_pop = nombottleneckpop.*(f_hat.^-1);
% actual bottleneck population (counting individuals)
true_pop = sum(outputmatrix, 2);
true_pop %true bottleneck population from counting total
bottleneck_pop %STAMP calculated population values

% combine values to get probability
true_bottle_prob_matrix = true_pop;
true_bottle_prob = mean(true_bottle_prob_matrix);
STAMP_bottle_prob_matrix = bottleneck_pop; % mean of all pools
STAMP_bottle_prob = nanmean(STAMP_bottle_prob_matrix); %mean of all
pools
%matrix of probability for output
bottle_probability = ones(1, 1)*p
%matrix of barcode total for output
barcode_number = ones(1, 1)*nbar
%insert results into table - true bottleneck, bottleneck STAMP,
probability

% add data from sims to table
new_sim_data_not_pooled_rerun = [true_bottle_prob_matrix
STAMP_bottle_prob_matrix inputsum barcode_number
bottle_probability];
sim_output_data_not_pooled_rerun =[sim_output_data_not_pooled_rerun;
new_sim_data_not_pooled_rerun];
end
end
end
% simulate more probability values
% repeat for 15, 20, 30, 50 barcodes per pool
%repeat for 100 iterations
%output the simulation data table
sim_output_data_not_pooled_rerun
%export data as csv
```

```
writematrix(sim_output_data_not_pooled_rerun,
'sim_output_data_not_pooled_variance_growth_rate_and_clonalexp_q_is_
0.1.xls')


% parasites and bottlenecks

%create output dataframe
sim_output_data_not_pooled_rerun = [];

for p = [0.0000001 0.0000005 0.000001 0.000005 0.00001 0.00005
0.0001 0.0005 0.001 0.005 0.01 0.05 0.1 0.5]; % individual's
probability of getting through bottleneck
% alter this as a loop for 0.000001, going up by 5x each run, up to
1
%simulating parasites traversing bottleneck (or not) with prob p of
each
%passing through. 70 barcodes total, 7 pools of 10 barcodes, then
try
% what is the 'real' bottleneck pop size (simulation) vs what is the
STAMP
% calculated pop size at different p values

% The aim here is to generate a vector of parasites passing through
the
% bottleneck
for nbar = [70 105 140 210 350]  % total number of barcodes used
(test eqv of 7 pools of 10, 15, 20, 30 and 50)
    % so we will need nbar barcodes in the experiement,
    % in total
inputmatrix = zeros(1, nbar);
outputmatrix = zeros(1, nbar);
foundervec = [];
for i=1:100 %100 iterations as repeat experiments, to compare to
pools version

 for looppower = 1
    % input vector of barcodes, call it initialvec
    % need to normalise number of parasites by using 10^7 x barcode
total
    % (7x pool sizes tested elsewhere)/7
    nindividperbarcode = floor(((10^7)*(nbar/7))/nbar);
    nindividperbarcode_int = int32(((10^7)*(nbar/7))/nbar) % number
of individiduals with a given barcode always
    initialvec = ones(1,nbar)*nindividperbarcode;


    q = [100] %10 100 1000 %inverse of prob of clonal expansion
occurring for individual
    clonal_rate = [10] % multiplier for rate of growth if clonal
expansion individual
    % see who gets through the BN, call it foundervec
  for count = 1:nbar
        foundervec(count) = binornd(initialvec(count), p);
        for value = floor(1 + (q - 1).*rand())
            if value == 2
                foundervec(1, count) = clonal_rate.*foundervec(1,
count);
            end
        end
  end

  growth_variance = 0.95 + (1.05-0.95).*rand(1, nbar);
  new_foundervec = foundervec.*growth_variance
```

```
    %inputs
    inputmatrix(looppower,:) = initialvec; %input for the experiment
(parasites for each barcode)
    % outputs
    outputmatrix(looppower,:) = new_foundervec; %This is the output
from the experiment
%     looppower
%     initialvec
%     foundervec
%     sum(foundervec>0) % number of barcodes through BN
%     sum(foundervec) % NB i.e. founder population
 end
inputmatrix
outputmatrix

% next task: use outputmatrix in MATLAB to calculate STAMP numbers
% which is a vector, size 1*looppower
% calculate input and output frequencies by dividing by sum totals
for each
% loop
% sum rows in outputmatrix
outputsum = sum(outputmatrix,2); %summing rows in matrix
output_freq = outputmatrix.*((outputsum).^-1);
output_freq %working to here
inputsum = sum(inputmatrix,2);
input_freq = inputmatrix.*((inputsum).^-1);
input_freq

% output - input
diff_matrix = output_freq - input_freq;
diff_square = diff_matrix.*diff_matrix;
%need matrix of 1*10 with all 1s to do 1-input_freq
one_matrix = ones(1,nbar);
to_sum = diff_square.*(input_freq.*(one_matrix - input_freq).^-1);
% square result
% divide by (input freq * (1- input freq)
denomfhat = ones(1,1).*nbar;
f_hat = sum(to_sum,2).*denomfhat; %for each loop, should have a row
each in final vector
% sum for each loop (matrix 1*looppower)
% divide by nbar and take inverse - STAMP bottleneck population
nombottleneckpop = ones (1,1);
bottleneck_pop = nombottleneckpop.*(f_hat.^-1);
% actual bottleneck population (counting individuals)
true_pop = sum(outputmatrix, 2);
true_pop %true bottleneck population from counting total
bottleneck_pop %STAMP calculated population values

% combine values to get probability
true_bottle_prob_matrix = true_pop;
true_bottle_prob = mean(true_bottle_prob_matrix);
STAMP_bottle_prob_matrix = bottleneck_pop; % mean of all pools
STAMP_bottle_prob = nanmean(STAMP_bottle_prob_matrix); %mean of all
pools
%matrix of probability for output
bottle_probability = ones(1, 1)*p
%matrix of barcode total for output
barcode_number = ones(1, 1)*nbar
%insert results into table - true bottleneck, bottleneck STAMP,
probability

% add data from sims to table
```

```
new_sim_data_not_pooled_rerun = [true_bottle_prob_matrix
STAMP_bottle_prob_matrix inputsum barcode_number
bottle_probability];
sim_output_data_not_pooled_rerun =[sim_output_data_not_pooled_rerun;
new_sim_data_not_pooled_rerun];
end
end
end
% simulate more probability values
% repeat for 15, 20, 30, 50 barcodes per pool
%repeat for 100 iterations
%output the simulation data table
sim_output_data_not_pooled_rerun
%export data as csv
writematrix(sim_output_data_not_pooled_rerun,
'sim_output_data_not_pooled_variance_growth_rate_and_clonalexp_q_is_
0.01.xls')
```

Appendix Script 5: MST exemplar script (R).

```
#Notes: exemplar code for AMA mice#

###amastigote mst run####


#data#
barcode_counts_ams <- read.csv("/cloud/project/2024-JCM-003-barcodes.csv")
View(barcode_counts_ams)

colnames(barcode_counts_ams)
barcode_counts_ams <- subset(barcode_counts_ams, X != "TGGATCTTAGGACGCAACAT")
barcode_counts_ams <- subset(barcode_counts_ams, INPUT1B > 500)

#long format#
barcode_counts_long_ams <- pivot_longer(
  barcode_counts_ams,
  cols = M10BM1:M9S,
  names_to = "Sample",
  values_to = "barcode_count"
)


#add column for mouse ID and tissue#

samples_ams = sapply(strsplit(barcode_counts_long_ams$Sample, split='_', fixed=TRUE),
function(x) (x[2]))
barcode_counts_long_ams <- mutate(barcode_counts_long_ams, "samples" =
samples_ams)

barcode_counts_long_ams$mouse_ID <- "NULL"
#mouse_ID = sapply(strsplit(barcode_counts_long_ams$Sample, split='_', fixed=TRUE),
function(x) (x[2]))
barcode_counts_long_ams$mouse_ID <-parse_number(sub(".*?(\\d+.*)", "\\1",
barcode_counts_long_ams$Sample))
```

```
Tissue_ams2 <- sub('^M[0-9]+', '', barcode_counts_long_ams$Sample)
barcode_counts_long_ams <- mutate(barcode_counts_long_ams, "Tissue" = Tissue_ams2)

#back to wide format for each mouse#
barcode_counts_values1_ams <- pivot_wider(subset(barcode_counts_long_ams, mouse_ID
%in% c("1")),
                          id_cols = c("X"),
                          names_from = Tissue,
                          values_from = barcode_count)
barcode_counts_values1_ams <- barcode_counts_values1_ams[-c(1)]



###READ IN FUNCTIONS####
#calcualte GD#
chord_dist = function(count1, count2){
 freq1 = count1/sum(count1)
 freq2 = count2/sum(count2)
 costheta = sum(sqrt(freq1*freq2))
 chord_dist = (2*(sqrt(2))/pi)*(sqrt(1-costheta))
 chord_dist
}


####FUNCTIONS END HERE


###for genetic distances####


#create dataframes for bet and harm metrics#
bet_data <- data.frame()
harm_data <- data.frame()

#find the mice in current study
mice_in_study <- unique(((barcode_counts_long_short_ams$mouse_ID)))
#remove all nas
mice_in_study <- sort(mice_in_study[!is.na(mice_in_study)])
#create wide format for each mouse here
#create list that will store all dfs
list_DFs_per_mouse <- list()
#run loop for all mice
for (mouse in mice_in_study){
 #back to wide format for each mouse#
 list_DFs_per_mouse[[mouse]] <- pivot_wider(subset(barcode_counts_long_short_ams,
mouse_ID == mouse),
                               id_cols = c("X"),
                               names_from = Tissue,
                               values_from = barcode_count)
 print(paste0("For mouse: ", mouse))
 print(head(list_DFs_per_mouse[[mouse]]))

 #now run code on subdivided dataframe for each mouse individually#
 #for GD calculation#
```

```
df <- list_DFs_per_mouse[[mouse]][-c(1)]
paired <- combn(colnames(df), 2)
chord_df <- matrix(ncol = 8, nrow = 8)
chord_df[,] <- 0
colnames(chord_df) <- colnames(df)
rownames(chord_df) <- colnames(df)

# Loop through all the pairs of columns, calculate the chord distance, and
# add the resulting value to the dataframe, using the pair of column names to
# put it in the right place

for (p in seq(1, ncol(paired))) {
  colA <- paired[1,p]
  colB <- paired[2,p]
  cd <- chord_dist(df[colA], df[colB])
  chord_df[colA,colB] <- cd
  chord_df[colB,colA] <- cd
}



#GD split by sample#
# matrix of the p-value of the GD?
p.mat <- cor.mtest(chord_df)


 png(paste0(file =
"project_output_GD_amastigote_new_corrected/barcode_correlation_",paste0("mouse_",
mouse),".png"), height = 10, width=15)
 corrplot(chord_df, method="pie", type = "upper", p.mat = p.mat, sig.level = 0.01, insig =
"blank", tl.col="black", tl.srt=45, order="hclust")
 dev.off()
 ################################
 ##############end###############
 ################################
 ################################
 ########MST+Clustering##########
 ################################
 ################################



 heatmap(chord_df)

 # Create a graph adjacency based on GD between cell types in  pairwise fashion.
 g <- graph.adjacency(chord_df, mode="undirected", weighted=TRUE, diag=FALSE)


 # Simplfy the adjacency object
 #g <- igraph::simplify(g, remove.multiple = TRUE, remove.loops = TRUE)
```

```
# Colour negative correlation edges as blue
E(g)[which(E(g)$weight>1)]$color <- "darkred"

# Colour positive correlation edges as red
E(g)[which(E(g)$weight<1)]$color <- "darkblue"

# Convert edge weights to absolute values
# DO NOT USE as genetic distance is never less than zero
#E(g)$weight <- abs(E(g)$weight)

# Change arrow size - NOT REQUIRED
# For directed graphs only
#E(g)$arrow.size <- 1.0

# Remove edges below a certain distance
#g <- delete_edges(g, E(g)[which(E(g)$weight>0.8)])

# Remove any vertices remaining that have no edges
g <- delete_vertices(g, igraph::degree(g)==0)

# Assign names to the graph vertices (optional)
#V(g)$name <- V(g)$name
#These woud be SITES eg Bonemarrow1 etc
V(g)$name <- colnames(chord_df)
# Change shape of graph vertices
V(g)$shape <- "sphere"
# Change colour of graph vertices
V(g)$color <- "lightblue"
# Change colour of vertex frames
V(g)$vertex.frame.color <- "white"

V(g)
E(g)

# Scale the size of the vertices to be proportional to the level of expression of
# each gene represented by each vertex
# Multiply scaled vales by a factor of 10
scale01 <- function(x){(x-min(x))/(max(x)-min(x))}
vSizes <- (scale01(apply(chord_df, 1, mean)) + 1.0) * 10

# Amplify or decrease the width of the edges
edgeweights <- E(g)$weight * 10.0
# Amplify or decrease the width of the edges
#edgeweights <- (1-E(g)$weight) * 2.0

# Convert the graph adjacency object into a minimum spanning tree based on Prim's
algorithm
mst <- mst(g, algorithm="prim")



bet <- betweenness(mst, cutoff = -1)
harm <- harmonic_centrality(mst, cutoff = -1)
bet <- bet[order(names(bet))]
```

```
 harm <- harm[order(names(harm))]
 #find communities
 #save betweenness and harmonic centrality scores#
 bet_data <- rbind(bet_data, bet)
 harm_data <- rbind(harm_data, harm)

 #find communities
 mst.communities <- cluster_louvain(mst)
 #plot the graph
 mst.clustering <- make_clusters(mst, membership=mst.communities$membership)
 V(mst)$color <- mst.communities$membership + 1


png(paste0("project_output_GD_amastigote_new_corrected/barcode_igraph","_mouse_",
mouse,".png"))
 plot(
  mst.clustering, mst,
  layout=layout.fruchterman.reingold,
  edge.curved=TRUE,
  vertex.size=vSizes,
  vertex.label.dist=-0.3,
  vertex.label.color="black",
  asp=FALSE,
  vertex.label.cex=1,
  edge.width=edgeweights,
  edge.arrow.mode=0,
  main=paste0("mouse_", mouse) # need this to run through mouse ID for title#
 )
 dev.off()
}

colnames(bet_data) <- names(bet)
write.csv(bet_data, "project_output_GD_amastigote_new_corrected/bet_data.csv")
colnames(harm_data) <- names(harm)
write.csv(harm_data, "project_output_GD_amastigote_new_corrected/harm_data.csv")
```

Appendix Script 6: Assembly process exemplar script (R).

#Notes: run on Viking, requires parallel computing#

```
#load packages#
library("tidyverse")
#load and filter dataset#
#complement mice#
barcode_counts_comp <- read.csv("barcode_counts_comp.csv")
barcode_counts2_c<-subset(barcode_counts_comp, subset =
B1P1_INPUT1_1>0)
barcode_counts3_c<-subset(barcode_counts2_c, subset =
B1P1_INPUT1_1>500)
barcode_counts3_c<-subset(barcode_counts3_c, subset = X !=
"TGGATCTTAGGACGCAACAT")
dist_mat_c <- read.csv("distance_matrix_comp.csv")
library(iCAMP)

#pivot data to long format#
```

```
long_data_c <- pivot_longer(
    barcode_counts3_c,
    cols = B1P1_DAY7F1:B1P2_M9LN,
    names_to = "Sample",
    values_to = "barcode_count"
    )
samples = sapply(strsplit(long_data_c$Sample, split='_', fixed=TRUE),
function(x) (x[2]))
long_data_c <- mutate(long_data_c, "samples" = samples)
Mouse_ID = sapply(strsplit(long_data_c$Sample, split='_', fixed=TRUE),
function(x) (x[2]))
long_data_c$mouse_ID[long_data_c$samples %in% grep("^M[0-9]",
long_data_c$samples, value = T)]<-parse_number(substr(grep("^M[0-9]",

                                                    samples,
value = T), 2, 4))
Tissue <- sub('^M[0-9]+', '', long_data_c$samples)
long_data_c <- mutate(long_data_c, "Tissue" = Tissue)

#amastigote mice#
#load and filter dataset#
barcode_counts_ams <- read.csv("2024-JCM-003-barcodes.csv")
barcode_counts2_a<-subset(barcode_counts_ams, subset = INPUT1B>0)
barcode_counts3_a<-subset(barcode_counts2_a, subset = INPUT1B>500)
barcode_counts3_a<-subset(barcode_counts3_a, subset = X !=
"TGGATCTTAGGACGCAACAT")
dist_mat_a <- read.csv("distance_matrix_ams.csv")


#pivot data to long format#
long_data_a <- pivot_longer(
    barcode_counts3_a,
    cols = INPUT1A:M9S,
    names_to = "Sample",
    values_to = "barcode_count"
    )


Mouse_ID_a <- str_replace_all(long_data_a$Sample, "INPUT1A", "NA")
Mouse_ID_a <- str_replace_all(Mouse_ID_a, "INPUT2A", "NA")
Mouse_ID_a <- str_replace_all(Mouse_ID_a, "INPUT2B", "NA")
Mouse_ID_a <- str_replace_all(Mouse_ID_a, "INPUT2C", "NA")
Mouse_ID_a <- str_replace_all(Mouse_ID_a, "INPUT1B", "NA")
Mouse_ID_a <- str_replace_all(Mouse_ID_a, "INPUT1C", "NA")
Mouse_ID <- parse_number(Mouse_ID_a)
long_data_a <- mutate(long_data_a, "mouse_ID" = Mouse_ID)
Tissue <- sub('^M[0-9]+', '', long_data_a$Sample)
long_data_a <- mutate(long_data_a, "Tissue" = Tissue)

#complement loop#
#create list that will store all dfs
list_DFs_per_mouse <- list()
#find the mice in current study
mice_in_study <- unique(((long_data_c$mouse_ID)))
#remove all nas
mice_in_study <- sort(mice_in_study[!is.na(mice_in_study)])
#loop per mouse#
for (mouse in mice_in_study){
```

```
  #back to wide format for each mouse#
  wide_df <- pivot_wider(subset(long_data_c, mouse_ID == mouse),
                                          id_cols = c("X"),
                                          names_from = Tissue,
                                          values_from =
barcode_count)
  print(paste0("For mouse: ", mouse))
  print(head(wide_df))

#now run code on subdivided dataframe for each mouse individually#

#transpose for correct orientation#
wide_df <- wide_df[-c(1,1)]
wide_df <- as.matrix(wide_df)
df_mouse <- data.frame(t(wide_df))
#name objects for bNTI#
comm=as.matrix(df_mouse)
#pd=as.matrix(dist_mat)#
#remove col and row with barcode names#
pdx= as.matrix(dist_mat_c[-c(1,1)])

#align column and row names to match across datasets#
colnames(comm) <- colnames(pdx)
rownames(pdx)<- colnames(pdx)
nworker=4 # parallel computing thread number#
rand.time=1000 # usually use 1000 for real data#
#bNTI calculation#
bNTI=bNTIn.p(comm=comm, dis=pdx, nworker = nworker, memo.size.GB = 20,
             weighted = FALSE, exclude.consp = FALSE, rand = rand.time,
             output.bMNTD = FALSE, sig.index = "SES", unit.sum = NULL,
             correct.special = TRUE, detail.null = FALSE,
             special.method = "MNTD")

#save as CSV#
write.csv(bNTI$index, paste0(file = "bNTI2_SES",paste0("_comp_mouse_",
mouse),".csv"))
#RC calculation#
RC_values=bNTIn.p(comm=comm, dis=pdx, nworker = nworker, memo.size.GB =
20,
             weighted = FALSE, exclude.consp = FALSE, rand = rand.time,
             output.bMNTD = FALSE, sig.index = "RC", unit.sum = NULL,
             correct.special = TRUE, detail.null = FALSE,
             special.method = "MNTD")

#save as CSV#
write.csv(bNTI$index, paste0(file = "RC_values2",paste0("_comp_mouse_",
mouse),".csv"))
#other RC calculator, based on BC dissimilarity#
RC_new=RC.pc(comm=comm, rand = rand.time,
         nworker = nworker, weighted = FALSE,
         sig.index="RC")

write.csv(RC_new$index, paste0(file =
"RC_values2_modBC",paste0("_comp_mouse_", mouse),".csv"))
}
```

```
#create list that will store all dfs
list_DFs_per_mouse <- list()
#find the mice in current study
mice_in_study <- unique(((long_data_a$mouse_ID)))
#remove all nas
mice_in_study <- sort(mice_in_study[!is.na(mice_in_study)])
#loop per mouse#
for (mouse in mice_in_study){
  #back to wide format for each mouse#
  wide_df <- pivot_wider(subset(long_data_a, mouse_ID == mouse),
                                          id_cols = c("X"),
                                          names_from = Tissue,
                                          values_from =
barcode_count)
  print(paste0("For mouse: ", mouse))
  print(head(wide_df))

#now run code on subdivided dataframe for each mouse individually#

#transpose for correct orientation#
wide_df <- wide_df[-c(1,1)]
wide_df <- as.matrix(wide_df)
df_mouse <- data.frame(t(wide_df))
#name objects for bNTI#
comm=as.matrix(df_mouse)
#pd=as.matrix(dist_mat)#
#remove col and row with barcode names#
pdx= as.matrix(dist_mat_a[-c(1,1)])

#align column and row names to match across datasets#
colnames(comm) <- colnames(pdx)
rownames(pdx)<- colnames(pdx)
nworker=4 # parallel computing thread number#
rand.time=1000 # usually use 1000 for real data#
#bNTI calculation#
bNTI=bNTIn.p(comm=comm, dis=pdx, nworker = nworker, memo.size.GB = 20,
             weighted = FALSE, exclude.consp = FALSE, rand = rand.time,
             output.bMNTD = FALSE, sig.index = "SES", unit.sum = NULL,
             correct.special = TRUE, detail.null = FALSE,
             special.method = "MNTD")

#save as CSV#
write.csv(bNTI$index, paste0(file = "bNTI2_SES",paste0("_ams_mouse_",
mouse),".csv"))
#RC calculation#
RC_values=bNTIn.p(comm=comm, dis=pdx, nworker = nworker, memo.size.GB =
20,
             weighted = FALSE, exclude.consp = FALSE, rand = rand.time,
             output.bMNTD = FALSE, sig.index = "RC", unit.sum = NULL,
             correct.special = TRUE, detail.null = FALSE,
             special.method = "MNTD")

#save as CSV#
write.csv(bNTI$index, paste0(file = "RC_values2",paste0("_ams_mouse_",
mouse),".csv"))
#other RC calculator, based on BC dissimilarity#
```

```
RC_new=RC.pc(comm=comm, rand = rand.time,
         nworker = nworker, weighted = FALSE,
         sig.index="RC")
write.csv(RC_new$index, paste0(file =
"RC_values2_modBC",paste0("_ams_mouse_", mouse),".csv"))
}

#for promastigote RC bray modified only#
#load and filter dataset#
barcode_counts_timeline <- read.csv("barcode_counts_timeline.csv")
barcode_counts2<-subset(barcode_counts_timeline, subset =
B2P2_INPUT1_1>0)
barcode_counts3<-subset(barcode_counts2, subset = B2P2_INPUT1_1>500)
names(barcode_counts3)[names(barcode_counts3) == "B2P2_M51LN2"] <-
"B2P2_M15LN2"
barcode_counts3<-subset(barcode_counts3, subset = Barcode !=
"TGGATCTTAGGACGCAACAT")
dist_mat <- read.csv("distance_matrix.csv")



#pivot data to long format#
long_data <- pivot_longer(
    barcode_counts3,
    cols = B2P1_M10GUT:B2P2_M9LUNG,
    names_to = "Sample",
    values_to = "barcode_count"
    )
samples = sapply(strsplit(long_data$Sample, split='_', fixed=TRUE),
function(x) (x[2]))
long_data <- mutate(long_data, "samples" = samples)
Mouse_ID = sapply(strsplit(long_data$Sample, split='_', fixed=TRUE),
function(x) (x[2]))
long_data$mouse_ID[long_data$samples %in% grep("^M[0-9]",
long_data$samples, value = T)]<-parse_number(substr(grep("^M[0-9]",

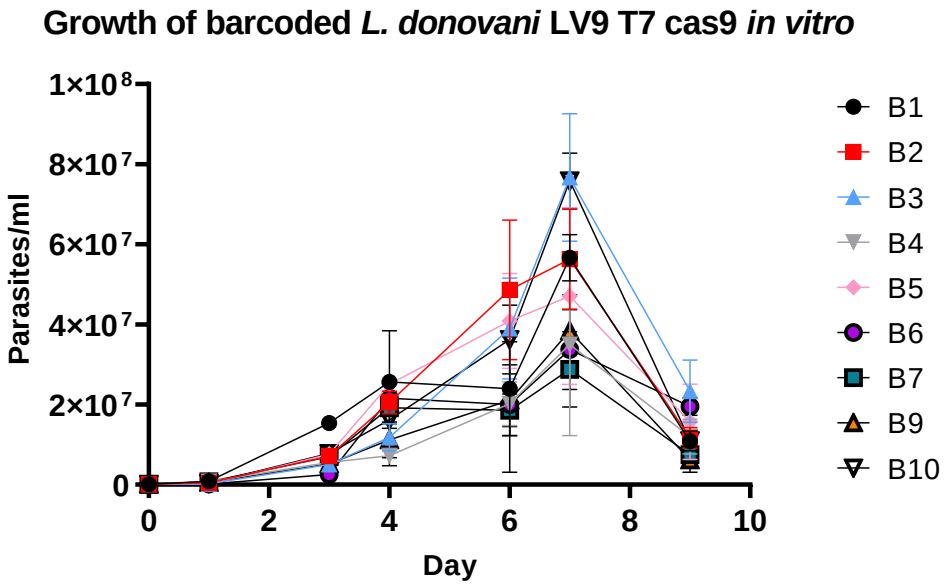
                                                    samples,
value = T), 2, 4))
Tissue <- sub('^M[0-9]+', '', long_data$samples)
long_data <- mutate(long_data, "Tissue" = Tissue)
#create list that will store all dfs
list_DFs_per_mouse <- list()
#find the mice in current study
mice_in_study <- unique(((long_data$mouse_ID)))
#remove all nas
mice_in_study <- sort(mice_in_study[!is.na(mice_in_study)])
#loop per mouse#
for (mouse in mice_in_study){
  #back to wide format for each mouse#
  wide_df <- pivot_wider(subset(long_data, mouse_ID == mouse),
                                          id_cols = c("Barcode"),
                                          names_from = Tissue,
                                          values_from =
barcode_count)
  print(paste0("For mouse: ", mouse))
  print(head(wide_df))

#now run code on subdivided dataframe for each mouse individually#
```

```
#transpose for correct orientation#
wide_df <- wide_df[-c(1,1)]
wide_df <- as.matrix(wide_df)
df_mouse <- data.frame(t(wide_df))
#name objects for bNTI#
comm=as.matrix(df_mouse)
#pd=as.matrix(dist_mat)#
#remove col and row with barcode names#
pdx= as.matrix(dist_mat[-c(1,1)])

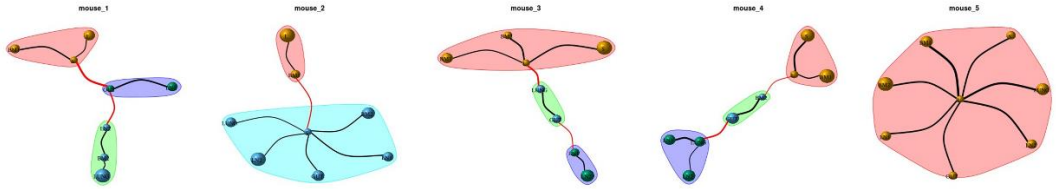
#align column and row names to match across datasets#
colnames(comm) <- colnames(pdx)
rownames(pdx)<- colnames(pdx)
nworker=4 # parallel computing thread number#
rand.time=1000 # usually use 1000 for real data#
#bNTI calculation#
bNTI=bNTIn.p(comm=comm, dis=pdx, nworker = nworker, memo.size.GB = 20,
             weighted = FALSE, exclude.consp = FALSE, rand = rand.time,
             output.bMNTD = FALSE, sig.index = "SES", unit.sum = NULL,
             correct.special = TRUE, detail.null = FALSE,
             special.method = "MNTD")

#save as CSV#
#including RC values#
write.csv(bNTI$index, paste0(file = "bNTI2_SES",paste0("mouse_",
mouse),".csv"))
#RC values#
#bNTI calculation#
RC_values=bNTIn.p(comm=comm, dis=pdx, nworker = nworker, memo.size.GB =
20,
             weighted = FALSE, exclude.consp = FALSE, rand = rand.time,
             output.bMNTD = FALSE, sig.index = "RC", unit.sum = NULL,
             correct.special = TRUE, detail.null = FALSE,
             special.method = "MNTD")
write.csv(RC_values$index, paste0(file =
"RC_values2_",paste0("_pro_mouse_", mouse),".csv"))
#other RC calculator, based on BC dissimilarity#
RC_new=RC.pc(comm=comm, rand = rand.time,
         nworker = nworker, weighted = FALSE,
         sig.index="RC")
write.csv(RC_new$index, paste0(file =
"RC_values2_modBC",paste0("_pro_mouse_", mouse),".csv"))
}
```

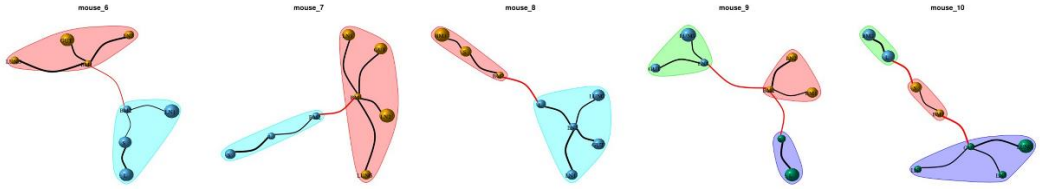


Appendix Figure 1: Growth curves for B1-10 *L. donovani* LV9 T7 cas9. Promastigote parasites from ten individual barcoded LV9 T7 cas9 lines (B1-10; see Appendix Table 1 for details) were seeded in triplicate at 1x10^5/ml in SDM-79 (completed media) and were countedat intervals over 9 days using a microscope. Error bars indicate standard deviation (n=3 per line). Figure created in GraphPad Prism.

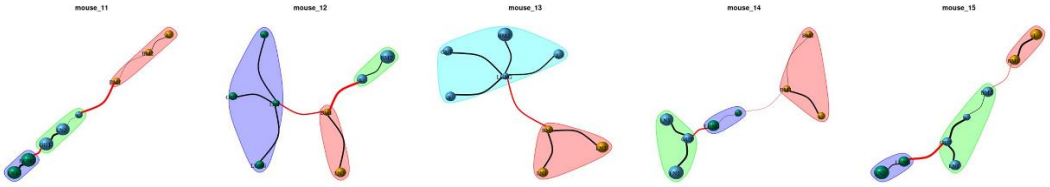
AMA; day 2; female



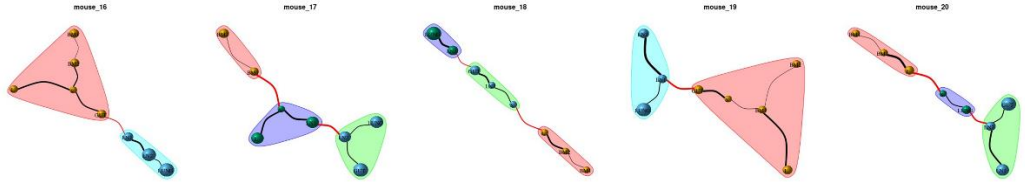
AMA; day 2; male



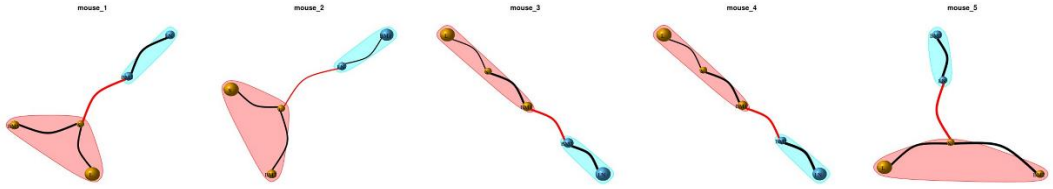
AMA; day 14; female



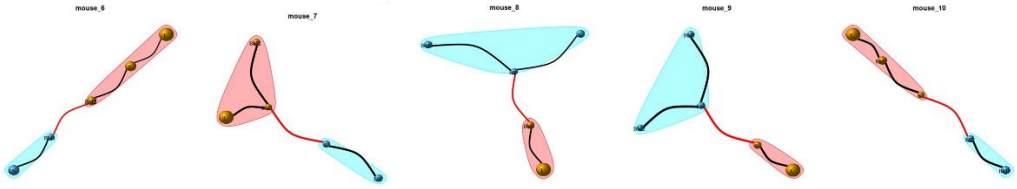
AMA; day 28; female

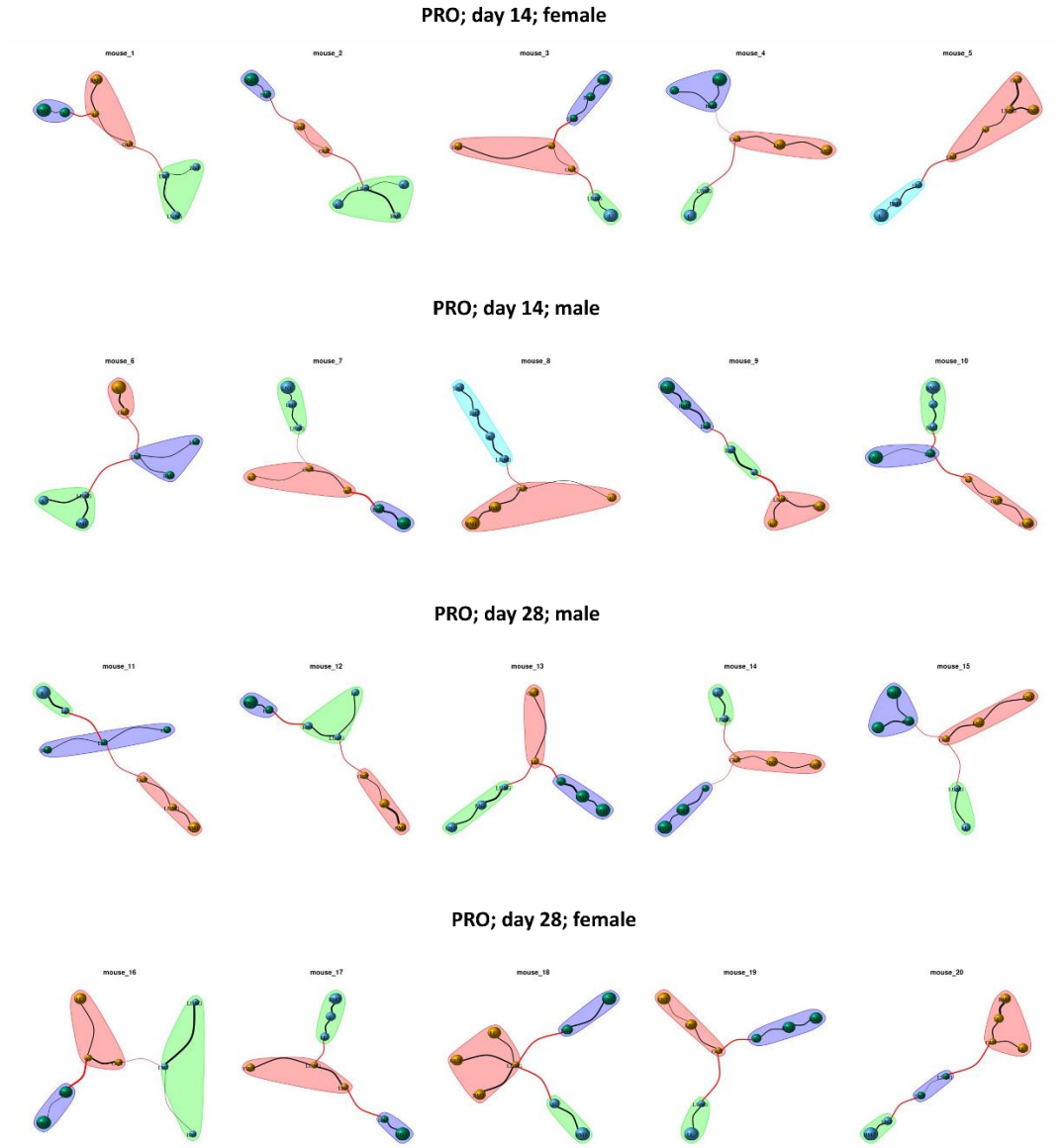


PRO; day 2; female



PRO; day 2; male





Appendix Figure 2: MST network diagrams for PRO and AMA mice. MST networks for all mice in the PRO and AMA study, calculated using the iGraph R package to generate an MST of tissue relatedness, using genetic distance (GD) as the basis, with Prim's algorithm and Louvain clustering. Cluster colours are arbitrary. Black lines indicate relatedness within clusters whilst red lines indicate relatedness between clusters. Line thickness is proportional to relatedness (1-GD). Tissues are labelled as: GUT, LUNG, L (LIVER), S (SPLEEN), LN1 or 2 (LYMPH NODE), BM1 or 2 (BONE MARROW).

Abbreviations:

AIC: Akaike information criterion

AMA: amastigote-infected mice (specifically those in Chapter 4 only)

B1-106: designation of barcoded LV9 T7 cas9 L. donovani library lines (see Appendix Table 1 for details)

B6: B6.CD45.1 or C57BL/6J mice

BIC: Bayesian information criterion

BMDC: bone marrow dendritic cell

BMDM: bone marrow derived macrophage

bMNTD: beta nearest taxon distance

bNTI: beta nearest taxon index

BSD: blasticidin S deaminase

C3KO: B6.129S4-C3tm1Crr/J (C3tm1) (in Chapter 5)

CCR: C-C chemokine receptor

CL: cutaneous leishmaniasis

CR: complement receptor

CRIg: complement receptor of the immunoglobulin superfamily

DC: dendritic cell

DEL-1: developmental endothelial locus-1

ECM: extracellular matrix

FP: founder population

GD: genetic distance

GP63: a zinc-dependent metalloprotease in Leishmania

HCR: hybridisation chain reaction

HIV: human immunodeficiency virus

HSC: hematopoietic stem cell

HYG: hygromycin resistance locus

IFN: interferon

IFNAR: IFNα/β receptor

IL-X: Interleukin-X

iNOS: inducible nitric oxide synthase

IRF: interferon regulatory factor

ISG: interferon-stimulated gene

iTreg: inducible Treg

IVIS: *in vivo* imaging system

LAMP1: lysosomal-associated membrane protein 1

LDU: Leishman-Donovan Unit

LPG: lipophosphoglycan

LPS: lipopolysaccharide

MAC: membrane attack complex

MBL: mannose-binding lectin

MCL: mucocutaneous leishmaniasis

MHC: major histocompatibility complex

MOI: multiplicity of infection

MST: minimum spanning tree

mTOR: mammalian target of rapamycin

MW: Mann-Whitney

MZ: marginal zone

NET: neutrophil extracellular trap

NK cell: natural killer cell

NO: nitric oxide

NTD: neglected tropical disease

pDC: plasmacytoid dendritic cell

p.i: post-infection

PAC: puromycin N-acetyltransferase

PKDL: post kala azar dermal leishmaniasis

PRO: promastigote-infected mice (specifically those in Chapter 4 only)

PRR: pattern-recognition receptor

RAG: *B6.Rag2*$^{-/-}$ CD45.1 mice

RC: Raup Crick index

RC_{BRAY}: Bray-modified Raup Crick Index

RNAi: RNA interference

ROS: reactive oxygen species

SES: standard effect size

SI: Shannon index

SNP: single-nucleotide polymophisms

STAMP: sequence tag-based analysis of microbial population dynamics

TGFβ: Transforming Growth Factor beta

TLR: toll-like receptor

TNFα: Tumour Necrosis Factor alpha

VL: visceral leishmaniasis

WT: C57BL/6J mice (in Chapter 5)

Zombie: Zombie is optical measurement of barcodes by *in situ* expression

βRC: beta Raup Crick