Investigating the role of the *Leishmania (Leishmania) major* HASP and SHERP genes during metacyclogenesis in the sand fly vectors, *Phlebotomus (Phlebotomus) papatasi* and *Ph. (Ph.) duboscqi*

Johannes Doehl

PhD

University of York
Department of Biology
Centre for Immunology and Infection

September 2013

I’d like to dedicate this thesis to
my parents, Osbert and Ulrike,
without whom I would never have been here.

Abstract

Leishmania parasites are the causative agents of a diverse spectrum of infectious diseases termed the leishmaniases. These digenetic parasites exist as intracellular, aflagellate amastigotes in a mammalian host and as extracellular flagellated promastigotes within phlebotomine sand fly vectors of the family Phlebotominae. Within the sand fly vector's midgut, *Leishmania* has to undergo a complex differentiation process, termed metacyclogenesis, to transform from non-infective procyclic promastigotes into mammalian-infective metacyclics. Members of our research group have shown previously that parasites deleted for the *L. (L.) major* cDNA16 locus (a region of chromosome 23 that codes for the stage-regulated HASP and SHERP proteins) do not complete metacyclogenesis in the sand fly midgut, although metacyclic-like stages can be generated in *in vitro* culture (*Sádlová et al. Cell. Micro.2010, 12, 1765-79).* To determine the contribution of individual genes in the locus to this phenotype, I have generated a range of 17 mutants in which target HASP and SHERP genes are reintroduced either individually or in combination into their original genomic locations within the *L. (L.) major* cDNA16 double deletion mutant. All replacement strains have been characterized *in vitro* with respect to their gene copy number, correct gene integration and stage-regulated protein expression, prior to phenotypic analysis.

HASPA1 was not detected in cultured promastigotes, but was expressed in mouse isolated amastigotes. Parasite mutant lines were passaged through susceptible BALB/c mice, during which HASPA2 gene containing mutant lines, in the absence of a HASPA1 gene, were shown not to develop lesions. Mouse-passaged parasites were used to infect the *L. (L.) major* specific sand fly vectors, *Ph. (Ph.) papatasi* and *Ph. (Ph.) duboscqi*. The progress of parasite metacyclogenesis was then monitored over twelve days, by midgut dissection and microscopy. Metacyclogenesis was not fully recovered in any of the replacement mutants tested. Surprisingly, HASPB protein expression could not be detected in the replacement mutants within the sand fly midgut, although HASPB protein was readily detected when the same parasite lines were cultured *in vitro*. The same was true for SHERP, although *in situ* expression was recovered in the presence of a HASPB gene, which itself did not expressed detectable HASPB protein levels. These observations suggest a requirement for one or multiple as-yet-unidentified regulatory component(s) for HASPB expression within the sand fly midgut and these are not required in culture. Quantitative PCR data suggested HASPB upregulation to be essential for metacyclogenesis completion, suggesting a sand fly specific function for HASPB.

Table of content

Abstract 3
Table of content 4
List of figures 9
List of tables 12
List of appendices 13
Acknowledgment 14
Author's declaration 15
1. CHAPTER I. – Introduction 16
1.1. The Leishmaniases 16
1.1.1. Epidemiology of the Leishmaniases 16
1.1.2. Clinical Manifestations of Leishmaniasis 18
1.1.3. Immunopathogenesis 20
1.1.4. *Leishmania* / Human Immunodeficiency Virus co-infection 25
1.1.5. Diagnosis of Leishmaniasis 25
1.1.6. Treatment 28
1.1.7. Vaccine development 29
1.2. *Leishmania* parasites 30
1.2.1. New World Model of *Leishmania* Origin 33
1.2.2. Old World model of *Leishmania* Origin 35
1.2.3. Separate origins of L. (Leishmania) spp. and L. (Viannia) spp. 35
1.2.4. *Leishmania* and its reservoirs 36
1.2.5. Leishmania lifecycle 36
1.2.5.1. Sand fly stage: Metacyclogenesis of *L. (Leishmania) spp.* 38
1.2.5.2. Sand fly stage: Metacyclogenesis of *L. (Viannia) spp.* 42
1.2.6. The *Leishmania* cell surface 42
1.2.6.1. Glycosylphosphatidylinositol-anchors 43
1.2.6.2. Lipophosphoglycans 43
1.3. Sand Fly Vectors 44
1.3.1. Sand fly Development 46
1.3.2. Structure of the sand fly alimentary canal 48
1.3.3. Midgut Physiology 53
1.4. Sand fly Vector and *Leishmania* Parasite 54
1.4.1. Manipulating host enzyme expression patterns 54
1.4.2. Evading the sand fly's complement system 56
1.4.3. Peritrophic Matrix 58
1.4.4. *Leishmania* interactions with the sand fly alimentary canal 60

1.4.5. Inhibiting sand fly gut peristalsis 60
1.4.6. Promastigote secretory gel 61
1.4.7. The stomodeal valve 65
1.4.8. Bacterial midgut flora 65
1.4.9. Sexual reproduction of *Leishmania* 67
1.4.10. Transmission to a mammalian host 67
1.5. The *Leishmania* genome 68
1.5.1. Gene transcription 69
1.5.2. Gene expression regulation 71
1.6. The *L. (L.) major* cDNA16 locus 72
1.6.1. HASPA1, HASPA2 and HASPB 74
1.6.1.1. HASPA1 and HASPA2 properties 77
1.6.1.2. HASPB properties 77
1.6.2. SHERP1 and SHERP2 properties 78
1.7. Leishmania (Viannia) braziliensis Orthologous HASP Locus 79
1.8. Project aims 82
2. CHAPTER II. – Materials and Methods 83
2.1. In silico work 83
2.1.1. Databases 83
2.1.2. Primer design 83
2.1.3. The Basic Local Alignment Search Tool (Blast) 83
2.1.4. CLUSTAL Sequence alignments 83
2.1.5. Sequencing Data Analysis 85
2.1.6. Restriction Site Determination Tools 85
2.1.7. Other Computer Softwares 85
2.2. *Leishmania* manipulation 85
2.2.1. *Leishmania* species and strains used 85
2.2.2. Culture media and culture conditions 88
2.2.3. Splitting and passaging Leishmania parasites *in vitro* 88
2.2.4. Cryo-samples 89
2.2.5. Artificial mouse infection with *L. (L.) major* 89
2.2.5.1. *L. (L.) major* passage through BALB/c mice 89
2.2.5.2. Amastigote generation and isolation 89
2.2.6. *Leishmania* homologous recombination mutant generation 90
2.2.6.1. Transfection 90
2.2.6.2. *Leishmania* clone selection 91
2.2.6.3. Growing up parasite clones 91
2.2.7. Parasite measurements 91

2.2.8. Growth Assay 92
2.2.9. Osmotaxis Assay 92
2.3. DNA Manipulation Protocols 93
2.3.1. DNA sequencing and processing 93
2.3.2. Genomic DNA extraction 93
2.3.2.1. Phenol/chloroform gDNA extraction 93
2.3.2.2. Genomic DNA extraction by blood and tissue kit (Qiagen) 94
2.3.3. PCR amplifications 94
2.3.3.1. Conventional PCR 94
2.3.3.2. Long range PCR 94
2.3.3.3. Reverse transcriptase – PCR 97
2.3.3.4. Quantitative Real Time – PCR 97
2.3.4. PCR product purification 98
2.3.5. Plasmid construction 98
2.3.5.1. 3’ A-overhang addition 98
2.3.5.2. Restriction digests protocols 98
2.3.5.3. DNA ligation 100
2.3.5.4. Transformation of chemically competent *E. coli* cells 100
2.3.5.5. Plasmid extraction from cultured *E. coli* cells 102
2.3.6. Gel electrophoresis 102
2.3.7. DNA agarose gel extraction 102
2.3.8. Ethanol precipitation 102
2.3.9. Southern blot 103
2.3.10. Amplification of genomic DNA extracts 105
2.4. mRNA Manipulation Protocol 105
2.4.1. mRNA extraction from cultured parasites 105
2.4.2. mRNA extraction from midgut derived parasites 105
2.5. Protein studies 106
2.5.1. Protein extraction 106
2.5.2. Western / Immunoblotting 106
2.5.3. Ponceau S stain of immunoblot membranes 106
2.5.4. Promastigote secretory gel (PSG) extraction 107
2.5.5. PSG detection by Dot-blot 107
2.5.6. Biotinylation assay 107
2.6. Sand fly manipulation 108
2.6.1. Sand fly strains 108
2.6.2. Artificial sand fly infections 108
2.6.3. Sand fly midgut dissection and analysis 111

2.6.4. Gene regulation in culture 112
2.7. Microscopy 112
2.7.1. Giemsa stained gut slide analysis by light microscopy 112
2.7.2. Confocal microscopy 113
2.7.2.1. Fixed parasite antibody staining 113
2.7.2.2. Live / dead staining 113
3. Chapter III. – Generating *Leishmania* HASP and SHERP replacement mutants 115
3.1. Introduction 115
3.2. Recombinant construct generation 116
3.3. *L. (L.) major* HASP and SHERP gene(s) mutant generation 119
3.3.1. Screening the *L. (L.) major* cDNA16 mutant genes 123
3.3.1.1. PCR screen of *L. (L.) major* cDNA16 mutant genes 123
3.3.1.2. Southern blots of *L. (L.) major* cDNA16 genes mutant 123
3.3.1.3. qPCR screen for replacement gene copy number 127
3.3.1.4. Western blot time courses to assess the expression and stage-regulation of the replacement genes 134
3.3.1.5. Assessing HASPA1 expression in amastigotes 135
3.3.1.6. Assessing HASPB surface localization *in vitro* by biotinylation assay 138
3.4. Parasite passage through mice 142
3.5. Growth Assay 143
3.6. Conclusions 143
4. Chapter IV. – Investigating metacyclogenesis in the sand fly 148
4.1. Introduction 148
4.2. HASP and SHERP mutant development in the sand fly midgut 148
4.2.1. Artificial sand fly infections 148
4.2.2. Sand fly dissections 149
4.2.3. Assessing *Leishmania* forward migration in the sand fly vector 150
4.2.4. Assessing *Leishmania* infection loads in the sand fly vector 151
4.2.5. Analysing *Leishmania* parasite morphology and metacyclogenesis in the sand fly vector 158
4.2.6. Effects of parasite morphological plasticity on morphological data 165
4.2.7. Discriminating between FVI metacyclics and mutant metacyclics 169
4.3. Conclusions 171
5. Chapter V. – Further investigation into the cDNA16 locus *in vitro* and *in vivo* 173
5.1. Introduction 173

5.2. Confocal microscopic analysis of HASP and SHERP localization in *Leishmania* parasites derived from culture and sand fly midguts 173
5.3. Parasite *in vitro* differentiation in 5% sucrose/PBS compared to M199 medium 180
5.4. Assessing the potential effect of sand fly midgut molecules on parasite growth in liquid medium 181
5.5. Looking at HASP and SHERP mRNA levels within the sand fly vector 183
5.6. Assessing osmotaxis capacity in cDNA16 mutant strains 189
5.7. Looking for promastigote secretory gel secretion *in vivo* 191
5.8. Conclusions 196
6. Chapter VI. – *L. (V.) braziliensis* orthologous HASP locus 198
6.1. Introduction 198
6.2. Variation in the orthologous HASP locus between clinical *Leishmania (Viannia) braziliensis* 198
6.3. Addressing the *Leishmania (Viannia) braziliensis* orthologous HASP locus repeat collapse 205
6.4. Generating *L. (V.) braziliensis* OHL double deletion mutants 220
6.5. Conclusions 226
7. Chapter VII. – Discussion 229
7.1. Part One: Metacyclogenesis of *L. (L.) major* HASP and SHERP mutants in the sand fly vector 229
7.1.1. Data summary 229
7.1.2. The HASPs and SHERPs in metacyclogenesis 231
7.1.3. Differences in mutant parasite behaviour *in vitro* and *in vivo* 234
7.1.4. Using parasite morphology for promastigote stage identification 241
7.1.5. Model of HASP and SHERP regulation mechanism 243
7.1.6. Perspective on future studies 246
7.1.7. Conclusion 247
7.2. Part Two: The Orthologous HASP Locus of *L. (V.) braziliensis* 248
7.2.1. Data Summary 248
7.2.2. The *L. (V.) braziliensis* orthologous HASP locus 248
7.2.3. Current Orthologous HASP Locus model 250
7.2.4. Perspective on future studies 250
7.2.5. Conclusion 250
Appendices 252
List of abbreviations 253
References 257

List of figures

Fig.1.1 – Geographical distribution of leishmaniasis burden 17
Fig.1.2 – Images of different leishmaniasis pathologies 19
Fig.1.3 – Immune response to Leishmania parasites inside the mammalian host 23
Fig.1.4 – Geographical distribution of VL and HIV co-infection 26
Fig.1.5 – Taxonomic tree of the mammalian infective Leishmania spp. 31
Fig.1.6 – Differences in *L. (Leishmania) spp.* and *L. (Viannia) spp.* development in the sand fly 32
Fig.1.7 – Representation of evolutionary spread of *Leishmania* across the globe 34
Fig.1.8 – Mammalian infective *Leishmania* digenetic lifecycle 37
Fig.1.9 – Schematic representation of *Leishmania* metacyclogenesis 39
Fig.1.10 – Difference in LPG side chain composition between *L. (L.) infantum (=chagasi)* and *L. (V.) braziliensis* 40
Fig.1.11 – Taxonomic tree of *Leishmania* transmitting phlebotomine sand flies 45
Fig.1.12 – Sand fly *Phlebotomus (Phlebotomus) papatasi* 47
Fig.1.13 – Sand fly development from the egg to the adult insect 49
Fig.1.14 – Schematic of the *Ph. (Ph.) papatasi* alimentary canal 50
Fig.1.15 – Images of sand fly foregut 51
Fig.1.16 – Image of dissected blood meal containing midgut 52
Fig.1.17 – Changes in mRNA levels in the sand fly midgut transcriptome 55
Fig.1.18 – Peritrophic matrix of a *Phlebotomus papatasi* female sand fly 59
Fig.1.19 – Filamentous proteophosphoglycan 63
Fig.1.20 – Changes in sand fly feeding due to PSG-plug 64
Fig.1.21 – Stomodeal valve of un-/infected *Ph. (Ph.) papatasi* & *Ph. (Ph.) duboscqi* 66
Fig.1.22 – Gene conservation in *Leishmania* 70
Fig.1.23 – Post-transcriptional gene regulation in *Leishmania* parasites 73
Fig.1.24 – Schematic of regulation of HASP and SHERP genes 75
Fig.1.25 – *L. (L.) major* cDNA16 locus and its genes 76
Fig.1.26 – OHL alignment to cDNA16 locus 80
Fig.1.27 – ClustalW alignment of *Lbr*M.1110 and *Lbr*M.1120 amino acid sequences 81
Fig.2.1 – Southern blot set-up 104
Fig.2.2 – Artificial sand fly feeding method 110

Fig.3.1 – Plasmids constructed for homologous recombination 118
Fig.3.2 – Schematic representation of mutant genotypes 121
Fig.3.3 – Relationships of *L. (L.) major* wild type and mutant lines used in this study 122
Fig.3.4 – Schematic representation of PCR screen 125
Fig.3.5 – PCR screen for construct integration in *Leishmania* mutant strains 126
Fig.3.6 – Schematic of *Sac*I restriction sites and DIG-probe binding sites 128
Fig.3.7 – Southern blot verification of construct integration 130
Fig.3.8 – qPCR for gene copy number verification 133
Fig.3.9 –Western blot time-courses of mutant lines 137
Figure 3.10 – Western Blot of Amastigote Lysates 139
Figure 3.11 – Biotinylation Assay 141
Figure 3.12 – Growth Assays 145
Fig.4.1 – Parasite localization in the sand fly midgut 152
Fig.4.2 – Apparent presence and absence of PSG in the TMG in different L. (L.) major strains 153
Fig.4.3 – Comparing light microscopic infection load results with qPCR results for day 12 PBM 154
Fig.4.4 – Parasite infection load by light microscopy in the sand fly midgut 156
Fig.4.5 – Variability of observed infection intensity 157
Fig.4.6 – Complete morphology data set of sand fly midgut derived parasites 160
Fig.4.7 – Representation of statistically significant difference in parasite form distribution of summarized data 162
Fig.4.8 – Parasite morphology at day 12 PBM in the sand fly midgut 163
Fig.4.9 – Variability of observed parasite morphology between infection rounds 164
Fig.4.10 – Limitations of measurements to determine parasite form 166
Fig.4.11 – Parasite morphology data with adjusted threshold 168
Fig.4.12 – Differences in cell shape among metacyclics from different strains 170
Fig.5.1 – Confocal images detecting HASPB in culture derived parasites 175
Fig.5.2 – Confocal images detecting SHERP in culture derived parasites 177
Fig.5.3 – Confocal images detecting HASPB in midgut derived parasites 178
Fig.5.4 – Confocal images detecting SHERP in midgut derived parasites 179
Fig.5.5 – HASP and SHERP expression in parasites grown in M199 versus 5% Sucrose/PBS 182

Fig.5.6 – Western blot to assess potential changes in HASPB expression profiles in the presence of midgut extracts 184
Fig.5.7 – HASP and SHERP mRNA levels relative to NMT mRNA levels 187
Fig.5.8 – PCR amplification to determine mRNA length 190
Fig.5.10 – Osmotaxis assay 192
Fig.5.11 – Dot blot of PSG containing supernatants from midgut extracts 194
Fig.5.12 – Dot blot of PSG extracts from midgut 195
Fig.6.1 – Repeat collapse in OHL 199
Fig.6.2 – Screen shot of Artemis software showing the reading depth for the OHL genes LbrM.23.1110 and LbrM.23.1120 200
Fig.6.3 – Southern blot of three gDNA digests for two L. (V.) braziliensis strains 202
Fig.6.4 – Expected and measured sizes for the gDNA digests of LTB 300 and *Lbr*2904, respectively 203
Fig.6.5 – Southern blot of *Sac*I digested gDNA of LTB 300 and *Lbr*2904 204
Fig.6.6 – Southern blot of *Sac*I digested GenomiPhi amplified gDNA from clinical *L. (V.) braziliensis* isolates to assess intra OHL variablilty 206
Fig.6.7 – Schematic for placement of primer in OHL used to amplify fragments for subcloning 207
Fig.6.8 – PCR amplification of parts of the OHL for subcloning 208
Fig.6.9 – Two heterozygous copies of LbrM.23.1120 210
Fig.6.10 – Amino acid alignment of 1120_v1 and 1120_v2 212
Fig.6.11 – Conserved intergenic region between OHL genes 214
Fig.6.12 – Sequence alignment of down-stream LbrM.23.1110 gene copy with 3'flanking region of OHL 216
Fig.6.13 – Multiple copies and polymorphisms in LbrM.23.1110 218
Fig.6.14 – Alignment of distinct 1110 amino acid sequences 219
Fig.6.15 – PCR screen for verification of OHL single deletion 221
Fig.6.16 – Orthologous HASP locus deletion constructs in plasmid vectors 222
Fig.6.17 – Southern blot to verify successful single OHL deletion 223
Fig.6.18 – qPCR verification of OHL copy number in single deletion mutants 225
Fig.6.19 – Proposed *L. (V.) braziliensis* 2904 OHL map 228
Fig.7.1 – HASPB detection in midgut extracts 232
Fig.7.2 – Differences in HASPB-GFP fusion gene expression 239
Fig.7.3 – Problems with plasticity of parasite morphology 242
Fig.7.4 – Model for HASP & SHERP regulation *in vivo* during metacyclogenesis 245

List of tables

Table 1.1 – Human infective *Leishmania spp.* according to disease tropism 21
Table 1.2 – Manipulation of sand fly midgut enzyme up- and downregulation during blood meal digestion by *L. (L.) major* 57
Table 2.1 - Bioinformatics tools used in this study 84
Table 2.2 – *Leishmania* mutant strains 86
Table 2.3 – PCR profile 95
Table 2.4 – Plasmids 99
Table 2.5 – Restriction Enzymes used in this study 101
Table 2.6 – Sand Fly Vector Strains used in this study 109
Table 3.1 – Expected band sizes according to primer pair used 126
Table 3.2 – Expected band sizes for probes per integrated construct 131
Table 3.3 – Statistics on growth rates (day 0 – 3) 146
Table 5.1 – Measurements of Parasites on Confocal Images 176
Table 5.2 – Attraction Coefficient, Means and Standard Deviations 192
Table 6.1 – Fragment size calculated by Fragment Size Calculator* 202
Table 6.2 – SNPs with the LbrM.23.1110 and Lbr.23.1120 ORFs 219

List of appendices

Appendix 1 – List of human infective Leishmania spp. 252
Appendix 2 – Primers used for qPCR 252
Appendix 3 – Primers used for construct generation 252
Appendix 4 – Primers used DIG-probe generation 252
Appendix 5 – Primers used for OHL re-sequencing 252
Appendix 6 – Other primers used in this study 252
Appendix 7 – Raw data of growth assay (parasites - ml) 252
Appendix 8 – Raw Data Parasite Localization in the Sand Fly Midgut 252
Appendix 9 – Raw Data of Midgut Infection Loads (from Light Microscopy) 252
Appendix 10 – Raw Data of Midgut Infection Loads (from qPCR) 252
Appendix 11 – Measurements of sand fly midgut derived Leishmania parasites 252
Appendix 12 – Hypothetical ORFs for SHERP homologue in intergenic region 252
Appendix 13 – Leishmania (Viannia) braziliensis orthologous HASP locus (Version 1) 252
Appendix 14 – Leishmania (Viannia) braziliensis orthologous HASP locus (Version 2) 252
Appendix 15 – Complete LbrM.23.1110 sequences flanked up- and downstream by intergenic region 252
Appendix 16 – Complete LbrM.23.1110 sequences flanked upstream by the intergenic region and downstream by OHL 3' flanking region 252

Acknowledgment

I would like to extend my deepest thanks to my supervisor, Deborah F. Smith, for her continuous support and the brilliant opportunity of this project. I also want to thank my training advisory panel members, Paul Kaye and Marjan van der Woude, whose continuous input helped me reach my goal.

From the Charles University in Prague, Czech Republic, I want to thank Petr Volf for having me in his lab and giving me great insights into the biology of sand flies. I need to thank Jovana Sádlová, who helped me do the sand fly infections and without whom I would not have been able to do this work. I learned a lot from her. I also must thank Jan Votýpka, who performed the qPCR on parasite infection numbers in the sand flies for me. My thanks also go out to the members of the Volf group for having been so kind to me, while working alongside them.

My thanks go to Matthew E. Rogers from the London School of Hygiene and Tropical Medicine, who supplied me with the LT15 antibody. I must thank from the University of York all the past and present members of the Smith group, who helped me to find my way and were always there to give me support, when I needed it; my thanks go in particular to Helen Price, Barbara Smith and Michael Hodgkinson. From the technology facility at the department of biology I must thank Karen Hodgkinson for helping me with the cell-sorting on the MoFlo and the rest of the imaging staff helping me solve my troubles with the confocal microscopes. I am also thanking the countless people, who gave me opportunity to discuss me work and gave me good ideas about how to proceed in my project.

Last but not least I must thank my family for their emotional support, in particular my parents Osbert and Ulrike and my brother Leonard, who endured me patiently during my write up. My thanks also go to my beloved girlfriend, Aline I. Arakaki, whose moral support has been invaluable to me over the final months of writing this thesis. I also cannot forget the friends I have in the various corners of the world, who always gave me something more and helped me with some laughs and distractions to remember the fun side of life and science.

Author's declaration

All the work presented in this thesis has been done by me, if not stated otherwise. The sand fly infection work presented in chapters IV was done in collaboration with Dr. Jovana Sádlová from Petr Volf's laboratory at the Charles University in Prague, Czech Republic.

1. CHAPTER I. – Introduction

1.1. The Leishmaniases

Leishmaniasis is the general term for a diverse spectrum of infectious diseases caused by kinetoplastid parasite species of the genus *Leishmania* (Appendix 1). They are transmitted by phlebotomine sand flies of the sub-family *Phlebotominae*. The leishmaniases are among the most complex of all vector borne diseases and occur in the sub-/tropical regions of all continents except Antarctica.

1.1.1. Epidemiology of the Leishmaniases

In 1991, Philippe Desjeux delivered a report to the World Health Organization (WHO) regarding the global burden of the leishmaniases (*1*). He considered 82 countries and territories to be endemic to at least one form of human leishmaniasis. In 1996, Desjeux published a second review considering 88 countries and territories to be endemic to the leishmaniases with 350 million people at daily risk of infection (Fig.1.1) (*2*). In 2012, Alvar *et al.* published the first update on the empirical database delivered by Desjeux (1991) considering 98 countries and 3 territories on five continents as being endemic to at least one form of human leishmaniasis (*3*). Australia was not included in this list, because the recently discovered *Leishmania* species (*spp.*) infecting red kangaroos has not yet been shown to be human infective (*4*). At least 72 of these 101 countries and territories are undeveloped countries, of which 13 are among the least developed in the world. 21 of the 101 countries and territories belong to the 'New World' (the Americas) with Brazil being the biggest disease hotspot followed by Colombia and Peru, while the remaining 80 countries and territories belong to the 'Old World' (17 of those in Europe around the Mediterranean and Black Sea) – the main hotspots being East Africa, Iran, Afghanistan and South Asia (*3*).

By 2004, a total of >12 million leishmaniases cases had been registered worldwide causing an estimated 2.4 million disability-adjusted-life-years, although the number of unregistered case is estimated to be much higher (*5*). This is because reporting of leishmaniasis is mandatory in only 33 of all endemic countries (*3*). Annual incidence rates for new leishmaniasis case are currently estimated at between 0.9 – 1.6 million cases worldwide, which is still considered to be underestimated due to under-reporting (*3*). Many cases occur in remote areas with low infrastructure and difficult access to medical facilities and, therefore, remain undetected, but also un- or misdiagnosis of

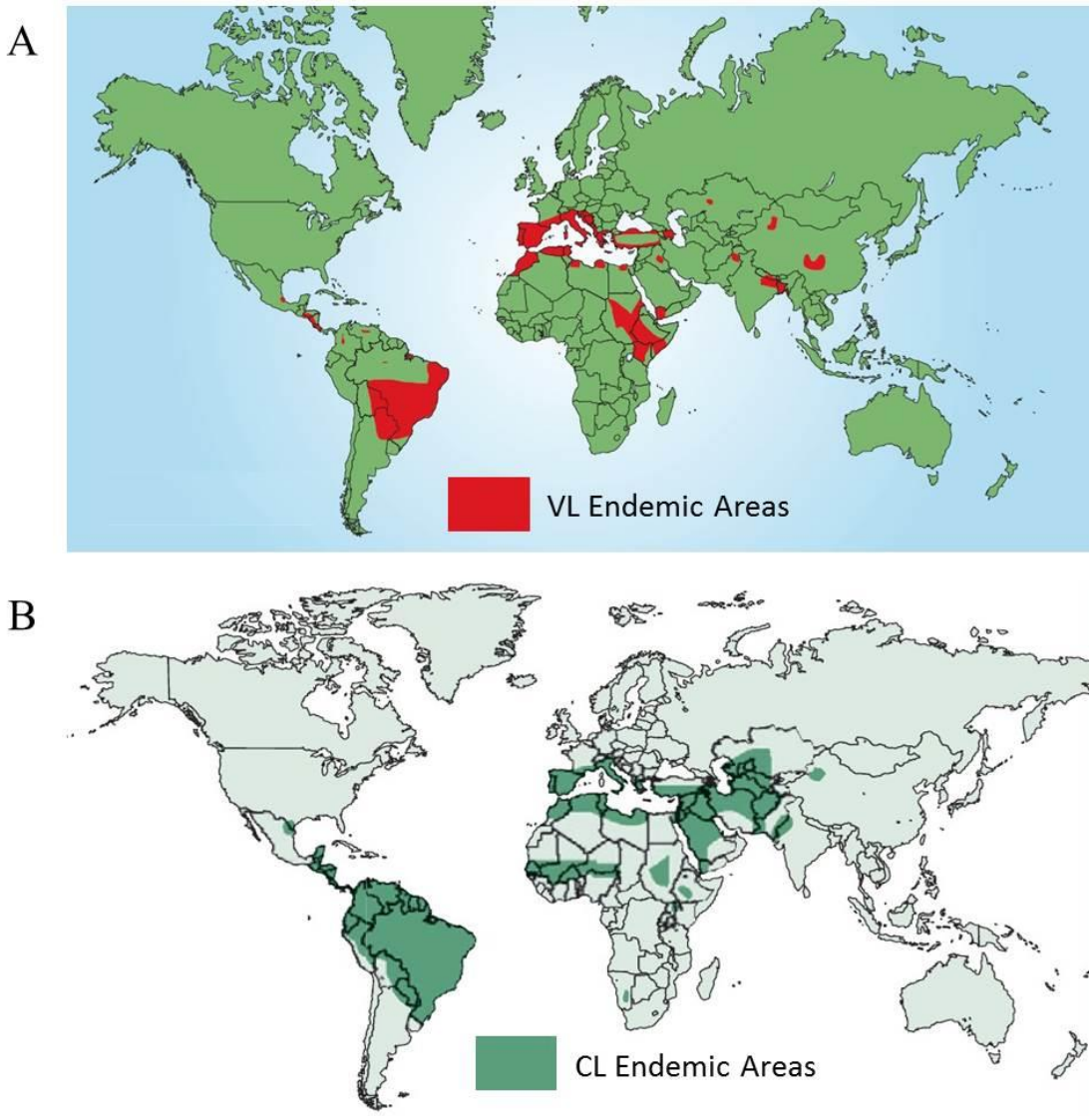


Fig.1.1 – Geographical distribution of leishmaniasis burden

No recent map of global leishmaniasis burden incorporating the recently published update on leishmaniasis epidemiology from Alvar *et al.* (2012) is currently available; hence, the maps shown are still giving an incomplete picture of the real global disease burden (*3*). A) shows the distribution of visceral leishmaniasis (VL) (*6*) and B) the distribution of all cutaneous leishmaniases (CL) forms (*7*).

infection is a common problem. Studies comparing active house-to-house surveys in endemic areas with official numbers showed that official numbers were on average 2 – 10-fold lower than those found by active case detection; in some cases even 40 – 47-fold lower (*8*, *9*). Although levels of morbidity due to cutaneous leishmaniasis (CL) are much higher than levels of mortality due to visceral leishmaniasis (VL), an estimated ~70,000 annual deaths worldwide are associated with the leishmaniases (~59,000 deaths due to VL; rest due to complications with CL forms) (*5*). Although estimates of mortality are difficult since data are sparse and mostly from deaths in hospitals, the leishmaniases are considered the second most severe protozoan disease after malaria and the third most important vector-borne disease in the world (*5*, *10*).

The leishmaniases are traditionally sylvatic zoonoses and human infection is generally considered to be accidental and occurs when humans invade endemic areas (*11*). However, since sand fly vectors were observed to adapt quickly to human induced environmental changes (*12*), they have adapted to peridomestic environments, bringing *Leishmania spp.*, which are able to use domestic animals, in particular domestic dogs, as reservoirs, into the proximity of human habitats (*13*). Some *Leishmania spp.* have even become anthroponotic in some parts of the world, such as the causative agent of VL in India, *L. (Leishmania) donovani* (*14*) and there are reports of anthroponotic CL due to *L. (L.) tropica* in Afghanistan (*15*). There are also observations of anthroponotic behaviour of *L. (L.) infantum*, primarily in human immunocompromised patients, like human immunodeficiency virus (HIV) co-infected patients (*16*). Vector-independent *Leishmania* transmission by needle sharing between HIV positive intravenous drug users (IVDU) (*17*) and blood transfusions from asymptomatic blood donors has also been reported (*18*, *19*). Fear of increased spread and parasite domestication of other *Leishmania spp.* in other parts of the world grows (*20*). Human movement and activity in formerly-sylvatic endemic areas due to urbanization, work and war have contributed markedly to the spread of the disease, forcing parasite and vector domestication and displacement of sylvatic *Leishmania* reservoirs (*5*, *11*). The increase of global travel and transport has added to disease spread (*21*).

1.1.2. Clinical Manifestations of Leishmaniasis

Three basic forms of leishmaniasis are clinically distinguished based on the tissues involved: CL, mucocutaneous leishmaniasis (MCL) (Fig.1.2H) and VL (Fig.1.2J). CL is further subdivided into non-/self-healing localized cutaneous

Fig.1.2 – Images of different leishmaniasis pathologies

A-C) Forearm with self-healing LCL skin lesions (*22*); D) Forearm with LRC (*23*); E) Forearm with non-self-healing ulcerating LCL skin lesion (Archives of UCSF); F) Dorsal view of DCL presenting male (*24*); G) Dorsal view of DL infected male (*25*); H) Severe MCL (*26*); I) PKDL (WHO 2013); J) VL (WHO 2004)

leishmaniasis (LCL) (Fig.1.2A-D) and the more unusual forms, diffuse cutaneous leishmaniasis (DCL) affecting several body parts (Fig.1.2F) and leishmaniasis recidiva cutis (LRC) (Fig.1.2E), a severe disease relapse in old scars. There is evidence that another CL form occurs in Brazil, termed disseminated cutaneous leishmaniasis (DL) (Fig.1.2G), which should be distinguished from DCL (*25*). Unusual CL forms are frequently misdiagnosed (*27*). In East Africa mucosal leishmaniasis (ML) is distinguished from MCL, because of its origin, but both affect the mucosa. MCL starts out as LCL, while ML is a secondary effect of VL without prior cutaneous involvement. Post kala-azar dermal leishmaniasis (PKDL) (Fig.1.2I) is a cutaneous infection, which can occur in patients after resolution of VL as a relapse form (*28*).

It is accepted that particular *Leishmania spp.* cause preferentially one disease type, but since disease development depends on the host immune response, too, some *Leishmania spp.* have been isolated in a broad variety of clinical manifestations (Table 1.1). For instance, *L. (L.) amazonensis* has been isolated from CL, MCL, PKDL and VL patients (*29*). The greatest global hot spots of LCL, DCL and DL cases (70-75%) are in Afghanistan, Algeria, Brazil, Colombia, Costa Rica, Ethiopia, Iran, North Sudan, Peru and Syria (*3*). MCL is primarily found in Brazil and >90% of all registered VL cases worldwide occur principally in Bangladesh, India, Nepal, Sudan and Brazil (*3*).

1.1.3. Immunopathogenesis

The immunology of leishmaniasis is highly complex and not fully understood (*30*, *31*). During *Leishmania* transmission to a new mammalian host by sand fly bite, 500 – 1000 metacyclic promastigotes are on average injected into the host‘s skin (*32*). In the mammalian host, parasites reside as amastigotes in phagolysosome-like structures, which are hostile environments designed for pathogen degradation. *Leishmania* has developed special mechanisms to promote its survival in the mammalian host’s macrophages (*33*). Before parasite reach the phagolysosome, however, they are immediately confronted with the host’s alternatively activated complement system after inoculation into the host skin. Metacyclics have a particularly dense lipophosphoglycan (LPG) coat, with specific side chain modifications and surface proteins, like gp63, which is the single most dominant surface protein in metacyclics. Gp63 – a surface glycoprotein – is a zinc-metalloprotease that converts the complement protein C3b – one of the most potent immune opsonins, which binds to LPG with high affinity, while others like C5b-9 cannot (*34*, *35*) – into

Table 1.1 – Human infective *Leishmania spp.* according to disease tropism

Subgenus	***L.(Leishmania)***	***L.(Leishmania)***	***L.(Viannia)***	***L.(Viannia)***
Old World	*L. donovani* *L. infantum* *L. tropica* º *L. major* º	*L. major* *L. torpica* *L. killicki* * *L. aethiopica* *L. infantum*		
New World	*L. infantum* *(= L. chagasi)* *L. amazonenesis*º *L. mexicana* º	*L. infantum* *(= L. chagasi)* *L. mexicana* *L. pifioni* * *L. venezuelensis* *L. gamhami* * *L. amazonensis*	*L. braziliensis* *L. guyanensis* *L. panamensis* *L. peruviana* *L. shawi* *L. naiffi* *L. lainsoni* *L. lindenbergi* *L. colombiensis*[+]	*L. braziliensis* *L. panamensis* *L. amazonensis* º *L. guyanensis* º
Principal tropism	Viscerotropic	Dermotropic	Dermotropic	Mucotropic

º Rare reported cases * Species status is under discussion [+] Taxonomic position is under discussion

This table was adapted from WHO control of leishmanasis (2010)

inactive C3b (iC3b) preventing its efficient binding to the *Leishmania* LPG coat and with that complement-mediated lysis (*37*). Protection against complement lysis is also achieved by shedding the membrane attack complex with the LPG from the surface of metacyclics (*35*).

Once past the complement system, *Leishmania* parasites encounter cells of the innate and adaptive immune system. Macrophages are the most dominant innate immune cells in the skin and will attempt to clear *Leishmania* parasites by phagocytosis. This internalization of *Leishmania* parasites effectively hides them from the humoral branch of the mammalian immune response. *Leishmania* have been shown to even possess mechanisms which actively enhance macrophage phagocytosis (*38*). However, for parasite persistence, nitric oxide (NO) synthesis needs to be prevented (*30*). Work done in *L. (L.) major* mouse (in particular, C57BL/6 and BALB/c) models revealed two distinct adaptive immunity pathways, which are predominantly mediated by T lymphocytes (Fig.1.3). One pathway is the T helper (Th) 1 response by classically activating macrophages (CAMΦ) and the other the Th2 response by alternatively activating macrophages (AAMΦ). In the draining lymph nodes, which are crucial for mounting a Th1 response (*39*), naïve T cells are stimulated to differentiate into $CD4^+$ Th1 and Th2-type effector cells by interleukin (IL) 12 and IL-4, respectively, which are mutually exclusive, secreted from antigen presenting cells (APCs) (*40*). Th1 cells in turn produce tumor necrosis factor (TNF-*α*) and, in particular, interferon-*γ* (IFN-*γ*), which promotes nitric oxide synthase (NOS) generation via a toll-like receptor (*41*). NOS synthesises NOs from L-arginine capable of killing *Leishmania* parasites. Conversely, Th2 cells produce IL-4, IL-10, IL-13 and transforming growth factor-β (TGF-*β*), which promote arginase-1 expression. Arginase-1 uses L-arginine as substrate, too, to produce L-ornithine, which is the first building block for polyamines essential for *Leishmania* survival (*42*). Therefore, *Leishmania* parasites seek to promote Th2 cell responses for their survival by actively manipulating the host immune response. Excessive L-arginine depletion by AAMΦ renders T cells hyporesponsive to antigen-stimulation and impairs their proliferation, cytokine production and T cell receptor signalling (*43*). The emergence of IL-10 secreting $CD4^+$ $CD25^+$ T regulator (T_{reg}) cells during a Th2 cell response are important for *Leishmania* persistence, too (*44*). IL-10 is an antagonist of IFN-*γ* (*45*). The Th2 cell secreted TGF-*β* is also a potent IFN-*γ* antagonist (*46*). Interestingly, TGF-*β* is also produced by neutrophils after exposure to apoptotic promastigotes, which

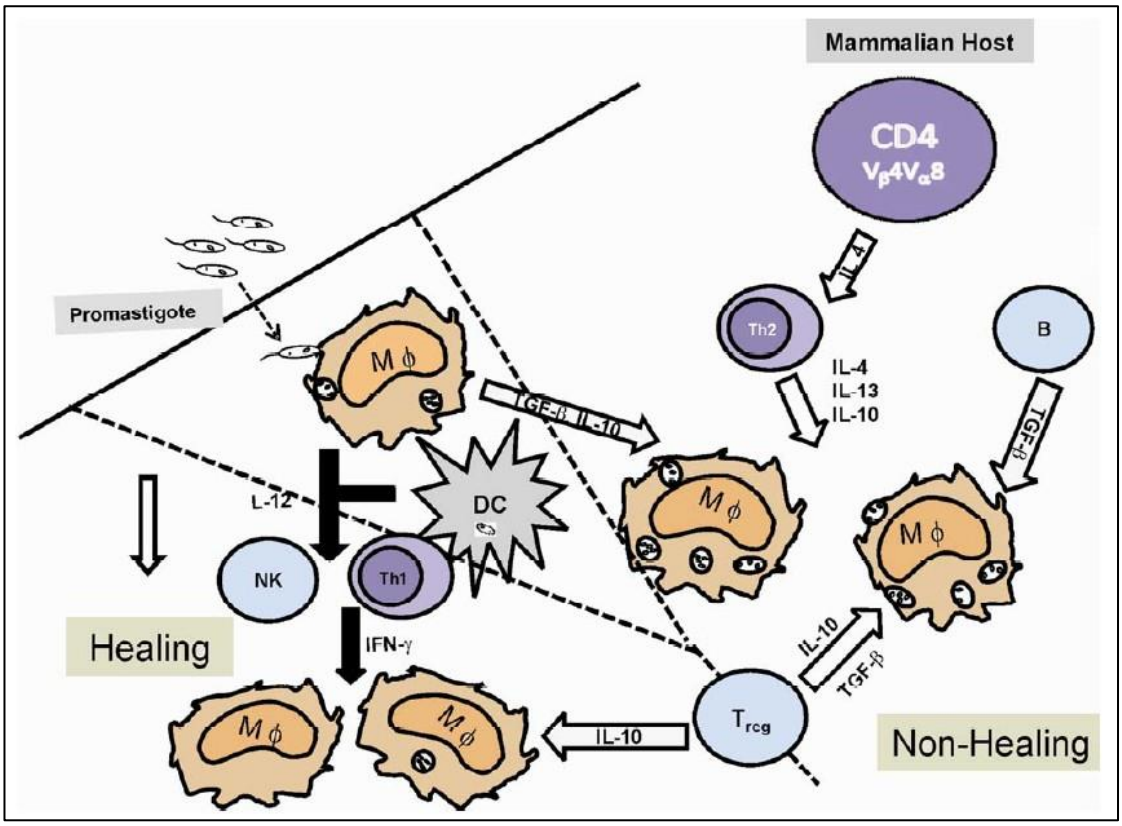


Fig.1.3 – Immune response to Leishmania parasites inside the mammalian host
Schematic representation of the differences between Th1 and Th2 cell response to *Leishmania*. If IL-12 is secreted by macrophages (MΦ) and antigen presenting dendritic cells (DC) naïve T cell will differentiate into T helper (Th) 1 cells, which secret Interferon-*γ* (IFN-*γ*), and T regulator cells (T_{reg}), which will secret IL-10, which both aid to the synthesis of nitric oxide synthase (NOS) for nitric oxide generation resulting in *Leishmania* killing. If IL-4 is secreted by antigen presenting cells instead of IL-12, then IL-4 secreting Th2 cells and transforming growth factor-*α* (TGF-*α*) secreting T_{reg} cells are generated, which promote Arginase-1 generation and parasite survival. The image was taken from Sharma & Singh (2009) (*30*).

are inoculated with infective metacyclics into the host skin (*47*). Apoptotic parasites have been proposed to be key in triggering the anti-inflammatory response, abolishing efficient killing of phagocytized viable parasites. This makes parasite apoptosis an important process during transmission (*48*). Macrophages also secrete TGF-*β* after *Leishmania* uptake (*49*). There may be slight variation and additional processes specific for some *Leishmania* species compared to others, too, that promote parasite survival (*33*).

Due to a lack of immunological studies, this dichotomy of T cell response in the mouse model has not been demonstrated as clearly in humans, in particular in the early phases of disease. Although a dichotomy was also reported in human patients, some postulated that a Th2 response precedes the Th1 response in LCL and may even be an essential prerequisite for an effective Th1 response (*31*). Another study showed that elevated levels of IFN-*γ* and TNF-*α* were not required in subclinical CL patients due to *L. (V.) braziliensis* to suppress pathology and parasite proliferation, but that in clinical CL patients persisting high levels of these cytokines increased pathology without enhancing parasite proliferation (*50*). It also appears that a balance between antagonising cytokines is more important in humans, as shown in the case of IL-10 and IFN-*γ* in VL patients (*51*). However, IFN-*γ* and TNF-*α* remain the primary cytokines involved in limiting and even clearing *Leishmania* infections. It is suggested that a balance between Th1 and Th2 cell responses is more relevant in humans (*30*). It has also been noted in VL patients that had low $CD4^+$ T cell levels (<100 cells/ml) at follow-up after treatment were indicative for disease relapse in particular in HIV-positive patients, although other factors had to participate in the control of leishmaniasis relapse too (*52*). In humans, the L-arginine depletion is more likely to be mediated through arginase-1 from active degranulating or dying neutrophils than macrophages (*53*) and immunosuppression is a common feature in non-healing CL and VL (*45*).

Leishmania derived secreted filamentous proteophosphoglycans (fPPG) and sand fly salivary proteins regurgitated together with the parasites by the vector have immunomodifying properties, too (*54*). For example, monocytes that can differentiate into mature macrophages are naturally recruited to the infection site due to the microvascular laceration caused by the sand fly bite and the presence of sand fly salivary proteins (*55*).

1.1.4. *Leishmania* / Human Immunodeficiency Virus co-infection

HIV is a retrovirus causing, in its final stages, acquired immunodeficiency syndrome (AIDS) (*56*). Globally, the number of people living with HIV is still growing (>34 million people by 2011 (*56*)) and >95% of HIV-positive cases are found in the developing world often in proximity to *Leishmania* endemic areas (*57*). While HIV in itself does not cause patient death, opportunistic infections in HIV-infected patients are often fatal. Although 'officially' still not recognised as a HIV associated opportunistic disease (*52*), *Leishmania* infection can be considered as an opportunistic disease in association with HIV (*58*). The risk of developing VL after infection with an appropriate *Leishmania spp.* is considered to be 100 – 2,320 times higher for HIV co-infected patients (*58*). In addition, HIV co-infection decreases *Leishmania* treatment success (*59*), changes *Leishmania spp.* associated pathology (*57*) and accelerates disease progression (*59*). Both pathogens target similar immune cells, resulting in synergistic detrimental effects on cellular immunity (*60*), and favour a strong Th2 response (*57*).

Leishmania/HIV co-infection has been reported in 35 *Leishmania* endemic countries to-date (Fig.1.4) (*57*). On average 2 – 9% of all reported VL cases worldwide showed HIV co-infection (*61*), while in some locations (e.g. in Ethiopia) the average lies at 15-30% (*62*). Since the introduction of highly active anti-retroviral therapy (HAART) in the European Union (EU) to combat the HIV pandemic, a clear decrease of leishmaniasis in HIV infected patients has been observed within the EU (*63*). However, it is not clear whether HAART also helps to reduce VL relapse risk (*52*). There is, however, a growing risk of co-infection due to urbanization of leishmaniasis in parallel to ruralisation of HIV causing greater regions of disease overlap (*57*). The epidemiological importance of Leishmania/HIV co-infected people is super-reservoir formation (*64*). Studies showed that sand flies are more likely to get infected with *L. (L.) infantum* by feeding on immunosuppressed hosts than on immunocompetent ones due to increased circulation of parasitized monocytes in the skin (*16*).

1.1.5. Diagnosis of Leishmaniasis

There is a range of diagnostic tools available to establish presence of *Leishmania* amastigotes in patient samples. Microscopic observation of amastigotes in macrophages within Giemsa-stained tissue samples from lesions (CL) or lymph nodes, bone marrow and spleen (VL) has been the

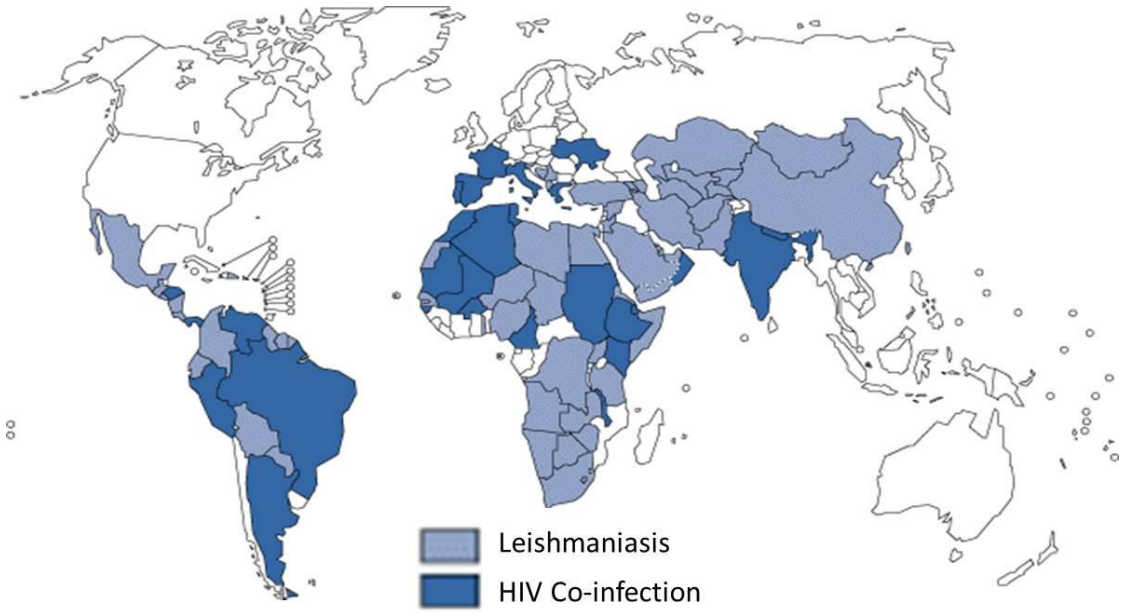


Fig.1.4 – Geographical distribution of VL and HIV co-infection

Distribution of leishmaniasis and HIV co-infection by countries with reported cases (*65*).

traditional form of diagnosis and remains broadly the gold standard (*66*). Its drawback is the low sensitivity of detection, making diagnosis of chronic lesion samples difficult, because of their lower amastigote counts compared to acute lesions (*67*). Histopathology of fixed lesion biopsies or of cultured parasites from these biopsies can be performed alongside the microscopy to increase certainty.

Current serological field tests for VL diagnosis are the easy-to-use direct agglutination and immunochromatographic dipstick tests, which have good diagnostic performance, but these tests are not applicable for CL diagnosis (*68*). The Montenegro skin test is occasionally used in scientific studies due to its high sensitivity and specificity, but is not used as a diagnostic tool in the field, because sophisticated culture facilities are required to raise the required antigens for the test (*69*). The development of molecular techniques has led to improved accuracy and sensitivity of parasite identification (*70*, *71*). Polymerase chain reaction (PCR)-based diagnosis has been shown to be highly sensitive (>90%) and specific (100%) (*72*). Conventional PCR in combination with restriction fragment length polymorphism (RFLP) analysis is probably the most commonly used PCR-based method, while real-time PCR, PCR-enzyme linked immunosorbent assay (ELISA) or PCR-oligochromatography (PCR-OC) in combination with a gold-conjugated probe dipstick can also be employed (*73*). All these techniques are, however, time consuming and require sophisticated laboratory equipment and trained personal (*72*). The loop-mediated isothermal amplification (LAMP) in combination with SYBR green is more suitable for the field (*74*). If LAMP were used as a substitute for the PCR in a thermocycler, it could be used in combination with the other PCR based methods as a good field method for diagnosis (*72*). PCR-based approaches can also be used to assess cure from VL, but requires modifications for CL, where >80% of scarred lesions test positive even 8 years after their clinical cure (*75*, *76*).

Medically differential diagnosis is essential since many other diseases often occur in the same endemic areas that show similar pathology spectra (e.g. skin cancer, tuberculosis, leprosy for CL and malaria, schistosomiasis for VL). Also, ~22 different human-infective *Leishmania spp.* with varying pathology complicates medical diagnosis. Identification of infecting *Leishmania spp.* is crucial to anticipate likely disease progression for best treatment choice, because therapeutic responses may be *Leishmania* species-specific (*72*).

1.1.6. Treatment

Trivalent antimonials were originally introduced by Vianna in 1912 in Brazil for CL treatment and in Italy by Di Cristina and Caronia in 1915 for VL treatment as the first commercially available drug. They were soon substituted with safer intravenously administered pentavalent antimonials in 1922 after successful treatment of VL with Hyper-Acid Antimonyl Tartrate + Urethane (*77*). Pentavalent antimonials have been the most commonly used first line drugs ever since. Although safer than trivalent antimonials, pentavalent antimonials are still highly toxic compounds, which cause in patients a broad spectrum of side-effects during the 20 – 30 day treatment period. Pentavalent antimonial drug resistance is becoming a serious problem, too, in e.g. India and Nepal (*78*). Alternative drugs are amphotericin B (AMB), lipid formulation of amphotericin B (L-AMB), miltefosine, paromomycin and pentamidine, which may be preferred over antimonials, depending on the success-rate of the respective drug against one or the other *Leishmania* species. AMB and L-AMB are preferentially used against VL in areas of pentavalent antimonial drug resistance. L-AMB has so far shown a better safety profile than AMB and a very good success-rate (>88%) (*79*). It is recommended in *Leishmania*/HIV co-infected patients, although it did not prevent relapses (*80*). Miltefosine is a relatively safe and successful drug (64% cure rate), which has the benefit of being an oral compound (*81*). Its gastrointestinal side effects are amplified in HIV co-infected patients and its potential teratogenic effects make it unsuitable for pregnant women. The choice of drug is not only related to the parasite species and patient, but is subject to regional preference (*82*). All of these drugs, however, are toxic and have a range of adverse side effects, what makes them theoretically inappropriate for treatment. However, no alternatives are available and to-date there is still no prophylactic vaccine. Combination therapy and secondary prophylaxis have been suggested as the way forward to avoid increased *Leishmania* drug resistance, to reduce treatment time and quantities of drugs administered (*83*). Secondary prophylaxis has been shown to be the only promising tool to reduce leishmaniasis relapses, too (*52*). While drug resistance can originate within the parasite by repeated drug exposure (*84*), some studies showed that host genetic characteristics may be a source of drug resistance, too (*85*).

The search for more appropriate drugs and drug targets is on-going. There are some alternative drugs in trial, but none have demonstrated their efficacy satisfactorily yet.

1.1.7. Vaccine development

There is no human prophylactic anti-*Leishmania* vaccine available to-date. Vaccine development has proven to be difficult. There are several requirements for an anti-*Leishmania* vaccine, which includes safety, easy low-cost production preferably in endemic countries, induction of long-term T-cell responses and prophylactic and therapeutic properties ideally with cross-protection properties against CL and VL, but so far cross-protection has been highly variable (*86*).

Leishmanization, in which *Leishmania* parasite are inoculated artificially into the person's skin to cause self-healing lesions, was the original vaccination strategy in practise (*87*). Although regionally very successful, leishmanization was largely abandoned, because of high safety risks and inapplicability in HIV or immunosuppressive drug patients. First-generation vaccines were then based on attenuated parasites and parasite material, which showed increased security, but none showed sufficient efficacy for large scale vaccination programmes (*88*). Alternative approaches (second-generation vaccines) including recombinant proteins, polyproteins and DNA vaccines in liposomal formula with dendritic cell and viral delivery systems were then adopted; candidate VL vaccines tested are reviewed by Evans & Kedzierski (2012) (*89*)). The development of second-generation vaccines, however, has been hampered by the lack of adequate animal models with appropriate reagents (*90*). A limited number of recombinant *Leishmania*-proteins have been tested in the murine model against VL, among which are the *L. (L.) donovani* dp72 (*91*), A2 cysteine proteinase (*92*), *L. (L.) infantum* BCG-LCR1 protein (*93*) and a KMP-HASPB recombinant fusion protein delivered as DNA in an adenoviral delivery system (*94*). Multicomponent vaccines have shown so far the greatest efficacy against VL. To-date, however, the LEISH-F1+MPL-SE vaccine consisting of three recombinant *Leishmania* polyproteins (TSA-LmSTI1-LeIF) with the adjuvant monophosphoryl lipid and squalene in a stable emulsion (MPL-SE) is the only multicomponent human anti-*Leishmania* vaccine in clinical trial, but is not promising (*95*). This vaccine has shown promising applications as immunotherapy in combination with chemotherapy in CL patients (*96*) and as a veterinary vaccine for dogs (*97*). For reservoir control there are currently two veterinary vaccines commercially available against canine visceral leishmaniasis, Leishmune® and Leishtec® (*98*).

The best studied DNA vaccine candidate is the *Leishmania* homologue of

receptors for activated C kinase (LACK) (*99*), which induced high level production of IFN-ɣ and IL-10 (*100*). The vaccine, however, showed significant inconsistencies in its capacity to protect against *Leishmania* (*101*). LeishDNAVAX is another DNA vaccine in preclinical trials with different delivery systems (*102*).

The inclusion of sand fly saliva components into second-generation vaccine formulations has been considered, too, due to their immunomodulating properties (*103*). Experiments have shown that pre-exposure to saliva or specific salvia components by injection or by repeated bites from uninfected sand flies conferred some protection against subsequent challenges with *Leishmania* by bites of infected sand flies or needle challenge (*104*, *105*), although co-infection of parasites with sand fly saliva in naïve mice enhances *Leishmania* infection (*106*).

With the increasing availability of *Leishmania* genome sequences (>80 to-date) (*107*–*109*) reverse vaccinology becomes possible (*110*) increasing the identification potential in combination with refined algorithms for new *Leishmania*-specific antigens for vaccines. This may accelerate the discovery of an effective anti-*Leishmania* vaccine, in the longer term.

1.2. *Leishmania* parasites

Lieutenant-General Sir William Leishman, after whom the genus was named, was one of the first to isolate *Leishmania* parasites in kala-azar patients in India in 1900. The *Leishmania* genus, together with 8 other genera, belongs to the family Trypanosomatidae (Fig.1.5) and was sub-divided by Saf'janova (1982) into the sub-genera *L. (Sauroleishmania)* and *L. (Leishmania)* (*111*); the latter being further divided by Lainson & Shaw (1987) into the *L. (Leishmania)* and *L. (Viannia)* sub-genera (Fig.1.5) (*112*). These two subgenera were principally distinguished by parasite localization in the vector's intestine (*113*); a distinction that was later supported by DNA sequence phylogenetic analysis and isoenzyme profiling (*114*, *115*). While *L. (Viannia) spp.* migrate first to the hindgut (HG) before developing towards the stomodeal valve (SV) (peripylarian development), *L. (Leishmania) spp.* have lost the HG colonization and colonize the abdominal midgut (AMG) of their vectors and begin their development from there towards the SV (suprapylarian development) (Fig.1.6) (*116*, *117*). It has been proposed that the peripylarian development is more primitive than the suprapylarian development and evolved from the hypopylarian development still observed in

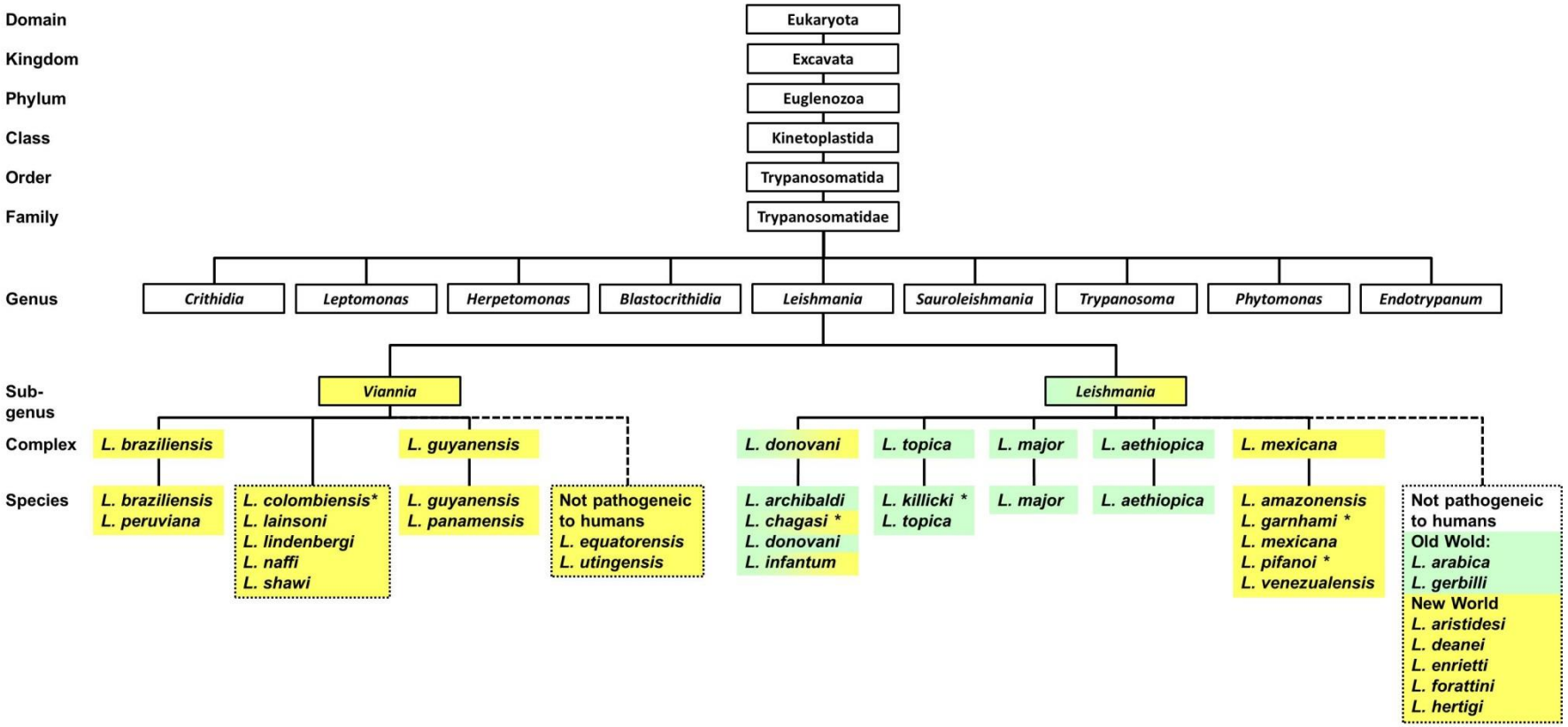


Fig.1.5 – Taxonomic tree of the mammalian infective Leishmania spp.

The tree was adapted from various sources (*36*, *118–120*). There is an on-going polemic to the exact complex clustering due to different technical approaches. For instance, Asota *et al.* (2009, 121, 352) clustered *L. (L.) tropica, L. (L.) killicki, L. (L.) major* and *L. (L.) aethiopica* into the same complex and merged the *L. (V.) braziliensis* and *L. (V.) guyanensis* complexes by cytochrome b gene sequencing. The five boxed *Viannia spp.* are still unassigned to a complex. The status and/or taxonomic position of species marked with an asterisk (*) are still under discussion. *L. (L.) chagasi* was accepted to be synonymous with *L. (L.) infantum*. *Leishmania spp.* found in the New World are in yellow; those found the Old World in green.

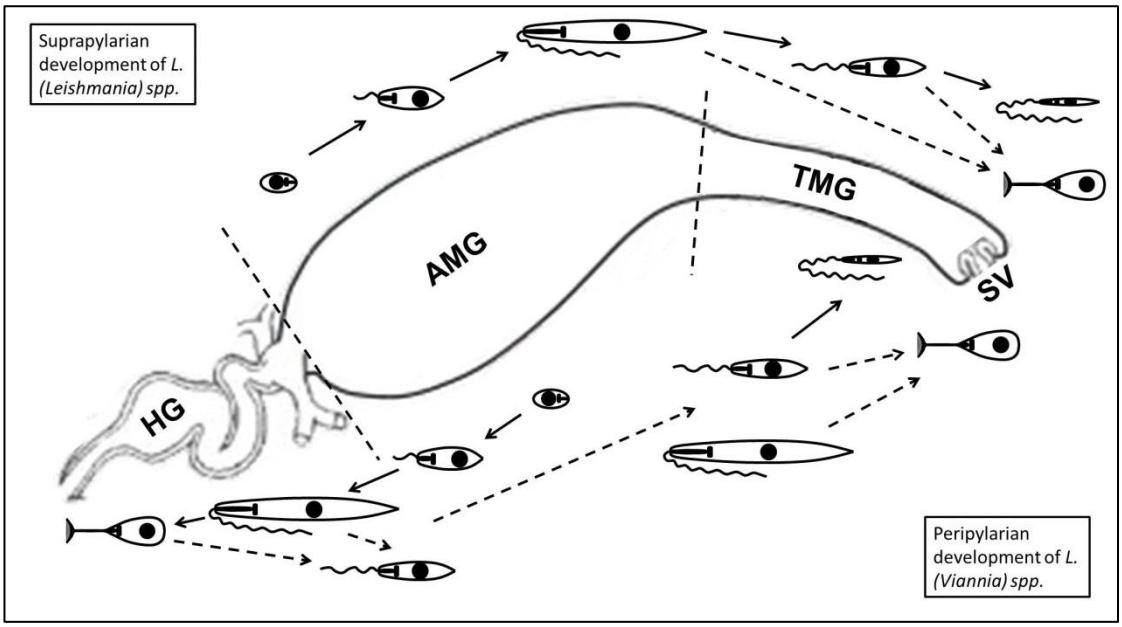


Fig.1.6 – Differences in *L. (Leishmania) spp.* and *L. (Viannia) spp.* development in the sand fly

At the top is a schematic of the suprapylarian development of *L. (Leishmania) spp.* starting within the blood meal in the abdominal midgut (AMG) and migrating forward towards the stomodeal valve (SV) as they differentiate. Below is the schematic of the peripylarian development of *L. (Viannia) spp.*, which is proposed to be evolutionarily more primitive than the suprapylarian development. Here the parasites move first into the hindgut (HG) before they migrate forward into the thoracic midgut (TMG) and to the SV. The dotted arrows indicated differentiation steps that are still being debated.

some *Sauroleishmania* species (e.g. *L. (S.) ceramodactyli* (Adler & Theodor, 1928) and *L. (S.) agamae* (David, 1929)) (*120*). It has been proposed that the genus *Leishmania* originated from monogenetic intestinal flagellates of invertebrates (*112*) and that phlebotomine sand flies were the primitive host for the *Leishmania* ancestor (*121*).

There are currently 31 species in the *Leishmania* genus (20 in the *Leishmania* subgenus & 11 in the *Viannia* subgenus) spread in several species complexes (*36*), but some species and complexes are still debated (Fig.1.5). *Leishmania* classification and species identification is difficult due to the extreme homogeneity of genomes (*109*), which is greater than within the closely related species of *Trypanosoma cruzi* (*122*), and the difficulty of morphological distinction between genus members. The origin of the *Leishmania* genus is also still a matter of debate. There are several hypothesis including Neotropic (*123*, *124*), Palaearctic (*125*, *126*) and separate origins of the *Leishmania* and *Viannia* subgenera (*127*).

1.2.1. New World Model of *Leishmania* Origin

In this model, it has been proposed that the origin of the *Leishmania* genus was in the Neotropics in the Palaeocene or Eocene 36 – 46 Million years ago (MYA) (Fig.1.7), because of the greater genetic diversity of *Leishmania* species in the New World (*123*, *124*) and the retention of the more primitive peripylarian development of *L. (Viannia) spp.* (*120*). *Leishmania* may have spread via island-hopping into the Nearctic as demonstrated by the spread of *L. (L.) mexicana* through the Caribbean (*128*) and some reservoir host then carried *Leishmania* via the Nearctic and Bering land bridge into Asia, before it became too cold for sand flies in the high north in the late Miocene (*129*). According to the New World model, the *Leishmania* and *Viannia* subgenera split in the early Miocene probably in the Nearctic, while *Leishmania* and *Sauroleishmania* did not split until the second half of the Miocene, probably as a consequence of the adaptation of *Sauroleishmania* to lizards and Old World sand flies from the genus *Sergentomyia* (*124*). The *Leishmania* subgenus diversified in Central and/or Southeast Asia during the Miocene (24-14 MYA) and spread to Europe and Africa (*114*, *124*). *L. (L.) infantum* was eventually reintroduced into the New World from Europe in historic times (~500 years ago) (*130*). This model suffers due to inconsistencies with the current sand fly classifications, which are postulated to have originated in the Old World (*131*).

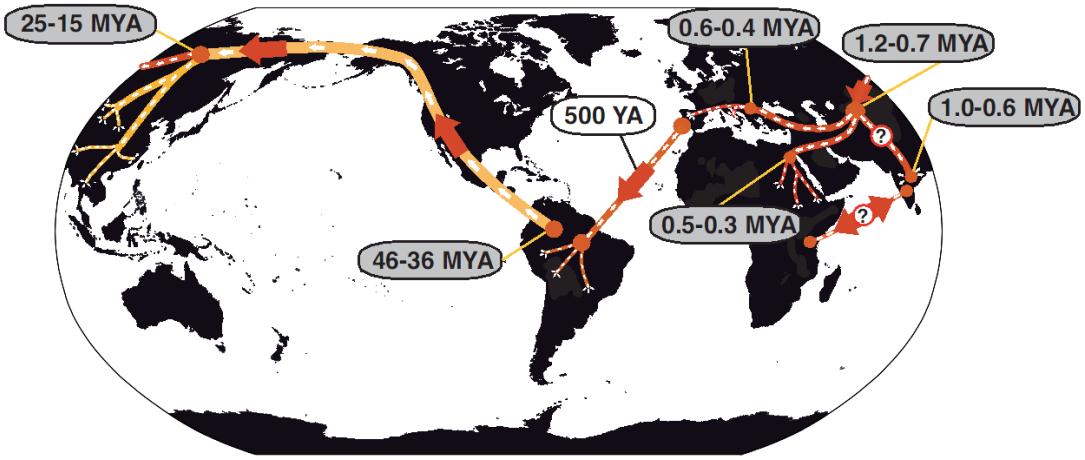


Fig.1.7 – Representation of evolutionary spread of *Leishmania* across the globe
This figure has been taken from Lukes *et al.* (2007) (*123*). It shows chronologically the proposed spread of *Leishmania spp.* across the globe from the Neotropic via the Nearctic and the Bering land bridge into Asia and from there into Europe and Africa until the re-introduction of *L. (L.) infantum* into the New World in historic times. The time points mark particular diversification and spreading events.

1.2.2. Old World model of *Leishmania* Origin

This model proposes an Old World origin of *Leishmania*, because of the proposed Palaearctic origin of the Phlebotomine sand flies in the Oligocene (potentially, in today's Lebanon) based on scares fossil records (*131*). Although *Leishmania spp.* may have vector species of not closely related sand fly species from *Phlebotomus* and/or *Lutzomyia* subgenera, *Leishmania* has been proposed to have co-evolved more closely with their sand fly vectors than with their vertebrate hosts (*132*). Therefore, it has been proposed that the parasites could only have originated and spread together with sand flies (*126*, *130*). It is important though to distinguish co-evolution from co-adaptation, which is difficult without well-supported vector/parasite phylogenies (*133*). The Palaearctic origin of murid rodents in the Oligocene (*134*) – the most important group of zoonotic *Leishmania* reservoirs in the world today – also supports the Palearctic origins of *Leishmania*. In this model *Leishmania* spread with its vectors and vertebrate hosts across Africa, the Mediterranean Europe and Asia. *Sauroleishmania* split first from the *Leishmania* genus due to the adaptation to lizards and stayed in Asia (*111*), while *Leishmania spp.* crossed the Bering land bridge into the Nearctic. Once the Panamanian land bridge was established *Leishmania* spread into the Neotropic, where *L. (Viannia) spp.* split from the *Leishmania* genus and rapidly diversified due to the climate changes and population isolation events (*125*, *126*). Alternatively, *Leishmania* may have entered the Neotropic via island-hopping (*135*), allowing a longer period for diversification of the *Leishmania* genus in the Neotropic (*125*). In this model, it has been proposed that the peripylarian development of *L. (Viannia) spp.* is not more primitive than the suprapylarian one of *L. (Leishmania) spp.* (*136*).

1.2.3. Separate origins of L. (Leishmania) spp. and L. (Viannia) spp.

In this model, it has been proposed that the sand fly genera *Phlebotomus* and *Lutzomyia* may have evolved in isolation ~120 million years ago during the Cretaceous period, while adapting to feeding on the spreading mammals from the original feeding on lizard (*131*). If the *Leishmania* and *Viannia* subgenus separated as early as ~90 million years ago (*137*), then *L. (Viannia) spp.* got separated from the *L. (Leishmania) spp.* with their *Lutzomyia* and *Phlebotomus* vectors, respectively, when Gondwana broke up (*127*). The ancestor of the *L. (L.) mexicana* complex could have entered the Nearctic via the Bering land bridge after the split with *Sauroleishmania*, but prior to the Panamanian land bridge establishment (*112*), explaining the global spread of

L. (Leishmania) spp. compared to restriction *L. (Viannia)* species to South and Central America. By correlating the origins of the sand fly genera and *Leishmania* subgenera, this model supports the proposed African origin of all Old World *Leishmania* species, while permitting the early development of *L. (Viannia) spp.* in the Neotropics (*127*).

1.2.4. *Leishmania* and its reservoirs

Leishmania has a broad reservoir host diversity reaching from insects over mammals to even plants depending on the *Leishmania* species. Reservoir hosts are essential for the survival of *Leishmania*, because they represent the natural long-term source of continuous infection of sand fly vectors, which can spread the disease to new reservoir and incidental hosts. A reservoir of infection is defined as “an ecological system in which an infectious agent survives persistently” (*138*). A good reservoir host usually bears some or all of the following characteristics: it must allow for parasite persistence, is abundant, social, long-lived, develops no acute disease, the parasites are present in either the skin and/or the blood circulation, where a vector can pick them up. As a general rule, it is not in the interest of a parasite to cause harm to its host and, therefore, these natural reservoirs remain generally asymptomatic to *Leishmania* infection. Natural mammalian parasite reservoirs are species from groups like rodents, canids, edentates, marsupials, procyonids, ungulates and even some primates among others (Appendix 1) (*139*). With the exceptions of *L. (L.) donovani* and *L. (L.) tropica* in some parts of the world, humans are incidental hosts of infection and acquisition of the parasites frequently happens when man invades endemic areas (*139*).

1.2.5. Leishmania lifecycle

All human infective *Leishmania spp.* are obligate digenetic kinetoplastid protozoa, which live as extracellular flagellate forms (promastigotes) in sand fly vectors and as an intracellular aflagellate forms (amastigotes) in mammalian macrophage phagosomes (Fig.1.8) (*140*). Intracellular amastigotes are considered to be the dominant morphological form in a mammalian host, although amastigote morphology may differ between *Leishmania spp.* (*141*) and other morphological amastigote forms may exist (*142*). Promastigote morphology, however, is more complex and several distinct morphological forms are distinguished in shape and order of appearance: procyclic promastigotes (short, ovoid, slightly motile – cell body width ≥ 4 µm, cell body length < 7,5 µm), nectomonad promastigotes (long,

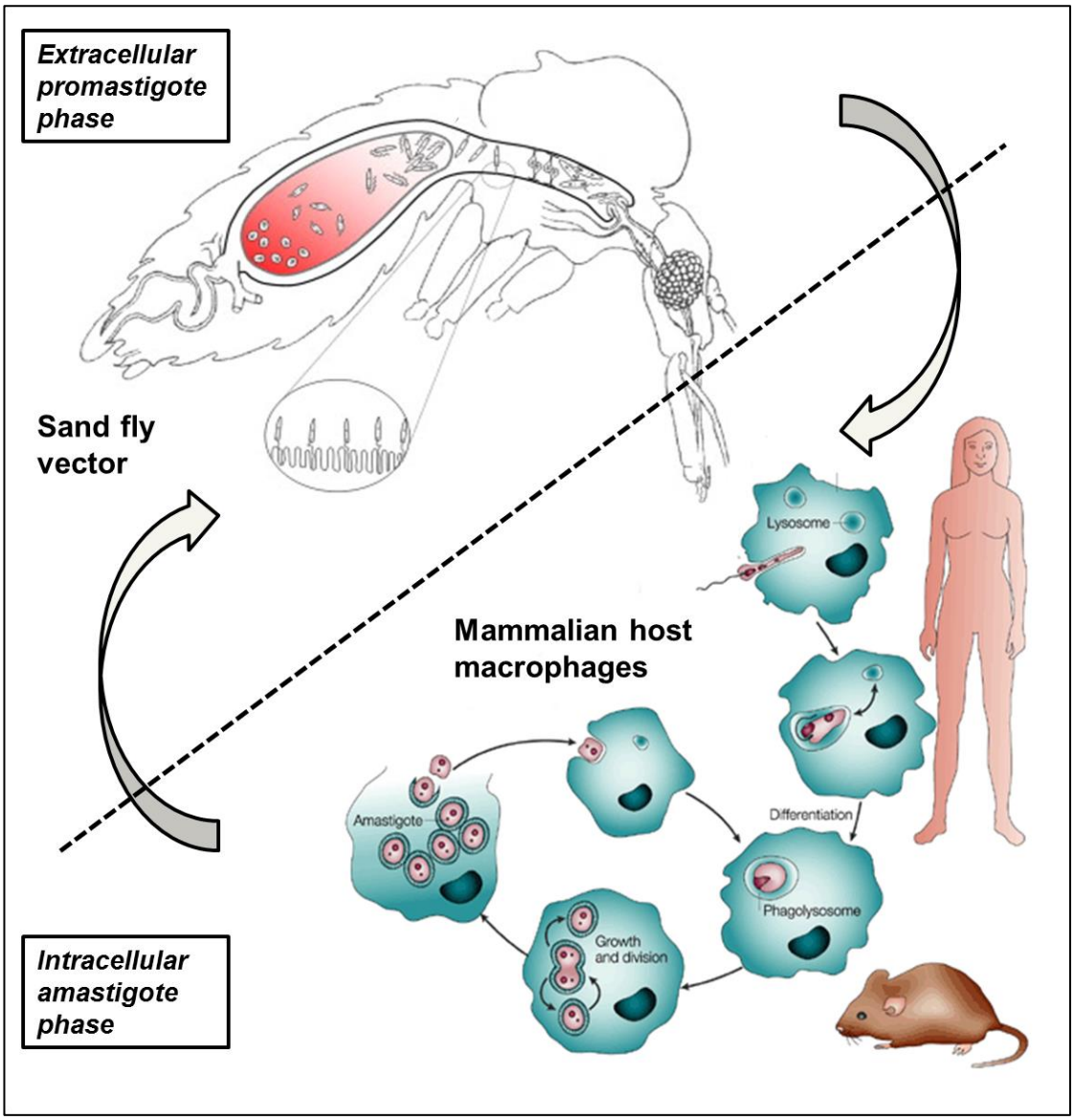


Fig.1.8 – Mammalian infective *Leishmania* digenetic lifecycle

Mammalian infective parasites from the genus *Leishmania* exist as two obligate life-cycle stages: as intracellular aflagellate amastigotes in the phagolysosomes of host macrophages and as extracellular flagellate promastigotes in the midgut of sand fly vectors of the family *Phlebotominae*. The image was adapted from Sacks & Noben-Trauth (2002) and Kamhawi (2006) (*143*, *144*).

slender, highly motile – cell body length > 14 µm), leptomonad promastigotes (short, broad formed, proliferating), metacyclic promastigotes (short, slender, highly motile, non-dividing, – cell body width < 4 µm, cell body length 7 µm < _ < 14 µm, flagellum length ≥2x body length) and haptomonad promastigotes (flaccid-looking, expanded flagellar tip, immobile) (Fig.1.9B) (*140*, *145*, *146*). The timing of appearance of the different morphological forms varies between *Leishmania spp. in vivo*, but the order is always the same. In culture, the same patterns are observed on inoculation with amastigotes, but parasite differentiation becomes rapidly desynchronised on subpassaging (*145*). The differentiation of promastigotes from procyclics to metacyclics and the accompanying migration towards the SV is termed metacyclogenesis.

1.2.5.1. Sand fly stage: Metacyclogenesis of *L. (Leishmania) spp.*

The timing of suprapylarian metacyclogenesis of *L. (Leishmania) spp.* presented here is sourced from Sacks & Kamhawi (2001) and Kamhawi (2006) (Fig.1.6 & 1.10) (*140*, *144*). Parasite differentiation steps are triggered in response to one or multiple of several described microenvironmental changes, such as changes in temperature, tetrahydrobiopterin, absence of haemoglobin or oxygen, exposure to sand fly saliva and decrease in pH (*147–150*).

Leishmania amastigotes are taken up in a blood meal by the sand fly vector. 12-18 h post blood meal (PBM), amastigotes transform within the peritrophic matrix (PM; described in more detail in 1.4.1.3) encapsulated blood meal into proliferative flagellate procyclics bearing a dense LPG coat, which confers resistance to midgut conditions. Temperature drops from ~37 ºC to ambient temperature and pH increases from pH ~5 to pH ~8 have been shown to trigger this transformation (*150*), although the underlying molecular mechanisms remain unknown. ~50% of amastigotes may be destroyed during procyclic generation and procyclics will proliferate intensely in the 24-48 h following their appearance to increase their numbers (*151*). The single flagellum extending from the flagellar pocket permits locomotion within the vector's gut lumen (and in the case of nectomonads and leptomonads attach to the vector's intestinal epithelia) (Fig.1.9A.1) (*152*). Proliferating procyclics form rosette shapes and accumulate at the anterior end of the PM. Some procyclics persist even after blood meal defecation and for the rest of the sand fly's life span, although it is not clear how they avoid expulsion since they are

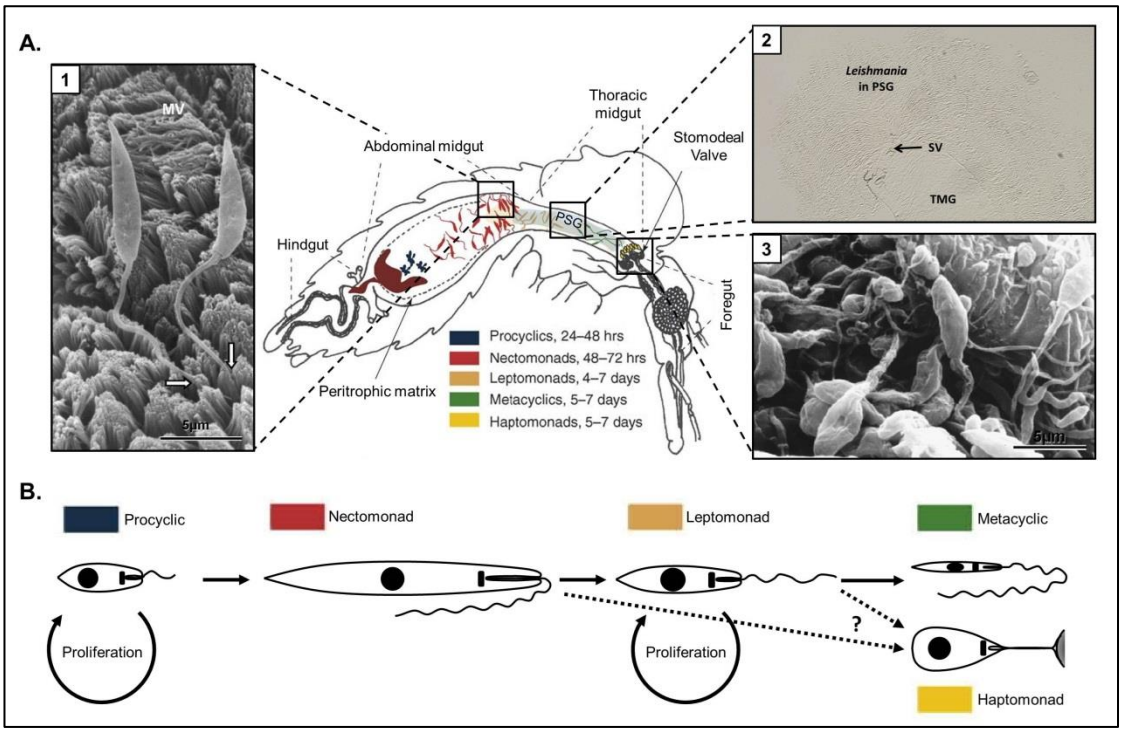


Fig.1.9 – Schematic representation of *Leishmania* metacyclogenesis

A. Parasites from the *L. (Leishmania)* sub-genus initially colonize the midgut from where they migrate forwards, while undergoing a process correlating with differentiation. 1) Midgut epithelium attachment of nectomonads by insertion of their flagellum in between the epithelial microvilli. 2) The promastigote secretory gel (PSG) plug was proposed to be essential for parasite transmission, blocking the midgut and forcing the sand fly to regurgitate it with infective metacyclic promastigotes immediately prior to blood feeding. 3) The stomodeal valve (SV) is a one way valve that prevents midgut contents from spilling back into the foregut. Haptomonads attach to the SV lining and disintegrate it by secreting chitinases. Adapted from Kamhawi (2006) and Warburg (2008) (*144*, *152*)**.**

B. Schematic representation of *L. (Leishmania) spp.* differentiation in the sand fly midgut. Two proliferative and three non-proliferative stages have been described. Although the succession of morphological forms from procyclics to nectomonads then leptomonads to metacyclics is generally accepted, it is still no clear whether haptomonads differentiate from nectomonads or leptomonads, but due to their size, shape and timing of appearance it is more likely that they also differentiate from leptomonads like metacyclics. Adapted from Gossage *et al.* (2003), Bates & Rogers (2004) and Bates (2007) (*145*, *150*, *153*)**.**

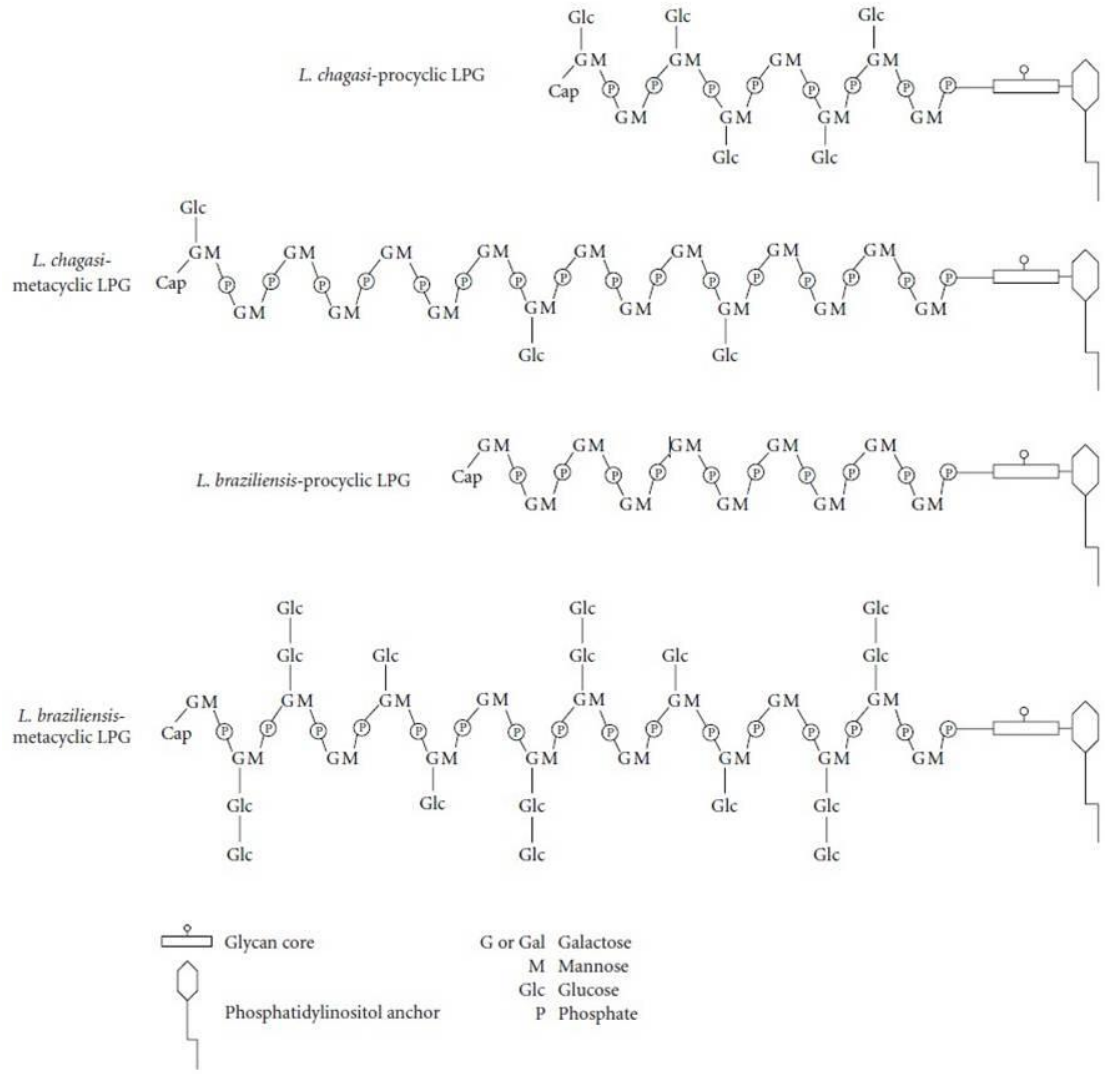


Fig.1.10 – Difference in LPG side chain composition between *L. (L.) infantum (=chagasi)* and *L. (V.) braziliensis*

While the length of metacyclic LPG increases in both species the short procyclic LPG chain of *L. (L.) infantum (=chagasi)* has many *β*(1,3)glucose residues, which are eliminated in the metacyclic LPG, the LPG of *L. (V.) braziliensis* does the opposite by increasing the number of *β*(1,3)glucose residues in the metacyclic LPG. Adapted from Soares *et al.* (2010) (*154*)**.**

outcompeted for epithelial attachment by nectomonads (*155*).

In culture, development of nectomonads from procyclics occurs principally in late exponential / early stationary phase in response to starvation due to nutrient depletion (*145*), which occurs after ~3 days PBM within the PM prior the blood meal defecation. The PM ruptures by about day 3 PBM – probably due to sand fly derived chitinases that aid excretion – and nectomonads enter the ectoperitrophic space to attach to the intestinal epithelium. This attachment has been proposed to be essential for infection persistence (*156*). The speed of escape correlates with the height of initial infection levels (*157*). Nectomonads are also proposed to migrate into the thoracic midgut (TMG) and towards the SV, but the attachment/detachment kinetics of nectomonads, which are proposed to depend on changes in the LPG structure (*158*, *159*), remain unknown. However, some nectomonads have been observed to remain unattached within the midgut lumen (*117*). It was, therefore, postulated that epithelium attachment may serve environment preadaptation for parasite survival and continuation of metacyclogenesis by, for instance, abolishing peristalsis by myoinhibitory neuropeptide secretion (*160*). During nectomonad migration, by day 4 – 5 PBM, nectomonads begin to transform into leptomonads (formerly short promastigotes or haptomonads, but renamed as distinct forms (*145*, *161*)), the second replicative promastigote form, which increase parasite numbers in the TMG/cardia generating massive infections by day 5 – 7 PBM (*145*). Leptomonads have been shown to secrete the promastigote secretory gel (PSG) (Fig.1.9A.2) (*161*). Flaccid looking true haptomonads, a terminal differentiation stage, are generated from leptomonads by day 5 – 7 PBM. They have a modified flagellum with which they can attach to the cuticle of the SV via hemidesmosomes-like structures (Fig.1.9A.3) (*155*). Haptomonads were proposed to be involved in the degeneration of the SV by secretion of parasite chitinases allowing infective metacyclic promastigotes to migrate into the sand fly foregut (FG) (*162*). Haptomonads never make up more than 10% of the total parasite load of an infected gut and may also be found attached to the HG and FG depending on infecting *Leishmania spp.* (*145*).

Mammalian infective metacyclics are also generated from leptomonads by day 5 – 7 PBM (*145*, *161*). *In vitro,* this differentiation step can be induced by lowering the pH, anaerobic condition and tetrahydrobiopterin depletion

(*148*, *149*). Metacyclic LPG is commonly longer than the procyclic form and varies in side-chains and cap composition, providing protection against the mammalian complement system, facilitating release from the midgut epithelium and preventing attachment to the PSG. The peak of metacyclic generation usually coincides with the sand fly's search for another blood meal allowing for optimal parasite transmission into a mammalian host (*163*). Therefore, metacyclogenesis is an essential prerequisite for parasite transmission to a mammalian host *in vivo*. By day 7 – 10 PBM, >60% of resident parasites are found near the SV and in the TMG (*161*).

1.2.5.2. Sand fly stage: Metacyclogenesis of *L. (Viannia) spp.*

A lot less is known about the peripylarian metacyclogenesis of *L. (Viannia) spp.* (Fig.1.6). They have the same morphological life-cycle stages within their sand fly vectors (*Lutzomyia spp.*) as *L. (Leishmania) spp.*, but the direction of their peripylarian development is different. Amastigotes are taken up with the blood meal, differentiate within the PM enclosed blood meal into procyclics and then proliferate at the anterior end of the PM just like *L. (Leishmania) spp.*, but once the nectomonads escape from the PM, the majority migrates into the pyloric region of the HG, where parasites attach to the cuticle as rarely dividing haptomonad-like forms just like at the SV (*116*, *164*). This may be integral to the establishment of a persistent infection, because there is little evidence that *L. (Viannia)* nectomonads actually attach to the midgut epithelium (*165*), although their LPG has been shown to bind (*154*). After the HG phase, there is a forward movement of nectomonads towards the SV and differentiation first into leptomonads and then into haptomonads and metacyclics (*116*). However, the details of the kinetics of this forward migration have not been fully explored, although the end result of metacyclogenesis in the *Viannia* subgenus is very similar to what is observed in the *Leishmania* subgenus (*153*). The degradation of the SV and secretion of the PSG is proposed to occur for *L. (Viannia) spp.*, too, so does the forward transmission during sand fly blood feeding.

1.2.6. The *Leishmania* cell surface

The cell surface of *Leishmania* parasites changes dramatically between lifecycle stages. While the extracellular promastigote forms are covered in a dense coat of predominantly glycosylphosphatidylinositol (GPI)-anchored glycoproteins, LPG, PPG and a family of free GPIs, the intracellular amastigotes are predominantly covered in a dense coat of free GPIs, lacking

LPG (*166*, *167*). The GPI-anchored glycoproteins and LPG are associated with different roles in the parasite's life within the midgut; for instance, in the invasion of macrophages, evading complement lysis and the specific midgut attachment of promastigotes (*168*, *169*). LPG is essential for the sand fly midgut attachment in specific vector – parasite combinations, but not in permissive vector – parasite ones (*169*, *170*).

1.2.6.1. Glycosylphosphatidylinositol-anchors

GPI-anchors are an abundant mean for all eukaryotic cells to tether proteins and other macromolecules to the plasma membrane. All examined *Leishmania spp.* have the same completely conserved 1-*O*-alkyl-2-*lyso*-phosphatidyl(*my*)inositol anchor with C_{24} and C_{26} saturated aliphatic chains (*171*). In *Leishmania*, GPI-anchors are used predominantly to tether proteins and LPG to the cell surface and they represent the sole class of free glycolipids on the cell surface. By targeting the dolichol-phosphate-mannose synthase (DPMS) of *L. (L.) mexicana*, it was determined that GPI synthesis is localized to a distinct tubular subdomain, termed the DPMS tubule, which appears associated with the subpellicular microtubules, the Golgi and the mitochondrion (*172*). The synthesis is tightly regulated to the parasite's developmental stages. GPI synthesising enzymes are down-regulated and re-localized to the multivesicular tubule lysosome in promastigotes with the approach of stationary phase in culture (*173*). In culture, the onset of stationary phase coincides with the initiation of metacyclogenesis due to starvation.

1.2.6.2. Lipophosphoglycans

LPG is the largest and most abundant surface glycoconjugate of promastigotes, covering the entire cell surface and flagellum. LPG consists of the conserved GPI-anchor linked to a conserved glycan core region (Gal(*α*1,6)-Gal(*α*1,3)-Gal_f(*β*1,3)-[Glc(*α*1)-PO_4]-Man(*α*1,3)-Man(*α*1,4)-GlcN(*α*1)) to which a capped oligosaccharide (Gal(*β*1,3)-Man(*α*1)-PO_4) backbone is attached (Fig.1.10) (*174*). The backbone varies in capping structure, in Glc(*β*1,3)-side-chain content and in length between different promastigote stages and *Leishmania spp.* (*175*). The backbone Glc(*β*1,3)-side-chains show inter- and intra-species specific polymorphisms in variability and size (*154*, *159*, *174*). Modes of LPG modification over the course of metacyclogenesis differ between *Leishmania* subgenera; while *L. (Leishmania) spp.* tend to reduce the number of side-chains towards the

metacyclic stage, the *L. (Viannia) spp*. increase the number and length of metacyclic LPG Glc(*β* 1,3)-side-chains (*159*, *165*). In case of *L. (Leishmania) spp.*, the reduction of the side-chains is proposed to be responsible for the detachment of metacyclics from the midgut epithelium (*159*). The epithelium attachment is proposed to be parasite LPG – sand fly lectin dependent in specific parasite – vector combination (*169*) and it has been shown that *L. (Leishmania) spp.* nectomonads and leptomonads have higher affinity for midgut epithelia attachment than procyclics and metacyclics, a feature attributed to their LPG modification (*155*). This is not the case for *L. (Viannia) spp*., where metacyclic derived LPG was able to out-compete all promastigote forms in midgut epithelium attachment (*154*), suggesting an alternative mechanism for metacyclic midgut epithelium detachment in *L. (Viannia) spp.*, than for *L. (Leishmania)* species. LPG is important for parasite survival in the sand fly vector for all *Leishmania spp.*, conferring resistance to sand fly immune responses and proteolytic digestion (*169*), while metacyclic specific LPG also protects against complement lysis in the human host (*34*, *35*).

1.3. Sand Fly Vectors

Phlebotomine sand flies are diptera insects from the Family Psychodidae. *Ph. (Ph.) papatasi* was described by Scopoli in 1786 (originally *Bibio papatasi*) as the first of ~900 sand fly species described to-date (*132*, *176*). All sand fly species belong to one of five accepted genera: *Phlebotomus* (Loew 1845) and *Sergentomyia* (França & Parrot 1920) in the Old World and *Brumptomyia* (França & Parrot 1921) *Lutzomyia* (França 1924) and *Warileya* (Hertig 1948) in the New World (Fig.1.11). Experimental *Leishmania* transmission by sand fly bite to a new mammalian host was incontrovertibly demonstrated for the first time by Adler & Ber (1941) with *L. (L.) tropica* in *Ph. (Ph.) papatasi* (*177*) and by Swaminath, Shortt and Anderson (1942) with *L. (L.) donovani* in *Ph. (Euphlebotomus) argentipes* (*178*). Only ~70 of all described sand fly species have ever been associated with *Leishmania* transmission to mammals and all of these belong either to the genera *Phlebotomus* (Old World) or *Lutzomyia* (New World), although *L (L.) major* has been isolated from *Sergentomyia spp.* in East Africa (*179*, *180*) and Iran (*181*). These species also take blood from mammals and not only from lizards (*182*), but these sand flies do not support full *Leishmania* development (*183*). There is also evidence that the vectors of *Leishmania* among marsupials in Northern Australia are ceratopogonid midges, while sand flies are scarce or absent in these areas of natural *Leishmania* transmission (*184*). To

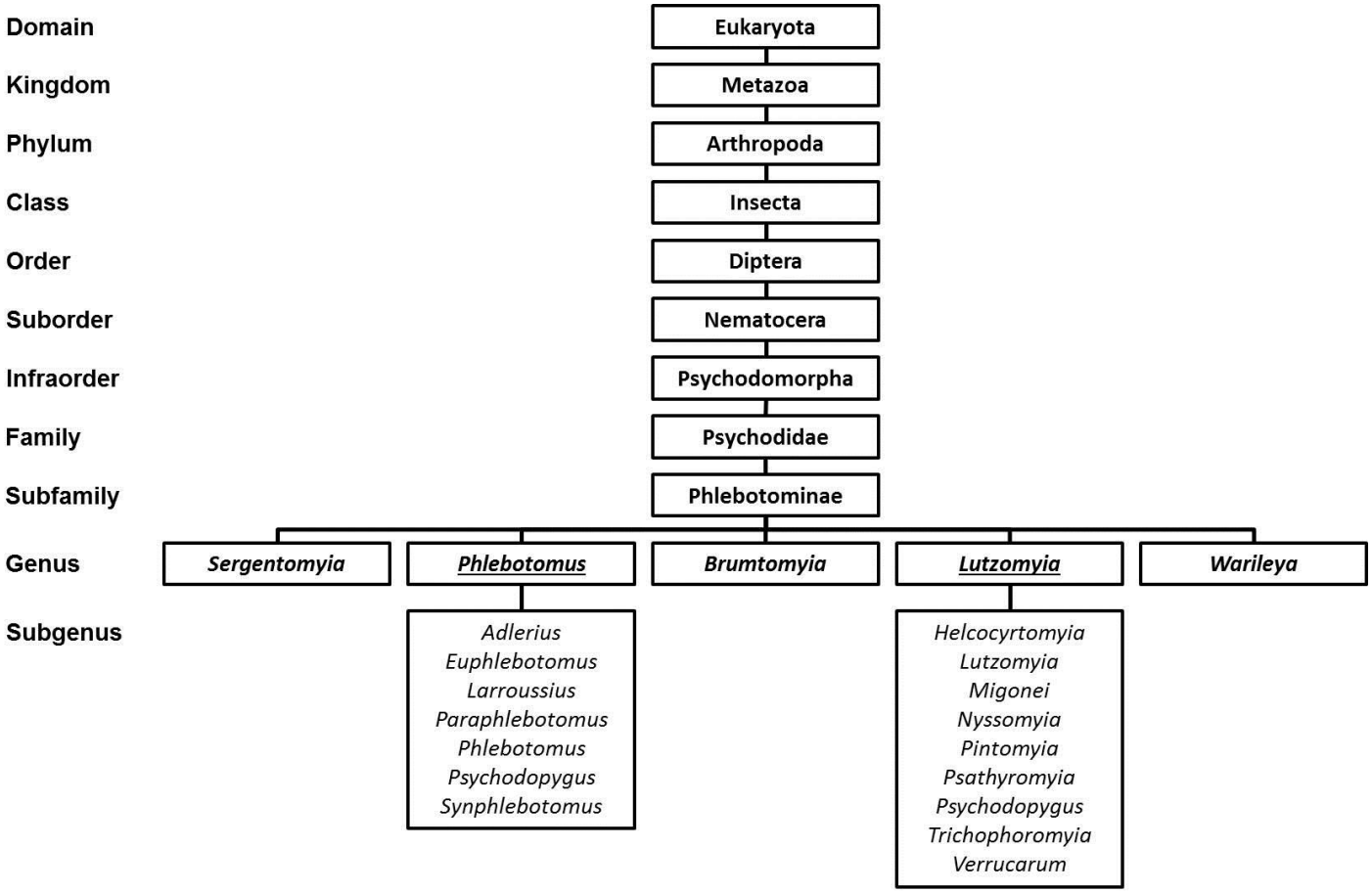


Fig.1.11 – Taxonomic tree of *Leishmania* transmitting phlebotomine sand flies

All the medically relevant sand fly subgenera are listed together underneath the respective genus. Species of only seven of the twelve *Phlebotomus* subgenera and nine of the 25 *Lutzomyia* subgenera and species groups are suspected or proven *Leishmania* vectors. The relevant sand fly vector species can be found in Table 1.1. The tree is based on NCBI's taxonomy browser and Killick-Kendrick (1999) (*185*).

date, there is convincing evidence for only ~30 sand fly species to be natural vectors for one or several human infective *Leishmania spp.* (*186*). The criteria for incriminating a natural vector are detailed in Ready (2013) (*132*), but a necessary characteristic for a vector of human-infective *Leishmania* is anthropophilicity. The fact that some *Leishmania spp.* have several different, often not closely related sand fly species as vectors demonstrates great plasticity of the parasites in their sand fly vectors and suggests that *Leishmania* may spread easily by quick adaption to a new sand fly vector (*132*), as in the case of *L. (L.) infantum* adaptation to *Lu. (Lu.) longipalpis*. This is one of the reasons why vector control is so complicated as a measure for leishmaniasis control.

Phlebotomine sand flies have a body length of about 2-3 mm and are covered in a dense coat of oily hairs and bristles (Fig.1.12). Their wings have a very characteristic V-shape posture when resting, which distinguishes them from other members of small diptera Families. Their usual time of activity is twilight and night with peak times between 3:30am – 5:30am, but they also may fly and bite during the day when disturbed from their resting places or present in dense shades or under sufficiently-day-light-reducing cloud cover. Adult sand flies are primarily phytophagous and their diet is sugar from plant sap and perhaps from aphid honeydew (*187*). Only mature female sand flies feed on blood and mammalian vector species blood-feed at least twice throughout their adult life. Females are telmophages (pool feeders), feeding from lacerations inflicted on the host's skin by inserting their saw-like mouthparts (*186*). Skin laceration has been proposed to promote up-take of amastigotes by sand flies, by releasing skin macrophages and/or freeing amastigotes from damaged macrophages into the pool (*153*).

1.3.1. Sand fly Development

Despite their importance as *Leishmania* vectors, not much is known about their natural behaviour and breeding sites of sand flies and most knowledge comes from laboratory observations. Sand flies occur in diverse environments from moist tropical rainforests to arid deserts. In nature, adult sand fly habitats do not necessarily coincide with breeding sites and it is proposed that female sand flies will find suitable breeding sites by recognition of natural attractants emitted by e.g. faeces and soil bacteria (*188*). In general, sand flies lay their eggs in organically rich moist soil protected from sunshine and rain, such as rain forest floor between tree roots and soiled animal shelters/burrows, but never in water as is typical for mosquitos. Davis (1967) published the most complete observation of the sand fly life stages (*189*). Briefly, the female sand

Fig.1.12 – Sand fly *Phlebotomus (Phlebotomus) papatasi*

The image shows *Ph. (Ph.) papatasi* blood feeding. The oily hairs on thorax, abdomen and wings are visible, which cause the sand fly's whitish appearance. The distinct V-shape of the wings is also discernible. The picture is an open source from the CDC homepage.

fly scatters about 30-70 eggs around relatively dry or moist soil rich in organic matter, which will be the feeding base for the hatching larvae. The larvae hatch through a J-shaped fissure in the egg shell and pass through four larval instars before pupating and emerging as adult sand flies (Fig.1.13). None of these stages takes the phlebotomine sand fly to aquatic environments, which distinguishes it from other haematophagous dipterans, like mosquitos. The whole lifecycle takes about 20-40 days. Temperature and precipitation levels can vary the length of the lifecycle in particular in the larval stages.

1.3.2. Structure of the sand fly alimentary canal

Warburg (2008) analysed the sand fly alimentary canal structure (Fig.1.14) in great detail by scanning electron microscopy (*152*). It begins with the cuticle-lined FG that consists of the proboscis (mouthparts) followed by cibarium and pharynx, which both have pumping activity to regulate the liquid food flow (Fig.1.15) (*189*, *190*). The pumping activity is facilitated by the cibarial valve, which separates the cibarium from the pharynx, and the SV, which separates the pharynx from the TMG. The valves prevent the back-spill of liquid food. Just ahead of the SV is the oesophagus, where the diverticulum (crop) attaches. The crop is a storage compartment exclusively for sugar-based foods. The SV has chemosensory activity on the side of the oesophagus, where basiconic sensilla (sensory organs with conical base protruding from the cuticle in arthropods) are present (*191*). This allows the SV to direct fluids either into the crop or the gut. When the sand fly takes a sugar meal, only a small amount is initially directed into the midgut, the majority is diverted into the crop and is then released only gradually into the midgut, while blood is always directed into the midgut in its entirety and digested as a single batch (*192*). The gut that follows the SV is divided into three sections that have different embryological origins: the TMG, the AMG and the HG (Fig.1.16). TMG and AMG make up the midgut, which is lined by a single layer epithelium with densely packed microvilli, that is separated by a fine basal lamina from the hemocoel (*193*). The midgut epithelium secretes the PM and digestive enzymes and is involved in absorption and transport of nutrients (*194*). The HG is cuticle-lined, like the FG, and consists of pylorus, ileum and rectum (*152*, *195*). While nutrient absorption is generally restricted to the midgut, water and salt are re-absorbed in the HG from urine and faeces (*194*). Sand flies are small (2 – 3 mm in body length) and with that the size of a meal – blood or sugar – is usually <0.8 µl (*196*). This is an important factor in *Leishmania* infection, because it is a limiting factor in how many parasites

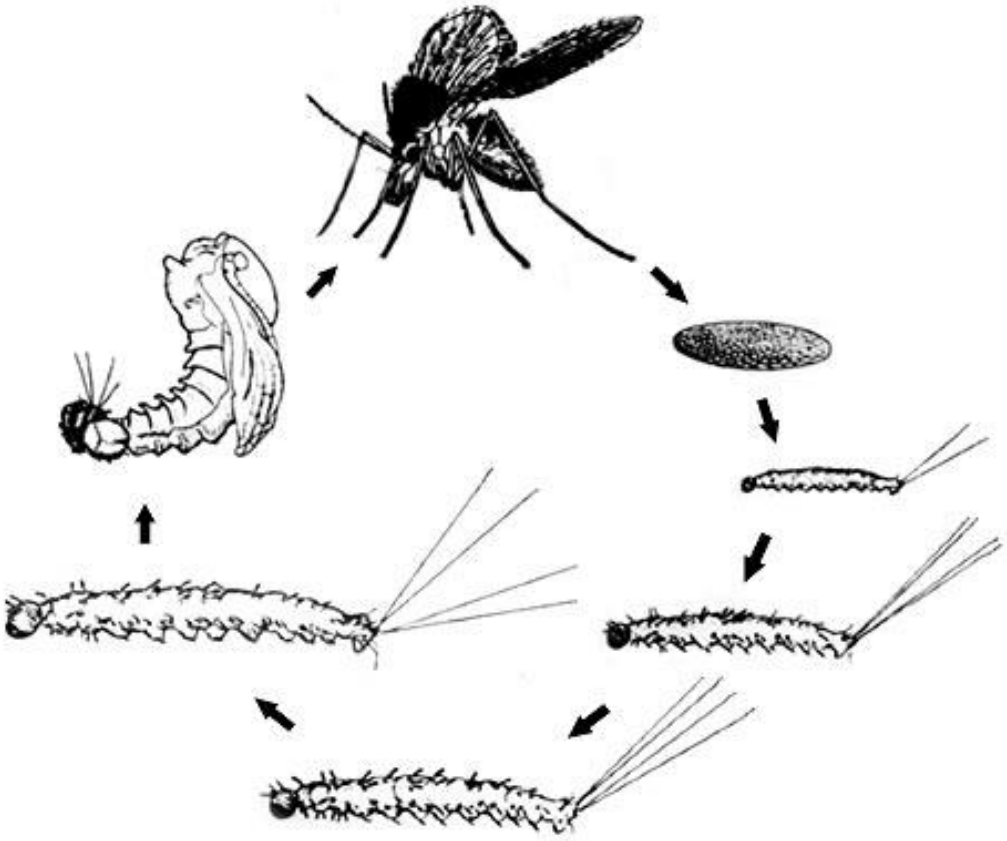


Fig.1.13 – Sand fly development from the egg to the adult insect

Phlebotomine sand flies are holometabolous insects developing from the egg through four larvae and one pupal stages to reach the adult stage (*197*).

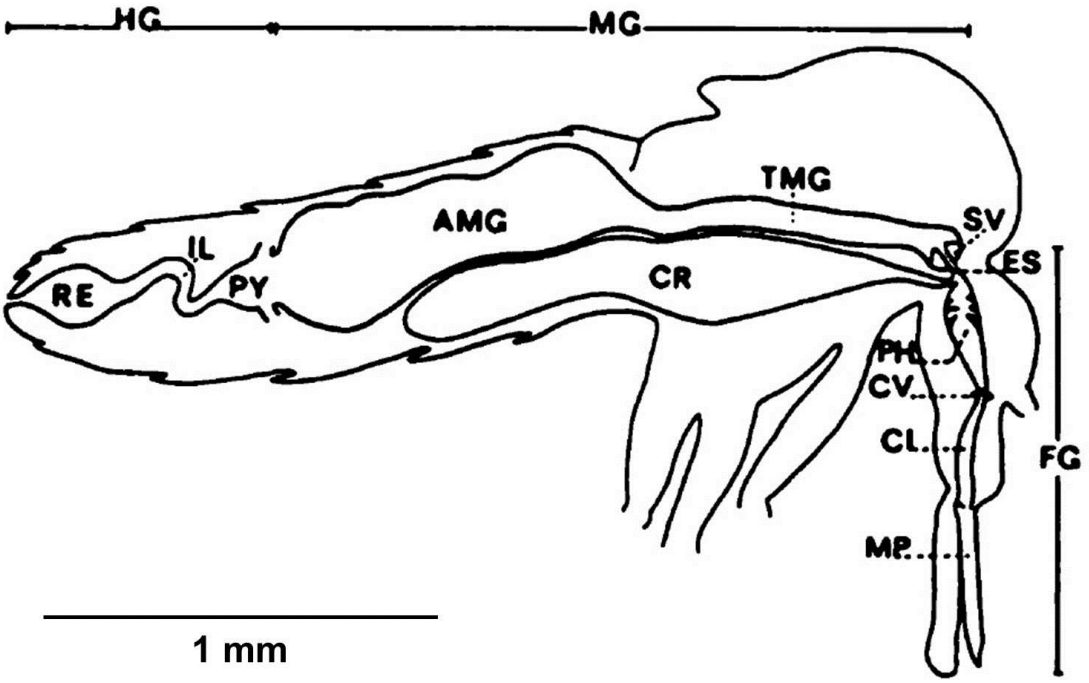


Fig.1.14 – Schematic of the *Ph. (Ph.) papatasi* alimentary canal

Foregut (FG), midgut (MG), hindgut (HG), proboscis or mouth parts (MP), cibarium (CI), cibarial valve (CV), pharynx (PH), esophagus (ES), stomodeal valve (SV), thoracic midgut (TMG), crop (CR), abdominal midgut (AMG), pylorus (PY), ileum (IL) and rectum (RE). Taken from Warburg and Schlein (1986) (*198*).

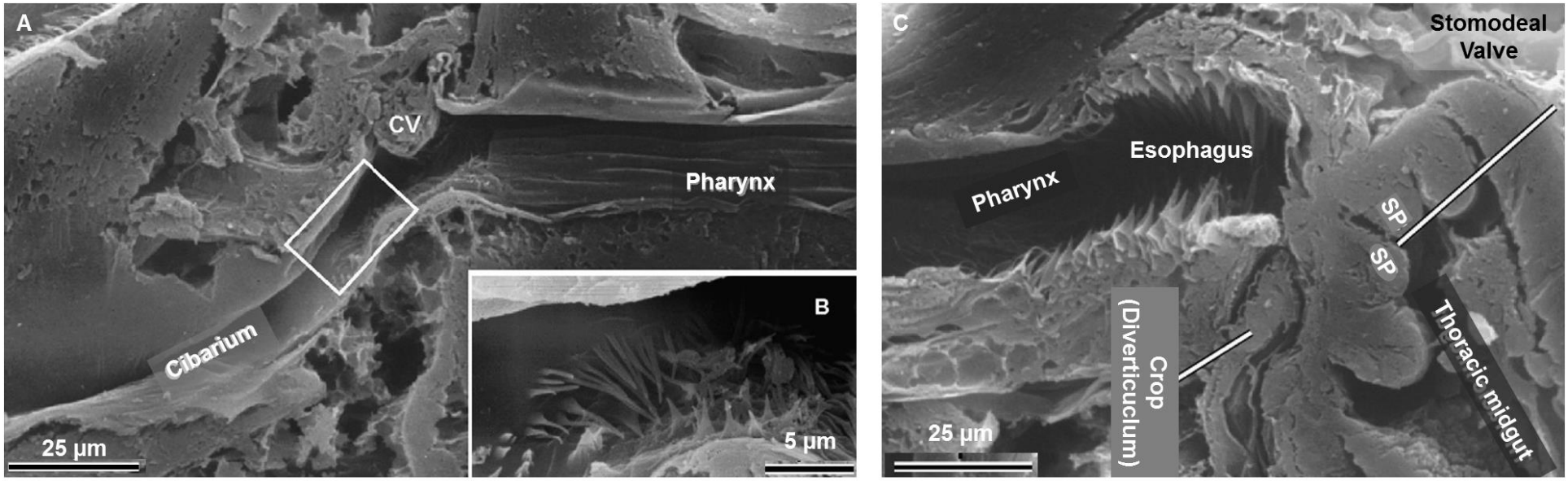


Fig.1.15 – Images of sand fly foregut

A - B) The cibarium, that follows the proboscis (mouth parts) is narrow and smooth till the posterior part ahead of the cibarial valve (CV), where several rows of thorn like shorter and flexible longer cuticular appendages are localized, whose function is elusive (see B). The anterior pharynx following the CV is narrow and lined by smooth cuticular plates with longitudinal ridges. C) Posteriorly, the pharynx opens into the esophagus which also contains rows of appendages. Here the diverticulum (crop) is attached ahead of the stomodeal valve, which regulates the flow of food by the cuticle-lined sphincter ring muscle (SP). For *Leishmania* transmission, the parasites have to travel up the FG against the direction of flow. These scanning electron microscopy images were adapted from Warburg *et al.* (2008) (*152*).

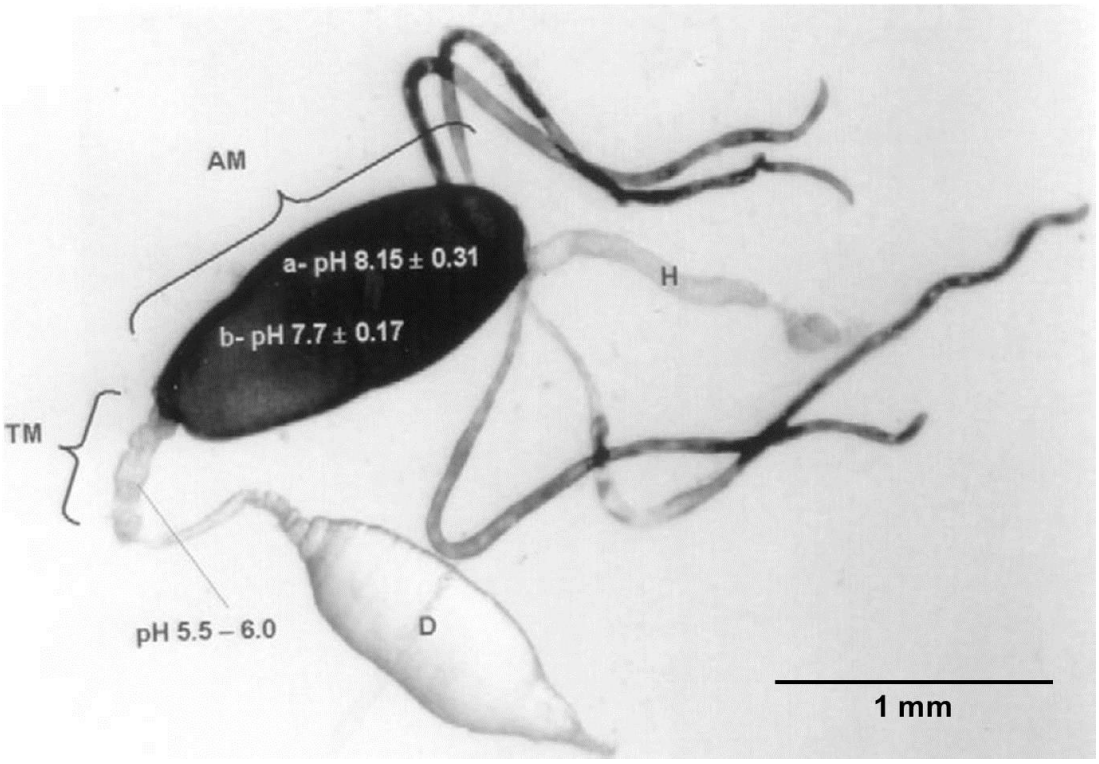


Fig.1.16 – Image of dissected blood meal containing midgut

The image shows the crop (D) filled with sucrose at the bottom. It attaches to the gut just ahead of the SV, which controls the influx into the midgut. The thoracic midgut (TM) is void of blood as can be seen in the image and its pH remains at sugar digestion levels, while the abdominal midgut (AM) is filled with blood and has an alkaline pH, which peaks after blood uptake at about pH 8 and then declines again over the course of digestion (about pH 7.7 after 32 h). The image was take from Santos *et al.* (2008) (*199*).

could potentially be taken up with a blood meal.

1.3.3. Midgut Physiology

To understand *Leishmania* metacyclogenesis, it is important to consider the physiological conditions and structure of the sand fly midgut, to which the parasites are exposed. Work done with the New World sand fly species *Lu. (Lu.) longipalpis* showed that the midgut was a highly dynamic environment (*190*). Digestion of sugars derived from plant saps is the main alimentary source for sand flies throughout their adult life stage. Since digestion is essentially an enzymatic process and enzyme activity is pH dependent, the pH is an important factor in the midgut environment and will be adapted to optimal digestive enzyme activity. Sugar digestion in the *Lu. (Lu.) longipalpis* midgut occurs at a pH of ~6, the optimal pH for the activity of the *Lu. (Lu.) longipalpis* midgut's *α*-glucosidase (pH 5.8) (*200*). This optimal *α*-glucosidase pH of 5.8 in *Lu. (Lu.) longipalpis* is not a universal constant among sand flies. Dillon and El Kordy (1997) found that the optimal pH for the *α*-glucosidase in Old World species, *Phlebotomus (Larroussius) langeroni*, was ~7.5 (*196*). This may suggest significant differences in midgut physiology between different sand fly subgenera or even species, which may aid the restriction of *Leishmania spp.* able to infect a specific sand fly species (*199*). *α*-glucosidase is a membrane bound enzyme and, interestingly, most of its activity is found in the TMG (*199*). The nutrients derived from the blood are almost entirely used for egg development, instead of sand fly alimentation. Since blood meal digestion occurs at about pH 8 and takes between 40 – 45 h in *Lu. (Lu.) longipalpis*, the female sand fly requires sugar digestion during blood meal digestion for sustenance, too. This need is resolved in the split of the midgut into two distinct environments. While the blood gets digested exclusively within the PM in the AMG, where the pH is changed to >8 within 10 min. after blood uptake, sugar digestion continues to occur in the TMG at pH 6 (Fig.1.16) (*199*). This explains why most of the *α*-glucosidase activity is found in the TMG even in unfed sand flies. Alkalization of the midgut is not unusual among insect species and, in particular, larvae of the suborder Nematocera, to which sand flies belong (*201*). It is proposed that at least two mechanisms are involved in midgut alkalization in Nematocera members. One is by CO_2 volatilization (*199*, *202*), which may be the principal mechanism *in vivo* (*203*). It is necessary for the sand fly to turn off the mechanisms that maintain the pH 6 in the midgut, otherwise CO_2 volatilization by itself would be an insufficient mechanism for alkalinisation and it was subsequently shown that any ingested

protein from blood at physiological levels could turn off the mechanism that maintains the pH 6 (*203*). A second mechanism has not been described in detail yet, but it was shown that the alkalization of the AMG also occurred significantly when blood was either depleted of CO_2 and then pH corrected or replaced completely with pH indicator dye containing solutions, which excluded CO_2 volatilization as the only alkalization process (*199*, *202*). Interestingly, lowering the pH of cultures was shown to be a trigger in metacyclogenesis (*149*, *204*).

1.4. Sand fly Vector and *Leishmania* Parasite

The level of *Leishmania* infection within a sand fly population rarely exceeds 1.5% of the total population in endemic areas, although infection rates of humans and dogs in the same area are ~18.9% and ~46.7%, respectively (*205*). Apart from mammalian host availability, aspects of the sand fly vector may contribute to limiting *Leishmania* infection in their populations. *Leishmania* parasites encounter various obstacles within the sand fly midgut, which they need to overcome to establish sand fly infection, complete metacyclogenesis and, ultimately, achieve transmission to a new mammalian host. These challenges require extensive and complex interaction at the molecular level between parasites and their sand fly vectors (reviewed in (*206*)). If *Leishmania* infection is established successfully, however, adult sand flies stay infected for the rest of their adult life.

1.4.1. Manipulating host enzyme expression patterns

The expression pattern of digestive midgut enzymes is dynamic and changes accordingly to whether a sugar meal or a blood meal is digested (Fig.1.17). Although no parasite mechanisms manipulating sand fly enzyme expression patterns have been discrebed in detail yet, distinct changes to the sand fly enzyme expression levels have been observed in the presence of *Leishmania* compared to their absence. For example, downregulation of digestive enzymes protects the parasite against proteolytic lysis, an additional protection to their LPG coat. It has been shown that *L. (L.) mexicana* survival is enhanced in *Lu. (Lu.) longipalpis* by experimental RNAi suppression of the major blood-meal induced trypsin (*207*). Active suppression or delay of the sand fly midgut trypsin peak activity and modulation of trypsin-like transcript levels was also shown for *L. (L.) major* and *L. (L.) infantum* in *Ph. (Ph.) papatais* and *Lu. (Lu.) longipalpis*, respecitvely (*208*, *209*). Peritrophins are another group of proteins to be markedly down-regulated in the presence of *Leishmania* parasites. Peritrophins bind chitin fibers via several chitin-binding

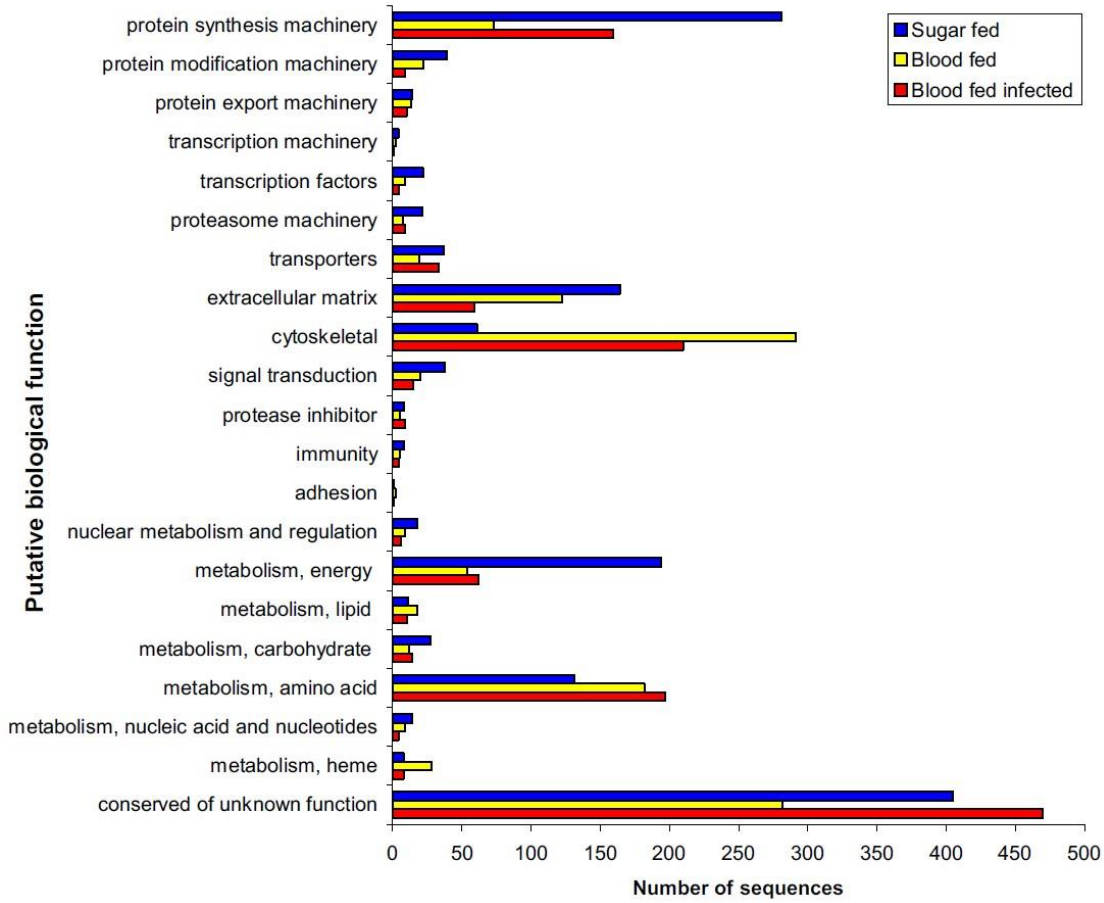


Fig.1.17 – Changes in mRNA levels in the sand fly midgut transcriptome

A number of predicted mRNA levels change during the switch from sugar to blood digestion in the sand fly midgut. The graphic groups identified sequences according to their biological function. Marked difference in sequence numbers can be observed between sugar fed, blood fed and *L. (L.) major* infected sand flies in the protein synthesis machinery, extracellular matrix, cytoskeleton, heme metabolism and conserved genes of unknown function. The graphic was adapted from Ramalho-Ortigão *et al.* (2007) (*208*).

domains, building the scaffold of the PM (*210*). A marked reduction in their expression could weaken the PM facilitating the escape of *Leishmania*. Other proteins, including chymotrypsins, microvilli proteins and the putative peptidoglycan recognition proteins (PGRPs), also showed different expression patterns between blood feed sand flies with and without *Leishmania* parasites (Table 1.2) (*208*, *211*). Interestingly, homologous of these PGRPs in *Drosophila melanogaster* and *Anopheles gambiae* were shown to be induced and involved in their immune defences against bacterial and *Plasmodium* infections, respectively (*212*, *213*). There is also evidence that *Leishmania* does not only affect expression of enzymes, but also affects their activity directly. During the genome analysis of *L. (L.) major*, serine protease inhibitors (ISPs) were identified, which, however, seemed to have no potential targets within the parasite genome (*107*). The ISPs were shown to be effective against sand fly midgut trypsin-like actrivity *in vitro* (*214*) and against vertebrate macrophage serine proteases, however, and one of them (ISP2) has been shown to enhance parasite survival in a murine model (*215*). Also secreted PPGs were shown to inhibit digestive enzyme activity (*216*).

1.4.2. Evading the sand fly's complement system

Not much is known about the sand fly's immune response to *Leishmania* infections. Examples from other parasite-vector relationships, however, have shown a series of complement mechanisms, including secretion of antimicrobial peptides (AMPs), NO, H_2O_2 and PGRPs, that confer control and elimination of midgut infections (*212*, *213*, *217*). For example, anopheline mosquito vectors of the malaria parasite secrete a complement C3 protein-like molecule called thioester-containing protein 1 (TEP1) (*218*). TEP1 was shown to eliminate ~80% of all *Plasmodium berghei* ookinetes in the gut of the susceptible *A. gambiae* G3 strain (*219*). The mode of action of TEP1 is still unknown. Other defence mechanisms include the expression of nitric oxygen synthase (NOS) in *Plasmodium* invaded gut epithelial cells along with peroxidases like heme peroxidases (HPX) and NADPH oxidases (NOX) (*217*, *220*). These enzymes generate nitric oxides (NOs) that act in toxic protein nitration. It was recently shown that sand flies can generate reactive oxygen species (ROS) too, which are used to combat microbial pathogens and to keep their commensal flora in check (*221*). It was shown that *Leishmania* is sensitive to ROS, but on infection with *Leishmania*, no increased ROS activity is observed in the sand fly midgut, probably because *Leishmania* do not induce sand fly immunomechanisms by tissue damage and invasion of the

Table 1.2 – Manipulation of sand fly midgut enzyme up- and downregulation during blood meal digestion by *L. (L.) major*

Putative function	Sugar fed	Blood fed	*L. major*
Microvilli protein (PpMVP1)	0 →	195 →	70
Microvilli protein (PpMVP2)	1 →	60 →	42
Microvilli protein (PpMVP3)	39 ⊣	8	
Microvilli protein (PpMVP4)	0 →	18	
Peritrophin (PpPer1)	0 →	54 →	16
Peritrophin (PpPer2)	152 ⊣	45 →	35
Ferritin light chain (PpFLC)	6 →	18 →	3
Chymotrypsin (Ppchym2)	0 →	36 →	8
Trypsin (PpTryp1)	86 ⊣	10 →	82
Trypsin (PpTryp4)	0 →	52	
Unknown (Cluster 73)	13 →	21 →	6
Unknown (Cluster 99)	0 →	29 →	5

Black arrows indicate gene upregulation; red bars indicate gene downregulation. The numbers indicate cluster overrepresentation.
The table was adapted from Ramalho-Ortigão *et al.* (2007) (*208*).

hemolymph like *Plasmodium spp.* in mosquitos. Limiting effects of the sand fly immune response on *Leishmania* were demonstrated by the downregulation of Caspar-like genes, which act as negative regulators of the sand fly's immune system (*222*). How exactly *Leishmania* evades the sand fly's complement system is unknown, but the parasite's surface PPG was proposed to play a vital role (*216*). The parasites LPG may also be involved, although there is evidence that LPG may actually induce the secretion of sand fly defensin, an AMP, secreted by the epithelium, which is effective against microbes in the midgut (*223*).

1.4.3. Peritrophic Matrix

The first defined physical obstacle for *Leishmania* parasites to overcome is the PM (also known as the peritrophic membrane (*224*)), which compartmentalizes the midgut into an ecto- and an endoperitrophic space (EnS) during blood digestion. The PM is an acellular envelope secreted by the intestinal epithelium of both AMG and TMG in most known haematophagous insects and consists of chitin, proteins and glycoproteins (Fig.1.18) (*225*, *226*). The secretion of the PM is stimulated by the distension of the midgut due to the consumption of blood (*227*). The PM protects the sensitive midgut epithelium against pathogens and abrasion by food particles (*210*, *225*, *226*) and plays a role in heme detoxification (*228*). Involuntarily, the PM confers some protection to *Leishmania* amastigotes within the EnS by preventing rapid diffusion of digestive enzymes, giving the parasite time to transform into LPG-covered procyclics (*151*).

By day 3 – 4 PBM, the PM becomes a trap for *Leishmania* parasites due to blood meal defecation; if they are not able to escape the PM, which happens in cases of non-permissive infections, the parasites will be excreted (*150*, *229*). The efficacy with which different parasites strains escape the PM is a defining factor in their development within a vector species (*157*). Although *Leishmania* accumulate at the anterior of the PM and secrete their own chitinases, which are essential in the degeneration of the SV (*230*), but are inhibited by haemoglobin (*231*), in most cases parasites escape into the AMG lumen occurs through a posterior rupture of the PM caused by sand fly-derived chitinases (*151*). For some natural parasite-vector combinations, like *L. (L.) major – Ph. (Ph.) papatasi*, however, it has been observed that the escape from the PM can occur through a rupture at the anterior end by the use of the parasite-derived chitinase (*232*, *233*), but that is not the case for *L.*

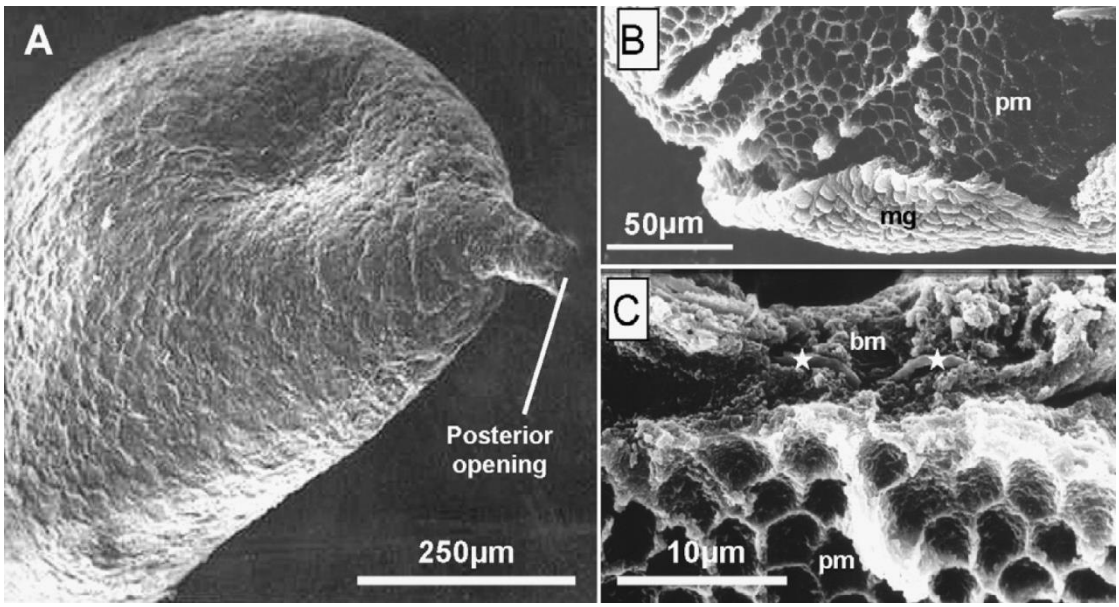


Fig.1.18 – Peritrophic matrix of a *Phlebotomus papatasi* female sand fly

A) Intact PM dissected from a *Phlebotomus papatasi* midgut 24 h PBM. B – C) Midgut (mg) with PM fixed in 2.5% glutaraldehyde in cacodylate buffer & post-fixed in 1% osmium tetroxide in cacodylate buffer 48 h PBM. A honey comb pattern is observed on the PM from the impression of the removed midgut epithelium cells. C) *Leishmania* parasites are visible (*) within the blood meal (bm) (*152*).

(L.) major in *Ph. (Ph.) duboscqi* (*156*, *234*). Parasite numbers decrease with the defecation of the blood meal initially and then stabilize due to the inability of nectomonads to proliferate. Numbers increase again in the TMG once proliferating leptomonads appear. It has been also observed that in most parasite – vector combinations the development of procyclics into nectomonads coincides with the breakdown of the PM (*156*).

1.4.4. *Leishmania* interactions with the sand fly alimentary canal

It has been proposed that to establish a persistent infection in the sand fly midgut, *Leishmania* parasites must attach to the midgut epithelium to prevent their excretion (*235*). This made the topology of the inner surface of the alimentary canal very important to the development of *Leishmania* (*152*). *Leishmania* parasites attach to the midgut's single layer of microvillar epithelium by insertion of their flagellum in between the microvilli (Fig.1.9A.1) (*152*). This anchoring is strengthened in case of *L. (Leishmania) spp.* by sand fly lectin-parasite LPG interactions against the peristaltic action of the midgut (*169*, *236*, *237*). It was also shown that nectomonads and leptomonads have much higher affinity for the intestinal epithelium binding than procyclics and metacyclics, which was attributed to their stage specific LPG modifications (*155*). Since different *L. (Leishmania) spp.* express differently modified LPG and different *Phlebotomus spp.* and *Lutzomyia spp.* present different lectins and lectin-like molecules, it was observed that vectors, whose naturally infecting *Leishmania spp.* present highly branched, species-restricted LPGs, had a much greater parasite-specific affinity than those that were naturally infected with *Leishmania spp.* presenting unsubstituted or poorly substituted LPG (*140*, *238*). In addition, an LPG-independent attachment mechanism has been observed in permissive parasite-vector combinations (*239*, *240*). *Leishmania* parasites has also been observed to colonize the FG (pharynx, cibarium and mouthparts) (*117*, *241*). Since the FG does not have a microvillar epithelium, but is cuticle-lined, the mode of parasite attachment here must differ from the mechanism in the midgut and is currently unknown. However, colonization of the FG was shown to be non-essential for the transmission of parasites and, therefore, may be only a secondary effect of the parasite-induced destruction of SV (*162*, *198*). In case of *L. (Viannia) spp.* much less is known about these processes.

1.4.5. Inhibiting sand fly gut peristalsis

Lysate proteins from cultured *Leishmania* parasites are able to reversibly stall

and even completely inhibit intestinal peristalsis of the sand fly mid- and hindgut (but not of the rectum) in a dose dependent manner (*160*, *242*, *243*). This activity peaks in late log-phase to early stationary phase parasites, coinciding with nectomonad escape from the PM. Lysate activity is then reduced as differentiation progressed. The capacity of the culture medium to inhibit peristalsis, however, persists much longer than that of parasite lysate, suggesting a parasite secreted inhibitor, whose activity could be abolished by trypsin and chymotrypsin digests. A 12 Kilo Daltons (KDa) peptide was identified, termed stambhanin, which was proposed to be the active compound modulating this effect (*160*). Stambhanin is likely to function by high-jacking a common ligand-receptor pathway for the insect's own neuropeptides that regulate the visceral muscle activity (*244*). This hypothesis was supported by the observation that *L. (L.) major* lysate proteins were also able to inhibit muscle contraction in oviducts and dorsal blood vessels even in non-*Leishmania*-vector insect species, like *Aedes aegypti* (*242*). Observations made in muscle tissues taken from rodents inhibited by *Leishmania* lysate proteins suggested that Stambhanin either inhibited the influx or promoted efflux of Ca^{2+}-ions from the susceptible muscle (*243*). Interestingly, Strambhanin peak activity in culture was proposed to coincided with a natural low level of trypsin and chymotrypsin in the midgut after blood meal digestion (*245*). Due to the lack of peristalsis, liquid accumulates in mid- and hindgut (*160*). After 30 min. of incubation with *Leishmania* lysate proteins, the midgut volume was enlarged by almost 50% and the HG by ~57%. This may allow unattached parasites to swim freely in the liquid medium by use of their flagellum. However, no *in vivo* work has been performed to confirm these *in* vitro results.

1.4.6. Promastigote secretory gel

The production and function of PSG was reviewed extensively by Rogers (2012) (*246*). Its secretion was proposed to be a universal *Leishmania* mechanism (*247*, *248*). PSG is a dense gelatinous matrix consisting mainly of fPPG, a high molecular weight glycoprotein, the largest molecule secreted by *Leishmania spp.* (3 – 6 nm in diameter and up to 6 µm in length), although other parasite and sand fly derived molecules may be incorporated into the PSG (*247*, *248*). Filamentous PPG is another [Gal-Man-PO_4] repeat unit containing phosphoglycans synthesised by *Leishmania*, like LPG and surface PPG, but is secreted from the parasite flagellar pocket (*237*, reviewed in *238*). It comprises of a serine-rich protein backbone with LPG-like phosphoglycans

attached to every second amino acid, which make ~75% of the molecules total mass and giving the molecule its filamentous character (Fig.1.19) (*250*). Gel formation occurs at fPPG concentrations >10 mg/ml and *L. (L.) mexicana* was shown to secrete up to 1 g/ml of fPPG in *Lu. (Lu.) longipalpis* (*161*, *247*, *248*).

The mature PSG, *in situ*, is a sausage-like obstruction in the entire TMG distending through the disintegrated SV into the FG (Fig.1.9) (*161*, *249*). Nectomonads and leptomonads attach by a yet to be identified mechanism to the PSG scaffold, while metacyclics can transverse the PSG unhindered accumulating at its poles to be in an ideal position for transmission (M.E. Rogers, 2013, unpublished). Leptomonads are by far the predominant morphological parasite form found in the PSG, although nectomonads bind to PSG with higher affinity, but it was proposed that most parasites have become leptomonads by the time of PSG formation (*161*). PSG can be detected as early as day 2-3 PBM and peaks at around day 5-7 PBM depending on *Leishmania spp.*-sand fly combination coinciding with the peak appearance of leptomonad forms in the sand fly (*145*, *161*, *251*).

The presence of the PSG-plug actively manipulates sand fly feeding behaviour causing difficulties in feeding (Fig.1.20), forcing them to probe more frequently, feed for longer and even change host more frequently, optimizing the chance of parasite transmission (*252*, *253*). It was shown that the amounts of regurgitated PSG during feeding correlated to the level of parasite burden and to the number of parasites transmitted (*254*).

Other secreted PPG functions included parasite protection from hydrolytic attack during blood meal digestion in early stage promastigotes, as shown *in vitro* (*169*, *216*), and support of metacyclic invasion of macrophages and parasite survival within the phagosomes (*54*, *255*). PSG also manipulates mammalian host immune responses by inducing arginase-1 expression in the Th2 response that promotes parasite survival (*254*). PSG is also able to recruit neutrophils and macrophages, which are the primary target cells of *Leishmania* infection, to the bite site (*140*, *254*). Interestingly, PSG is even proposed to improve the transmission of parasites from the mammalian host to the sand fly (M. E. Rogers, unpublished).

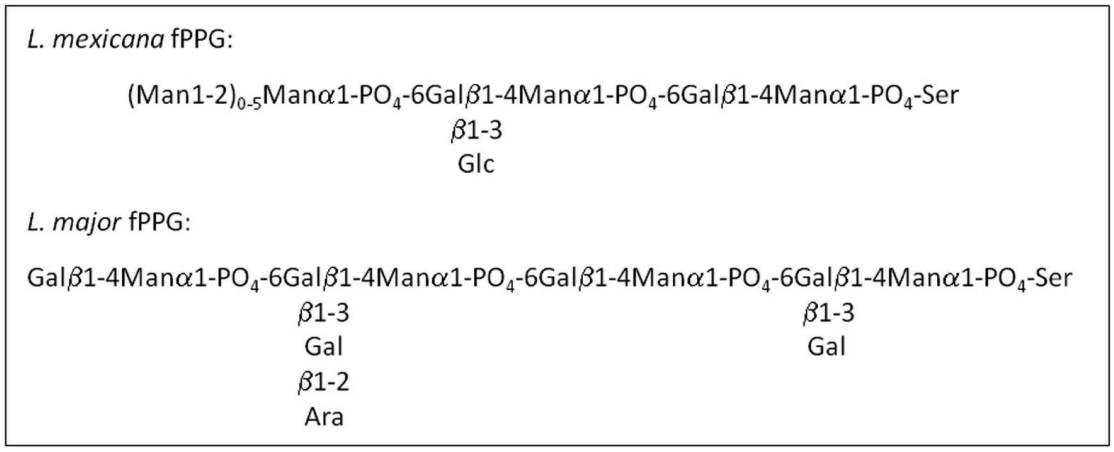


Fig.1.19 – Filamentous proteophosphoglycan

Structure of the main PSG component, filamentous proteophosphoglycan (fPPG), from two *L. (Leishmania) spp.* (*246*)

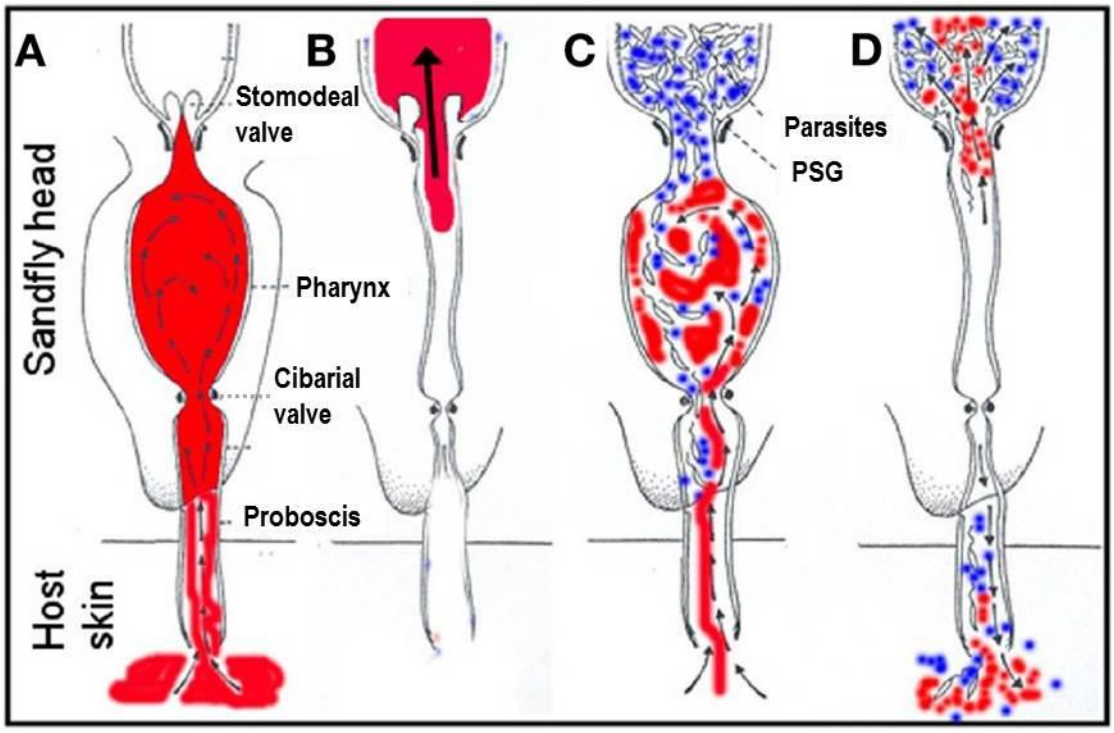


Fig.1.20 – Changes in sand fly feeding due to PSG-plug

The pumping action of the pharynx is facilitated through the presence of the cibarial valve between proboscis and pharynx and the SV between pharynx and TMG. A-B) In uninfected sand flies blood gets pumped from the blood pool in the skin through the proboscis into the pharynx and from there through the SV into the TMG. C-D) In an infected sand fly the PSG plugs up the SV and prevents immediate blood uptake into the TMG, so that the blood mixes with the parasite infested PSG in the pharynx before being regurgitated again. This process is repeated a few times before the SV and TMG are cleared enough for blood uptake. In this fashion parasites are effectively introduced into the blood pool and the gelatinous PSG dissolves again in the blood pool. The image was modified from Rogers (2012) (*246*)

1.4.7. The stomodeal valve

The SV is a unidirectional valve that prevents reflux of midgut contents, effectively forming a barrier between the midgut and the FG. It consists of “a cuticle-lined sphincter, which projects into the lumen of the TMG and an additional circular lobe external to it” (Fig.1.21A) (*152*). During *Leishmania* infection, the SV becomes heavily colonized with flaccid-looking, non-motile haptomonads via hemidesmosome-like structure on their modified flagellum (Fig.1.9.A.3 & 1.21B). It has been shown that these haptomonads secrete chitinases, which act in the degradation of the SV permitting metacyclics passage into the FG (*54*, *233*, *256*). The PSG was proposed to help to pry the SV apart, because it extends into the FG, but it is not clear, if this is only a consequence of SV degradation. SV degradation was proposed to be essential for parasite transmission, despite the lack of clear supporting evidence for this theory (*162*).

1.4.8. Bacterial midgut flora

The digestive system of all living animals is home to a variety of microorganisms, which naturally exist there and make up the commensal flora. Some of them are beneficial to the host as they support digestion and nutrient absorption, while others are more parasitic in nature. The sand fly microbiota is a consequence of the sand fly’s life-style of visiting plants and animals for alimentation, while their larvae feast on detritus, exposing the insect to a multitude of microorganisms of the phyllosphere, soil and fauna (*257*). There is very little known about the sand flies commensal flora and studies are only beginning to emerge (*257–262*), but a recent study in wild caught *Lu. (Lu.) longipalpis* showed that the array of bacterial phylotypes present in its midgut is comparatively limited (*263*) compared to other insects such as cockroaches, termites and crickets (*264*). Among a set of phyllosphere bacteria, it was found that sand flies also carried plant and human bacterial pathogens, like *Ralstonia spp.* (causative agents of bacterial wilt) and *Chryseobacterium meningosepticum* (causative agents of neonatal meningitis), potentially acting as vectors for these pathogens, too (*258*, *263*). Infection studies with different protozoan parasites have shown that bacteria in the midgut of insect vector species impact on the survival of the parasites in an either direct (by competition) or indirect manner (by modulating sand fly immune responses). Comprehensive studies addressing this topic are still few, however (*221*, *261*, *265–270*), although the idea of commensal bacterial in competition with *Leishmania* was first voiced by Theodor in 1957 (*271*).

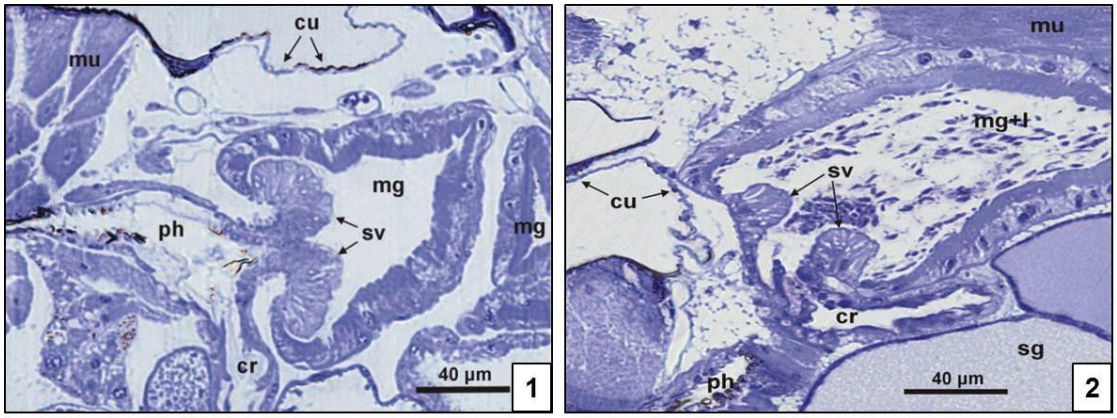


Fig.1.21 – Stomodeal valve of un-/infected *Ph. (Ph.) papatasi* & *Ph. (Ph.) duboscqi*

1) Semithin section of *Ph. (Ph.) duboscqi* embedded in LR White resin and stained by toluidine blue. The filamentous structures of the SV are indicated by arrows. 2) Semithin section of *Ph. (Ph.) duboscqi* infected with *L. (L.) major* treated in the same way as in image 1. The degradation of the filamentous structures is indicated by arrows. Figures were adapted from Volf *et al.* (2004) (*162*)

Abbreviations: cr, crop; cu, cuticle; mg, midgut; mg+l, midgut containing *Leishmania*; mu, muscles; ph, pharynx; sg, salivary glands; sv, stomodeal valve.

Evidence suggested that there is competition between *Leishmania* and certain bacteria, which can limit parasite infection in the sand fly (*261*, *265*, *266*).

1.4.9. Sexual reproduction of *Leishmania*

It has been shown that *Leishmania* can reproduce sexually in the sand fly midgut (*272*), but it is not clear if this can also occur in the amastigote stage. It had been believed that *Leishmania*'s primary mode of reproduction was clonal (*119*, *273*, *274*), but experimental data (*272*, *275*) and isolation of naturally occurring *Leishmania* hybrids (*276–280*) suggested that sexual reproduction may occur naturally within the sand fly midgut by either endogamy, autogamy, automixis or selfing, as suggested by the high levels of observed homogeneity in *Leishmania* populations (*281*, *282*). Experimental verification of genetic exchange within the sand fly was delivered by Akopyants *et al.* (2009) using two transgenic strains of *L. (L.) major* resistant to different antibiotics resulting in hybrids resistant to both antibiotics after sand fly passage, but with kinetoplast DNA (kDNA) maxicircles content derived from only one parent (*272*). Interestingly, seven of 18 characterized clones from this study were triploid, rather than diploid. Aneuploidy is well tolerated by *Leishmania* and even exploited as a mean of gene copy multiplication (*109*). Sex in the sand fly indicated it as the primitive *Leishmania* host and field isolation of naturally occurring *Leishmania* hybrids (Old World: *L. (L.) donovani*/*L. (L.) aethiopica* (*276*); New World (*280*): *L. (V.) braziliensis*/*L. (V.) panamensis* (*277*)*; L. (V.) braziliensis*/*L. (V.) guyanensis* (*278*)*; L. (V.) braziliensis*/*L. (V.) peruviana* (*279*)) suggested that genetic exchange was not just occurring among closely related *Leishmania* strains, but is also possible among closely related *Leishmania* species. Also, the isolation of different parasite strains/species from the same mammalian host suggests that co-infection of multiple *Leishmania* strains/species occurs in the same sand fly allowing natural hybrid generation (*181*). However, the exact mechanism of *Leishmania*'s genetic exchange remains uncharacterized.

1.4.10. Transmission to a mammalian host

In nature, *L. (Leishmania)* parasite transmission to a new mammalian host occurs during a second blood meal, which the sand fly seeks 1-2 weeks after the first blood meal, depending on species. Parasite transmission had been proposed to occur in two opposing ways: regurgitation of parasites from the TMG during blood meal uptake (*161*, *248*) or by host inoculation with parasites present in the FG only on the insertion of the sand fly's mouthparts

into the host's skin (*252*, *283*, *284*). With the characterization of the PSG the regurgitation model for parasite transmission was generated that incorporates all the available data on parasite transmission. In this model, the TMG blockage by PSG forces sand flies to regurgitate the metacyclic loaded PSG-plug before a blood meal can be taken, which may take several attempts and increases sand fly probing rates. This model also explains why at least 10x more parasites are found in bite sites than are present on average in the FG (*54*, *162*). The SV degradation is considered to be important for successful parasite transmission, too. Once inoculated into the host skin, metacyclics are endocytosed by epithelial macrophages. Once the parasites reach the phagolysosomes, they begin to transform into amastigotes (small, round, aflagellate, proliferating) (*145*).

Much less is known about transmission of *L. (Viannia) spp.*, but it has been proposed that transmission occurs via the forward route, too, which was suggested by the observation that members of the *Viannia* subgenus do also secrete PSG or at least a PSG-like molecule in the TMG (*117*, *153*).

1.5. The *Leishmania* genome

The genome of *L. (L.) major* Friedlin VI strain (MHOM/IL/81/Friedlin) (FVI) was published as the first completely sequenced *Leishmania* genome in 2005 (*107*) in parallel with the genomes of the related human infective kinetoplastid parasite species, *Trypanosoma brucei* (*285*) and *T. cruzi* (*286*). The *L. (L.) major* genome consists of ~32.8 Mb spread over 36 chromosomes, which is the typical chromosome number for Old World *Leishmania spp.* (*287*) – some New World species like *L. (L.) mexicana* and *L. (V.) braziliensis* have only 34 and 35 chromosomes, respectively, due to chromosome fusion events (*108*). All published trypanosomatid genomes are sequence repeat rich – in *T. cruzi* up to 50% of the whole genome (*286*) – and it was shown that of the 8272 protein-coding genes identified in *L. (L.) major*'s genome, 6158 genes had orthologous in the *T. brucei* and *T cruzi* genomes with a high level of synteny (*288*). Considering the long evolutionary divergence (>200 Million years (*289*)) between the two genera *Leishmania* and *Trypanosoma*, this high level of genome similarity was surprising (*288*). When the genome sequences of *L. (L.) infantum*, *L. (V.) braziliensis* (*108*) and *L. (L.) mexicana* (*109*) were published, it was shown that the content of *Leishmania* species-specific genes within the genus was surprisingly low (2, 14, 19 and 67 species-specific genes for *L. (L.) mexicana, L. (L.) major, L. (L.) infantum* and *L. (V.) braziliensis*, respectively), while 7392 genes

were shared by all four species with conserved gene synteny, despite the 46 million years of evolutionary divergence (Fig.1.22) (*109*, *123*). Compared to the *L. (L.) infantum* genome, all species-specific genes are either present or not clearly absent from analysed *L. (L.) donovani* strains (*290*). Further work showed that species-specific genes are generally conserved among members of the same species complex and strains of the same species (*109*). It was proposed that these few species-specific genes encoding predicted proteins of mostly yet unknown functions contribute to parasite tropism and differences in pathology (*108*, *291*), although transgenic studies have not broadly supported this to-date (*292*). Differential gene expression and/or structural and functional components of the genome may be important for parasite pathology, too (*293*). *Leishmania* genomes contain many highly repetitive tandem arrays, which pose a significant problem for automated and de novo genome assemblies, because they cause repeat collapses of unknown length, as has been shown for the *L. (L.) major* cDNA16 locus and the *L. (V.) braziliensis* orthologous HASP locus (OHL) in the past (*107*, *293*, *294*).

Leishmania parasite sequences show inter- and intra-specific variability in their ploidy (*109*). While the *L. (L.) major* genome is diploid with the exception of chromosome 31, which is multiploid with intra-specific variation in chromosome number, the sequenced genomes of *L. (L.) mexicana* and *L. (L.) donovani* strains show a higher level of ploidy variation and the genome of *L. (V.) braziliensis* is generally triploid with the exception of chromosome 29 and 31, which are multiploid. It has been observed that *Leishmania* parasites can adjust their ploidy as a whole or for particular chromosomes, without suffering any fitness disadvantage, after genetic manipulation, growing in culture for long periods or after drug selection (*295–297*), suggesting that *Leishmania* uses this mechanism as an alternative to forming multi-gene-copy arrays to increase gene copy number (*109*), while eliminating undesired messenger RNAs (mRNAs) by stability factors and/or protein degradation mechanisms (*298*). Unusual for eukaryotic cells, *Leishmania spp.* and *Trypanosoma spp.* have their genes arranged in directional polycistronic gene clusters (PGCs), containing tens to hundreds of functionally unrelated protein-genes (*107*, *299*).

1.5.1. Gene transcription

Transcription of PGCs is constitutive (*300*) and polycistronic, catalysed by the trypanosomatid-specific RNA polymerase II (RNAP II), which shows several differences to RNAP II in other eukaryotes (*107*, *301*). The only functionally

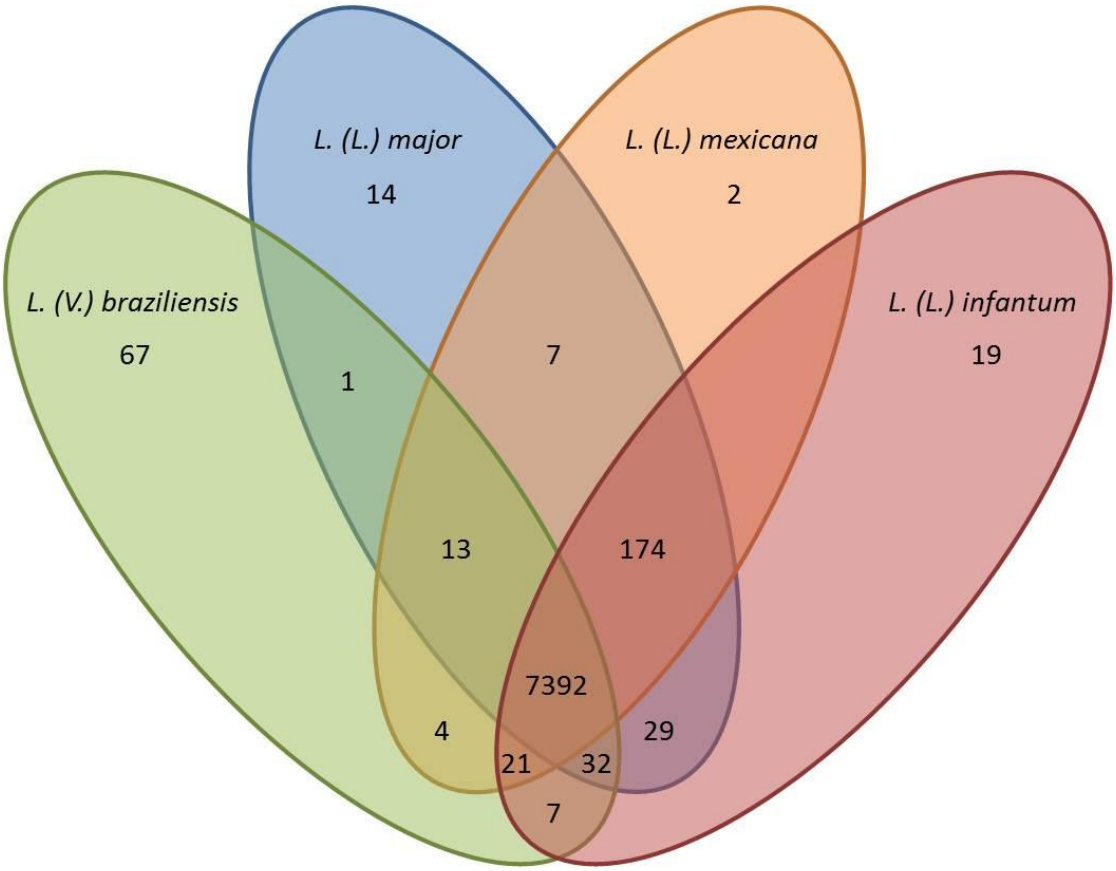


Fig.1.22 – Gene conservation in *Leishmania*

Venn diagram showing number of conserved genes in L. mexicana U1103, L. major Friedlin, L. infantum JPCM5, and L. braziliensis M2904 (*109*).

characterized promoter in the *Leishmania* genomes is the promoter of the splice-leader (SL) sequence genes (*302*) and transcription of PGCs can occur bidirectionally from divergent strand switch regions (SSRs), which have no recognisable promoter motifs, towards convergent SSRs or telomeres by RNAP II (*300*, *303*). Interestingly, TATA-less promoters in higher eukaryotes show similarities to divergent SSRs in trypanosomatids (*304*).

Not much is known about transcription termination in trypanosomatids; no protein factors have been identified yet. In case of RNAP II transcribed SL RNA genes, however, downstream T-tracks support transcription termination, but this is not the case for PGC transcription, where T-rich sequences are common in the intergenic regions (*305*). It was observed that transcription termination of PGCs coincided within tRNA genes at convergent SSRs (*306*), where stable, unmodified H2A/H2B/H3V/H4V octamers were bound, which may support RNAP II transcription termination (*307*). Transcription terminating at convergent SSRs elegantly prevents RNAP II collisions in trypanosomatids.

Polycistronic transcripts are co-transcriptionally processed into monocistronic mRNAs by 5' *trans*-splicing and polyadenylation. In the former, a 39 – 41-nucleotide (nt) splice-leader sequence with a hypermethylated guanosine cap (cap-4), far more complex than the typical m7G cap in other eukaryotes, is spliced onto the 5'-terminus of every mRNA (reviewed in (*308*) &(*309*)). The cap-4 structure is essential for 5' t*rans*-splicing and translation to occur (*310*). The efficiency of *trans*-splicing is proposed to depend on the length of the polypyrimidine tract (*311*) and the 5' untranslated region (UTR) sequence of the downstream mRNA (*312*). There are no consensus polyadenylation signals in trypanosomatids, but polypyrimidine tracts flanking polyadenylation sites, which are 100 – 300 nt upstream of 5' *trans*-splicing sites, help determine polyadenylation sites supported by the coupling of 3' polyadenylation to 5' *trans*-splicing (*313*, *314*).

1.5.2. Gene expression regulation

The constitutive polycistronic transcription in trypanosomatids means that all genes within a PGC are initially transcribed at the same level, and regulation is achieved by post-transcriptional and translational events. In *Leishmania* 2 – 9% of genes are regulated at the mRNA level and another 12 – 18% at the protein level (reviewed in (*315*, *316*)). It has been proposed that in trypanosomatids all the elements of post-transcriptional regulation are

focused in RNA regulons, ribonucleoprotein complexes able to regulate mRNA fate, which are found in other eukaryotes, too (*317*). These control the large scale changes in transcriptome and proteome during life-cycle changes (*316*). Regulation at the mRNA level is characterized by at least three factors (Fig.1.23): *cis*-acting signals, *trans*-acting factors and mRNA degradation machineries (reviewed in (*318*)). Among *cis*-acting signals are SIDER1 & SIDER2 motifs in the 3' UTRs of target mRNAs (*303*, *315*), which in conjunction with RNA-binding proteins (*318*) play an important part in mRNA de-/stabilization (*319*, *320*). U-rich instability elements – similar to the AU-rich elements in mammalian 3' UTRs – are also found in 3' UTRs, although they are more common in *Typanosoma spp.* than *Leishmania* (*321*). These 3' UTR elements may confer stability to stage-specific mRNAs in one life-cycle stage, while conferring instability in another as is the case for the URE in *Leishmania* amastin (*321*). Gene regulation mechanisms have been shown to take cues from environmental signals like chemical triggers, changes in temperature and pH, which have a global impact on mRNA regulation and translation within trypanosomatids (*322*).

In *Leishmania*, mRNA and protein levels frequently do not correlate well and variations in protein levels are often greater than in mRNA levels (*323*). Therefore, post-translational regulation mechanisms are proposed to be of increased importance in these parasites (*324*). 3' UTR motives are important in translation regulation in trypanosomatids: protein factors can bind to or secondary structures may be formed at these regions, hindering efficient translation (*325*). Trypanosomatid genomes also contain a large number of putative kinases, which are proposed to regulate translation by phosphorylation events (*326*). The relative importance of protein stability and sorting in gene regulation is still unknown, however.

1.6. The *L. (L.) major* cDNA16 locus

Differentially expressed genes may be involved in stage-specific developments, like metacyclogenesis, and are rare in *Leishmania* parasites (*327*, *328*). A study using a complementary DNA (cDNA) library to screen *L. (L.) major* procyclic and metacyclic promastigote mRNA extracts for differentially expressed genes, rendered only 4 clones from 25,000 independent recombinants as being differentially expressed between parasite stages (*329*). The *Lm*cDNA16 recombinant, after which the corresponding locus on the *L. (L.) major* chromosome 23 was named, was one of these four and recognised four distinct

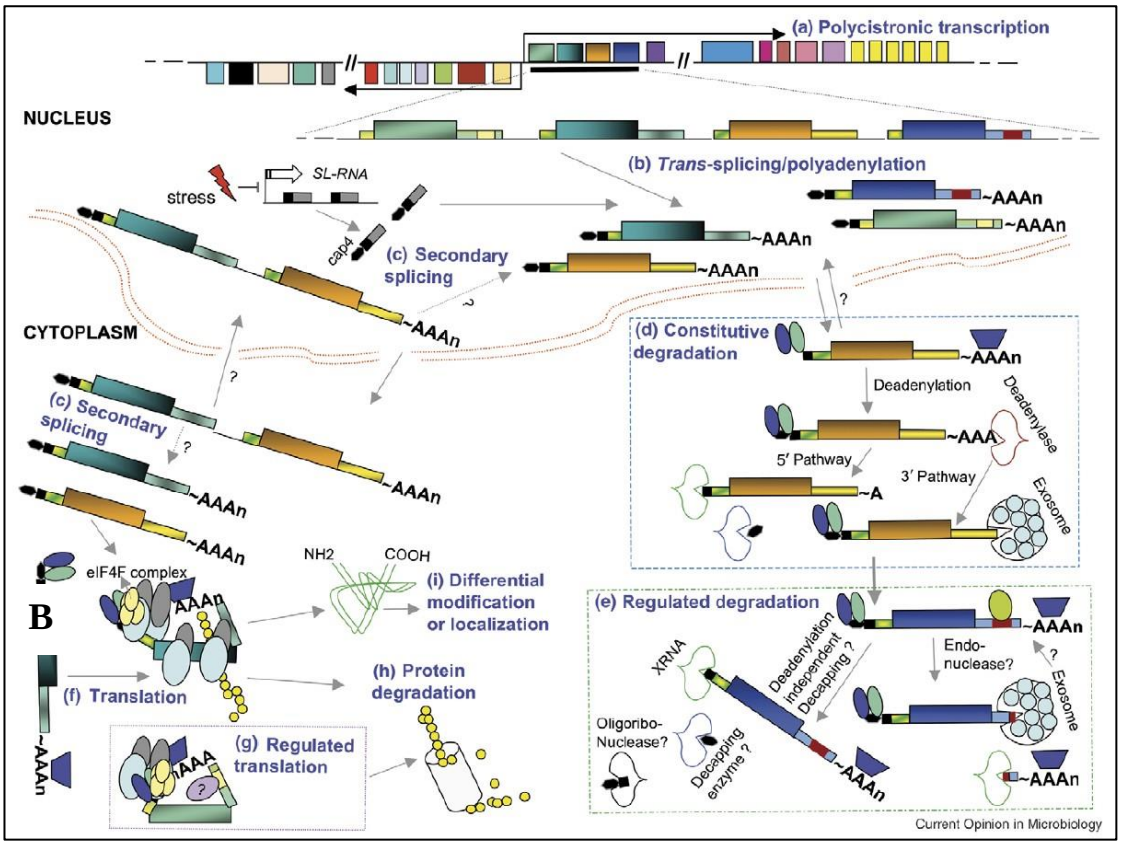


Fig.1.23 – Post-transcriptional gene regulation in *Leishmania* parasites

a) Polycistronic transcription from divergent SSRs across the polycistronic gene cluster (PGC). b) Polycistronic transcripts are processed into monocistronic mRNAs by coupled 3'-polyadenylation and 5'-*trans*-splicing, which adds the splicer leader sequence hypermethylated guanosine cap. The repression of *SL*-sequence transcription during stress was implicated as mean of post-transcriptional regulation of protein genes. c) Secondary splicing was implicated as another mean of gene regulation. Only after the secondary splice event can the mRNAs be translated successfully. The location of this event has not been established yet. d) This constitutive degradation pathway was proposed to work by release of poly(A)binding proteins (PABs) followed by 3'-poly(A)-degradation and 3'exsosomal degradation of the mRNA with cytosolic processing bodies (P-bodies). e) The deadenylation-independent regulated degradation pathway was proposed to act in the degradation of stage-specifically regulated and instable mRNAs. f) 3' UTR elements like SIDER1, SIDER2 and U-rich instability elements (UREs) can prevent effective mRNA translation in a life-stage specific manner. Translation regulation by g) protein degradation, h) protein targeting for degradation and i) post-translational protein modifications, like phosphorylation have all been implicated as means of gene regulation. The image was taken from Haile (2007) (*315*).

transcripts of varying expression levels between life-cycle stages, since renamed as the HASPA1, SHERP, HASPB and HASPA2 genes (Fig.1.24). The cDNA16 locus was sequenced and assembled manually (accession no. AJ237587) and its organization confirmed by Southern blotting and hybridization across the region. According to Flinn & Smith (1992), HASPA1 is the most upstream gene in the *Lm*cDNA16 locus in *L. major*, followed by two identical copies of SHERP (SHERP1 and SHERP2), then HASPB followed by HASPA2 (Fig.1.25.A) (*330*). Comparative studies have since shown that all HASP and SHERP genes are found in the same chromosomal region in all *L. (Leishmania) spp.*, while in *L. (Viannia) spp.*, related but divergent sequences are found (*108*, *294*, *331*).

In *L. (L.) major*, deletion of the entire cDNA16 locus prevented completion of metacyclogenesis within the sand fly vector (*146*), but not in culture (*332*) and did not perturb the parasite's virulence *in vitro* or *in vivo* (*332*). Because of this deletion, mutant metacyclogenesis was stalled mostly in the nectomonad stage and parasites never colonised the SV. The overexpression of the entire *Lm*cDNA16 locus by episomal replacement into the null background, however, caused avirulence in the deletion mutant and both null and overexpression mutants were sensitive to complement-mediated lysis (*332*).

1.6.1. HASPA1, HASPA2 and HASPB

HASP stands for Hydrophilic Acylated Surface Protein (*331*). HASPA1, HASPB and HASPA2 are highly related members of the same gene family, but non-identical (*330*). All three HASPs share the same *N*-terminal (first 17 amino acids) and *C*-terminal (last 35 amino acids) regions, but differ in the central section (Fig.1.25.B & C) (*333*, *334*). The HASP *N*-terminus contains a SH4 domain, which is co-translationally *N*-myristoylated at glycine 2, which targets the proteins transiently to the cytosolic side of the Golgi, where it gets post-translationally palmitoylated at cysteine 5 (*335*). This dual acylation has been shown to be sufficient and essential for HASPB trafficking to the cell surface and its tethering, in the absence of a secretory signal sequence, a GPI-anchor consensus sequence or a membrane spanning domain (*336*–*340*). This was shown by substitution of either glycine 2 or cysteine 5, which caused the HASPs to become exclusively cytosolic and trapped at the Golgi, respectively (*335*). HASPB was observed, however, to be shed by metacyclics *in vitro* (*336*). Therefore, the HASPs were classified as non-classically secreted proteins that were transferred to the cell surface from the cytosolic phase of the Golgi by a Golgi-vesicle-independent mechanism to

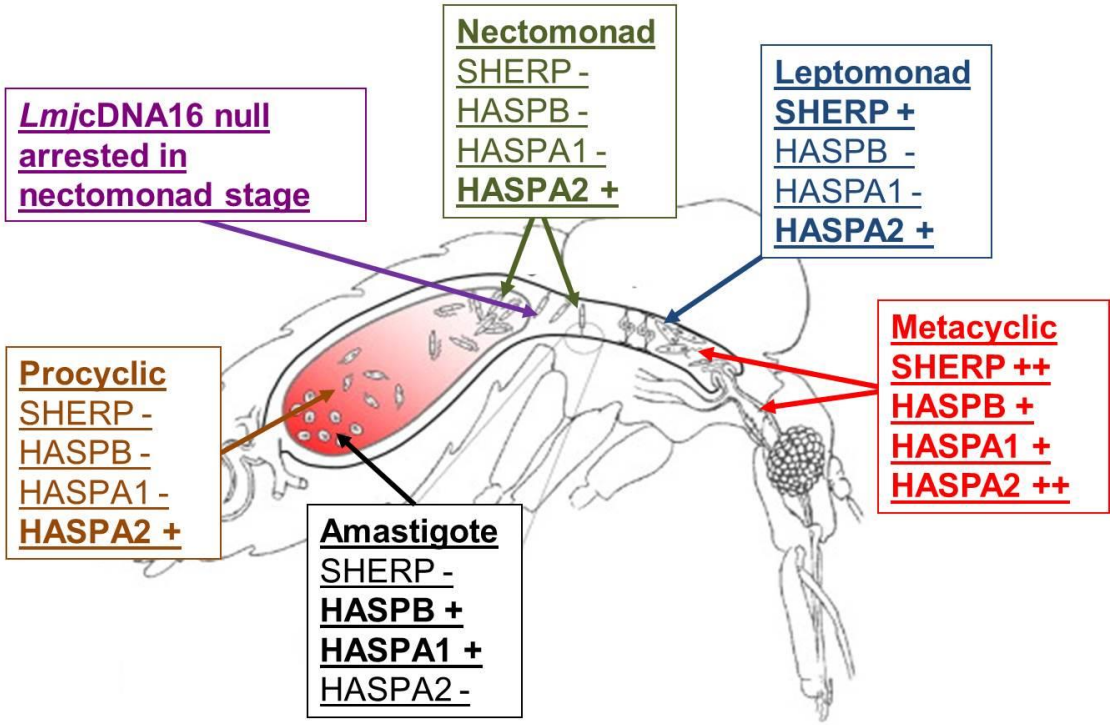


Fig.1.24 – Schematic of regulation of HASP and SHERP genes

There is only reliable data for HASPB and SHERP upregulation at the protein level (*331*, *341*). HASPA1 and HASPA2 regulation has only been shown at the mRNA level to-date (*330*). Based on the available data, HASPA2 is up-regulated as early as the procyclic stage after differentiation from amastigotes and is continuously expressed until the differentiation into amastigotes. SHERP is up-regulated in the late leptomonad stage and peaks in the metacyclic stage, while HASPB and HASPA1 are up-regulated in the metacyclic stage and both continue to be expressed in the amastigote stage, but are down-regulated after differentiation into procyclics. HASPB can be detected even weeks after macrophage infection (communication from Helen Price) and its metacyclic specific expression among promastigotes was shown by confocal microscopy (*146*). The peak expression of HASPB varies slightly between *L. (Leishmania) spp.*; e.g. in *L. (L.) major* HASPB expression peaks in the metacyclic stage (*331*), while in *L. (L.) mexicana* it peaks in the amastigote stage (*294*). The image was adapted from Oliveira *et al.* (2009) (*342*).

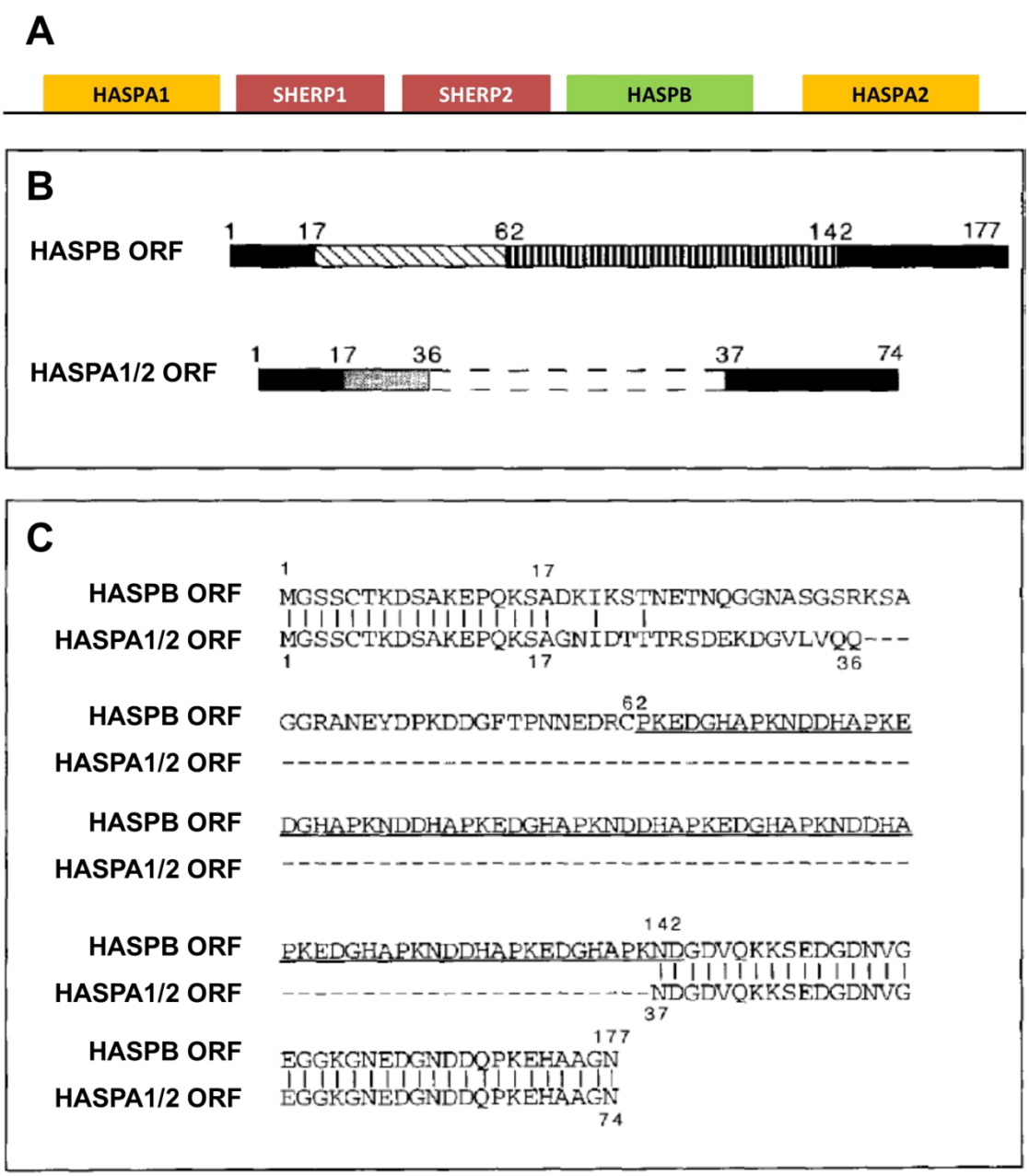


Fig.1.25 – *L. (L.) major* cDNA16 locus and its genes

A) Schematic representation of the *Lm*cDNA16 locus and its 5 genes (not to scale) (*330*). B) Schematic comparison between HASPA1/2 (Gene A/C) and HASPB (Gene B) structure. The repeat region in HASPB resides between amino acid 62 and 142, while only a short sequence is present between amino acids 17 and 36 in HASPA1 and HASPA2. C) Sequence alignment of HASPB and HASPA1/2 based on the amino acid sequence in *L. (L.) major*. The 5.5x PKEDGHA and PKNDDHA repeats in HASPB are between amino acid 62 and 142 (*334*).

the flagellar pocket and then by membrane shedding to the cell surface (*336*, *343–346*). The SH4 domain was also found in other non-classically secreted proteins, like the Src protein family members (*347*). Observations made in live metacyclics showed that HASPB could recycle from the cell surface back to the flagellar pocket and transfer from the cell surface to the flagellum surface, but not vice versa (*336*). It has been proposed that phosphorylation of the threonine 6 in the SH4 domain promoted internalisation of mammalian SH4 domain bearing proteins (*340*). Deletion of this phosphorylation site, however, had no effect on HASPB localization to the plasma membrane (*336*). An unusual characteristic of all three HASPs is their abnormal migration in sodium dodecyl sulphate (SDS) polyacrylamide gel electrophoresis, making them appear about twice their predicted molecular mass, a property attributed to the acidity and the high proline/lysine content of these proteins (*334*).

1.6.1.1. HASPA1 and HASPA2 properties

HASPA1 and HASPA2 genes (formerly known as gene A/C protein [GA/CP]) have substantial sequence identity with identical 5' UTRs and open reading frames (ORFs), but distinct 3' UTRs. Both code for 74 amino acid proteins with a predicted mass of 7.6 KDa each, which bear a protein-specific 19 amino acid sequence in their central region (*334*). They are highly hydrophilic with a low pI of about 4.2. However, their RNA expression pattern is distinct from one another (*330*). HASPA2 is already present at low levels in early log-phase promastigotes and more strongly expressed in metacyclic promastigotes, but undetectable in amastigotes. In contrast, HASPA1 is expressed in metacyclics only and is then maintained in amastigotes (*329*, *331*). No function has yet been assigned to HASPA1 and HASPA2.

1.6.1.2. HASPB properties

HASPB (formerly known as gene B protein [GBP]) is also a highly hydrophilic protein consisting of 177 amino acids (about 18.7 KDa) with a pI of almost 4.8 (*333*). Conversely to HASPA1 and HASPA2, HASPB has an extensive proline rich repeat region between its *N*-terminal (first 62 amino acids) and *C*-terminal (35 amino acids) region, which makes up 45% of the protein (Fig.1.25.C). In *L. (L.) major*, this repeat region consists of 5.5 tandem repeats of a 2x 7-amino acid sequences (PKEDGHA and PKNDDHA). Most amino acids in this repeat region and the *C*-terminus are hydrophilic with merely a few strongly hydrophobic amino acids at the *N*-

terminus. Despite its overall hydrophilicity, HASPB fractionates with the parasites LPG and glycoinositolphospholipids (GIPLs) (*333*), but no interaction between HASPB and LPG could be demonstrated (*337*). HASPB also contains two potential *N*-linked glycosylation sites, which, however, were shown to be unmodified by *N*-glycanase treatment (*333*). The specific function of HASPB remains unclear, although there are strong indicators for an importance in the progression of metacyclogenesis (*146*). In *L. (L.) major*, HASPB is up-regulated strongly in metacyclics and then continuously expressed in amastigotes, but this pattern is not universal; in *L. (L.) mexicana*, HASPB expression occurs later and peaks in amastigotes rather than metacyclics (*294*).

1.6.2. SHERP1 and SHERP2 properties

SHERP stands for Small Hydrophilic Endoplasmic Reticulum-associated Protein and the gene (formerly known as gene D) codes for a small (6.2 KDa), acidic (pI 4.6) protein, which is highly expressed in metacyclic parasites (*341*). SHERP is expressed from two tandem repeated copies within the *Lm*cDNA16 locus (Fig.1.25.A) (*330*), which share 98.8% identity and 100% identity throughout the ORF in *L. (L.) major*. The protein itself is hydrophilic, lacks transmembrane domains and has a high α-helical content with a helix-turn-helix (HTH) motif, which gives the protein a globular fold. However, the secondary structure of SHERP is condition dependent. In an aqueous solution of anionic lipids or detergent SHERP has a highly unordered structure. In a mixture of neutral and anionic phospholipids (DOPC and DOPG) in equimolar amounts or if an amphipathic, anionic detergent (SDS) is added to the solution, the protein adopts its mainly α-helical structure (*348*). This suggested that SHERP's structure is dependent on an anionic environment, which may drive its function. SHERP is a membrane associated protein, which localises to the cytosolic phase of the endoplasmic reticulum (ER) and mitochondrial membranes, where it is in close proximity to phospholipids (*341*). Due to the orientation of its residues, SHERP is amphiphilic, when folded, which suggests protein-protein interaction abilities. The protein also contains two potential phosphorylation sites, which, however, do not appear to be utilized *in vivo*. SHERP's association with the ER and mitochondrial membranes was proposed to be exclusively by weak protein-protein interactions with yet to be identified targets. One potential target could be subunit B of the vacuolar H^+-ATP synthase protein complex, which functions in subcellular compartment acidification and to which SHERP can bind stably (K_D = 2.0 ± 0.1 μM) (*348*).

This may be relevant for metacyclic parasites due to the importance of autophagy for differentiation (*349*). Autophagy is a cellular process common to eukaryotic cells. It serves firstly as a survival mechanism during cell starvation by recycling cytoplasmic constituents, secondly as a mechanism for clearance of damaged and redundant cellular constituents and thirdly, is involved in cell re-modelling during differentiation (*350*). *Leishmania* metacyclogenesis is essentially a cell differentiation process under cell starvation conditions and a functional autophagosomal system has been shown to be essential for completion of metacyclogenesis (*349*).

1.7. Leishmania (Viannia) braziliensis Orthologous HASP Locus

Comparative studies have shown that in *L. (Viannia) spp.*, different but related genes are present in the same chromosomal context as the cDNA16 locus (Fig.1.26) (*108*, *331*). According to chromosomal assemblies from GeneDB, there are two different ORFs (LbrM.1110 and LbrM.1120) in this orthologous HASP locus (OHL) in *L. (V.) braziliensis*, similar to those found in the same chromosomal region in the related insect parasite, *Leptomonas seymouri* (*294*). Analysis of these two ORFs in *L. (V.) braziliensis* showed structural and biochemical similarities of the protein products to the HASPs in *L. (Leishmania) spp.* (*294*). Both ORFs contained a *N*-terminal SH4 domain required for *N*-myristoylation and palmitoylation (Fig.1.27) and at least LbrM.1110 has been shown to localize to the cell surface of the rudimentary flagellum in amastigotes, but not in metacyclics (*294*). These proteins contain a varying number of central tandem repeats, which consist of 30 nt per repeat and have antigenic properties, like those of HASPB. LbrM.1110 expression patterns were similar to *L. (L.) major* HASPB expression patterns. Southern blot analysis of the *Hind*III/*Xho*I-digested genomic DNA (gDNA) of *L. (V.) braziliensis*, *L. (V.) guyanensis* and *L. (V.) peruviana* suggested that the currently available sequence of the locus (~7 Kb in length) was falsely assembled by automated analysis with reference to the *L. (L.) major* genome (*294*). Tandem repeat collapses of a highly conserved ‘AB’- motifs (~3,2 Kb) were found to be responsible (*294*). Similar misassemblies had also occurred in the cDNA16 locus of *L. (L.) major*. To-date the misassembled OHL has not been corrected in TriTrypDB.

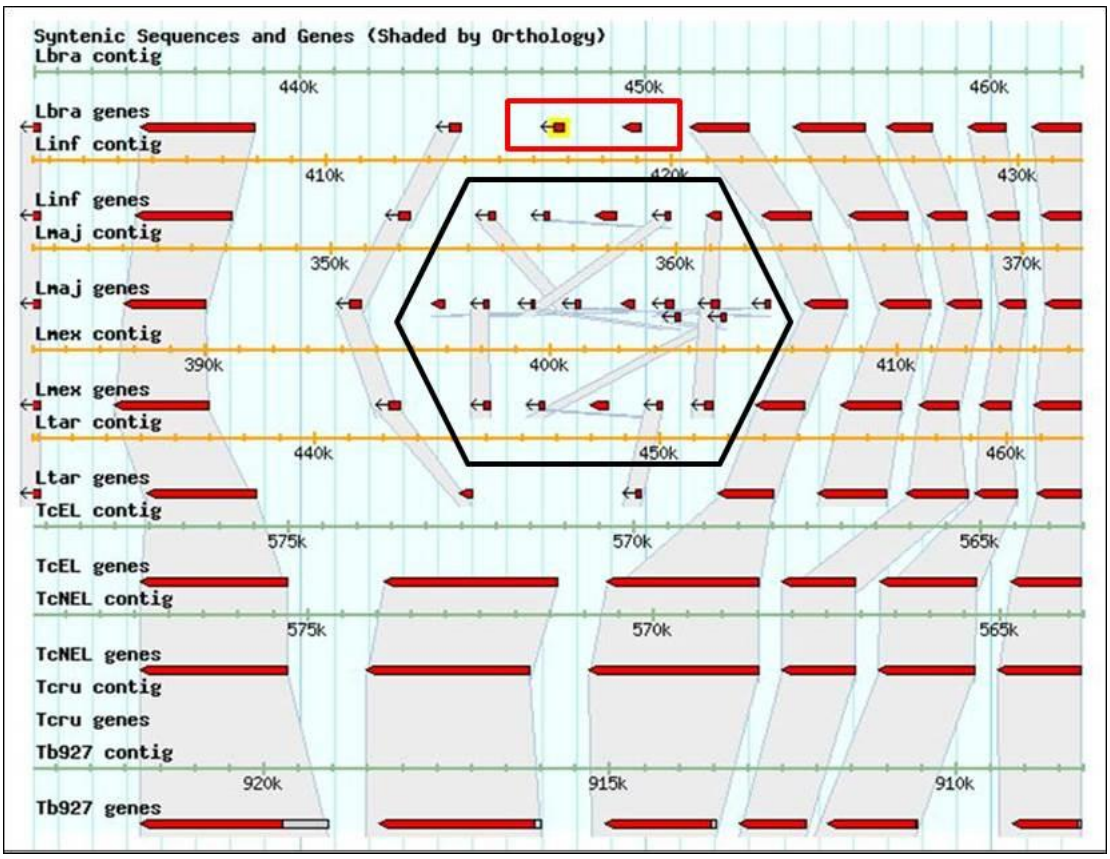


Fig.1.26 – OHL alignment to cDNA16 locus

Alignment of the OHL and cDNA16 locus region on chromosome 23 of *Leishmania* and *Trypanosoma spp. Lbr*M.23.1110 and *Lbr*M.23.1120 (red box) are not conserved in the genomes of any *L. (Leishmania) spp.* presented here, which have the HASP and SHERP genes (black box) in place of the OHL locus genes. Image taken from TryTripDB.

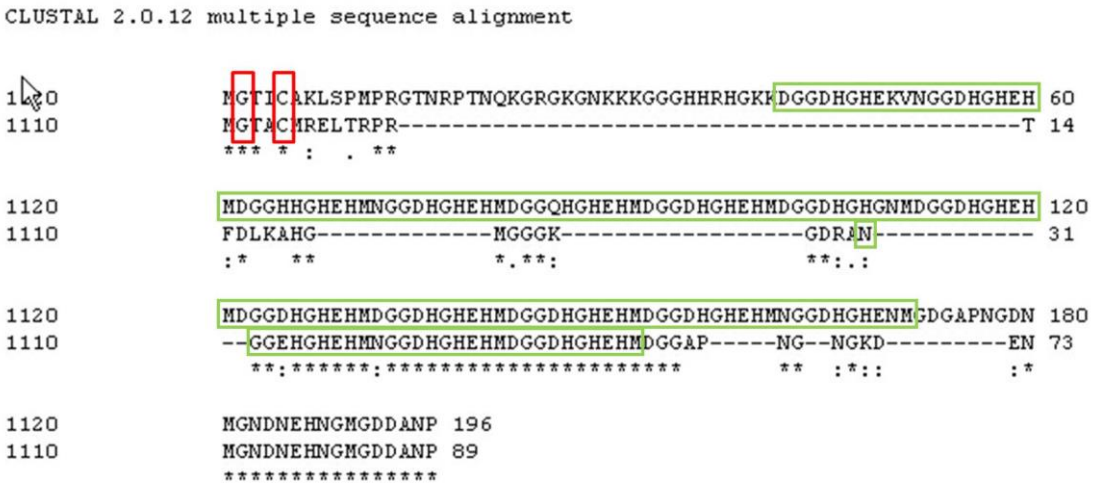


Fig.1.27 – ClustalW alignment of *Lbr*M.1110 and *Lbr*M.1120 amino acid sequences

The red boxes mark the glycine[2] (G) and the cysteine[5] (C) of the N-terminal SH4 domain essential for *N*-myristoylation and palmitoylation, respectively. The green box marks the internal tandem repeat areas (10 amino acids = 30 nucleotides per repeat). Adapted from Depledge thesis (2009).

1.8. Project aims

The study of metacyclogenesis is key to the understanding of *Leishmania spp.* transmission from the sand fly vector to a mammalian host. This event has been studied in members of the *L. (Leishmania)* sub-genus and it has been shown that the genes of the cDNA16 locus are essential for metacyclogenesis completion, although their functions, interaction and individual importance to the process are not yet known. To date, metacyclogenesis has not been investigated in the *L. (Viannia)* subgenus, which bears genes of divergent sequence to the cDNA16 genes in the same chromosomal region with predicted protein products bearing structural and biochemical similarities to the HASPs.

The main focus of this study is the function of HASPs and SHERP in metacyclogenesis within the sand fly vector. Several questions will be addressed:

- Is there one particular key player among the HASPs and SHERP for metacyclogenesis completion?

- Is there an interdependence between the HASPs and SHERP in their expression regulation?

- Are there differences in the expression patterns of HASP and SHERP genes between *in vitro* and *in vivo* development?

- Are there specific vector components that influence HASP and SHERP gene regulation?

- What parasite processes are affected by HASPs and SHERP deletion *in vivo* and does this knowledge lead to hypotheses on their function?

- What is the true sequence, gene content and arrangement of the OHL in *L. (V.) braziliensis* and what further parallels can be drawn to the cDNA16 locus?

2. CHAPTER II. – Materials and Methods

2.1. In silico work

A set of online tools, softwares and databases were used in this study, which are all listed in Table 2.1.

2.1.1. Databases

NCBI's PubMed and Google Scholar were the main search engines used for the literature review. TriTrypDB, GeneDB and NCBI were the main databases searched for DNA sequences.

2.1.2. Primer design

All primers were designed with the Primer3plus web program (*351*, *352*) with the exception of the quantitative Real Time – PCR (qPCR) primers, which were designed with the Primer Express software (Applied Biosystems [AB]). All primers were purchased from Sigma-Aldrich.

2.1.3. The Basic Local Alignment Search Tool (Blast)

The BLASTn tool on the NCBI and TriTrypDB homepages was used to identify sequence identities and similarities in the published genomes of *Leishmania spp.* (*353*). BLAST exploits a substitution matrix to search for sequences of a specific length that contain one or several high-scoring sequence pairs (HSPs), which are in turn calculated by a scoring matrix, that finds homologues in the query sequence to score above a threshold value, which is dependent on the speed and sensitivity of the search. From these HSPs, the program extends the sequence in either direction to achieve an alignment of the query sequence to the found sequences, which need to exceed another threshold to be considered.

2.1.4. CLUSTAL Sequence alignments

Multiple DNA and protein sequence alignments were performed using CLUSTALW2 - Multiple Sequence Alignment program hosted by EBI (*354*, *355*). CLUSTAL differs from BLAST by not searching a database, but by merely aligning two or several given sequences to a best fit. For that submitted sequences are pairwise aligned and stored in a distance matrix, for which a phylogenetic tree is created based on a neighbour-joining clustering algorithm that successively aligns sequences starting from the most closely related sequences.

Table 2.1 - Bioinformatics tools used in this study

Site	URL	Used for
BLASTn (NCBI)	http://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastn&BLAST_PROGRAMS=megaBlast&PAGE_TYPE=BlastSearch&SHOW_DEFAULTS=on&LINK_LOC=blasthome	Basic Local Alignment Search Tool (BLAST)
Calculator for determining the number of copies of a template	http://www.uri.edu/research/gsc/resources/cndna.html	Determining the number of copies of a template
CAP3 Sequence Assembly Program	http://pbil.univ-lyon1.fr/cap3.php	Sequence assembly
ClustalW2	http://www.ebi.ac.uk/Tools/msa/clustalw2/	Nucleic Acid sequence alignment
Fragment Size Calculator	http://www.basic.northwestern.edu/biotools/SizeCalc.html	DNA fragment size calculation on Southern blots
GeneDB	http://www.genedb.org/Homepage	Genomic sequence searches
Google Scholar	http://scholar.google.co.uk/	Literature searches
Nucleic Acid Sequence Massager	http://www.attotron.com/cybertory/analysis/seqMassager.htm	Manipulation of nucleic acid sequences
Molarity Calculator	http://www.graphpad.com/quickcalcs/molarityform.cfm	Conversion and calculation of molarity
ORF Finder	http://www.bioinformatics.org/sms2/orf_find.html	Open Reading Frame identification
Primer3Plus	http://www.bioinformatics.nl/cgi-bin/primer3plus/primer3plus.cgi	Primer design
PubMed	http://www.ncbi.nlm.nih.gov/pubmed	Literature searches
TheLabRat	http://www.thelabrat.com/restriction/index.shtml	Source of restriction sites and enzymes
Transcription and Translation Tool	http://www.attotron.com/cybertory/analysis/trans.htm	Conversion of DNA sequences to mRNA sequences to protein sequences and vice versa
TriTryp	http://tritrypdb.org/tritrypdb/	Genomic sequence, homologous and other searches
Uniprot	http://www.uniprot.org/taxonomy/	Taxonomy searches
Webcutter	http://rna.lundberg.gu.se/cutter2/	Identification of restriction sites
WHO	http://www.who.int/en/	Source of information

2.1.5. Sequencing Data Analysis

DNA sequencing data were analysed employing Sequence Scanner software v1.0 (AB) and to convert them into fasta format files, which were submitted to the CAP3 Sequence Assembly Program hosted by PBIL, France (*356*) for assembly into larger contigs. These contigs were then analysed either in CLUSTALW2 or BLASTn.

2.1.6. Restriction Site Determination Tools

Vector NTI was used to visualize DNA sequences and to plan plasmid assemblies and restriction digests (Version 11 - Invitrogen) (*357*). Webcutter (version 2.0) was also used to analyse DNA sequences for restriction sites (Heiman, 1997) (*358*).

2.1.7. Other Computer Softwares

ZEN light (Zeiss) was used for confocal microscopy and image analysis; ImageJ (*359*) for Giemsa stained parasite image analysis; StepOne™ Software (Version 2.2.2) (AP) for qPCR analysis; FlowJo (Tree Star, Inc.) software for flowcytometer data analysis; SPSS v.19 & v.20 (IBM) for statistical data analysis; SigmaPlot (Systat Software Inc.) and Prism (GraphPad Software) for graph generation; Mendeley (Mendeley Ltd.) for citations (*360*); Sequence Massager (supported by cybertory.org) for sequence manipulation.

2.2. *Leishmania* manipulation

2.2.1. *Leishmania* species and strains used

All *Leishmania* strains used in this study are listed in Table 2.2. *L. (L.) major* Friedlin VI (FVI) (MHOM/IL/81/Friedlin/VI) was used as a wild type and is the parental line to all mutant lines generated and used in this study. All cDNA16 locus gene replacements were done by homologous recombination into the former cDNA16 locus in the *L. (L.) major* 4.8 *ΔcDNA16* double deletion mutant (*Lmj*cDNA16 dKO) (*ΔcDNA16::HYG/ΔcDNA16::PAC*), generated from FVI previously (*332*).

L. (V.) braziliensis 2904 (*Lbr*2904) (*MHOM/BR/75/M2904*) was used as a wild type for the study of the orthologous HASP locus (OHL) and served as a parental line for the attempted full OHL deletion. *L. (V.) braziliensis* LTB300 (*MHOM/BR/83/LTB300*) and gDNA from several clinical *L. (V.) braziliensis*

Table 2.2 – *Leishmania* mutant strains

Species	Strain	Type	Source	Mutation	Cultured	gDNA
Leishmania (Leishmania) major	MHOM/IL/81/Friedlin/VI	Reference strain	Smith lab cryobank		√	√
L. (L.) major	cDNA16 dKO	Mutant line		*ΔcDNA16::HYG/ΔcDNA16::PAC*	√	√
L. (L.) major	cDNA16 sKI	Mutant line		*ΔcDNA16::HYG/ΔcDNA16::PAC/ΔPAC::cDNA16+NEO*	√	√
L. (L.) major	HASPB sKI	Mutant line	Generated in this study	*ΔcDNA16::HYG/ΔcDNA16::PAC/ΔHYG::HASPB+NEO (or ΔPAC::HASPB+NEO)*	√	√
L. (L.) major	SHERP sKI	Mutant line		*ΔcDNA16::HYG/ΔcDNA16::PAC/ΔHYG::SHERP2+NEO (or ΔPAC::HASPB+NEO)*	√	√
L. (L.) major	S2+HB sKI	Mutant line		*ΔcDNA16::HYG/ΔcDNA16::PAC/ΔHYG::SHERP2+NEO/ΔPAC::HASPB+BSD*	√	√
L. (L.) major	S2/HB sKI	Mutant line		*ΔcDNA16::HYG/ΔcDNA16::PAC/ΔPAC::SHERP2-HASPB+BSD* (or *ΔHYG::SHERP2-HASPB+BSD)*	√	√
L. (L.) major	HASPB dKI	Mutant line		*ΔcDNA16::HYG/ΔcDNA16::PAC/ΔHYG::HASPB+NEO/ ΔPAC::HASPB+BSD*	√	√
L. (L.) major	HASPA1 sKI	Mutant line		*ΔcDNA16::HYG/ΔcDNA16::PAC/ΔPAC::HASPA1+BSD (or ΔHYG::HASPA1+BSD)*	√	√
L. (L.) major	HASPA2 sKI	Mutant line		*ΔcDNA16::HYG/ΔcDNA16::PAC/ΔPAC::HASPA2+NEO (or ΔHYG::HASPA2+NEO)*	√	√
L. (L.) major	HASPA1/2 sKI	Mutant line		*ΔcDNA16::HYG/ΔcDNA16::PAC/ΔPAC::HASPA1-HASPA2+NEO (or ΔHYG::HASPA1-HASPA2+NEO)*	√	√
L. (L.) major	HA1+HB sKI	Mutant line		*ΔcDNA16::HYG/ΔcDNA16::PAC/ΔHYG::HASPB+NEO/ ΔPAC::HASPA1+BSD*	√	√
L. (L.) major	HA2+HB sKI	Mutant line		*ΔcDNA16::HYG/ΔcDNA16::PAC/ΔPAC::HASPA2+NEO/ΔHYG::HASPB+BSD*	√	√
L. (L.) major	HA1/2+HB sKI	Mutant line		*ΔcDNA16::HYG/ΔcDNA16::PAC/ΔPAC::HASPA1-HASPA2+NEO/ΔHYG::HASPB+BSD*	√	√
L. (L.) major	HA1+S2 sKI	Mutant line		*ΔcDNA16::HYG/ΔcDNA16::PAC/ΔHYG::SHERP2+NEO/ΔPAC::HASPA1+BSD*	√	√
L. (L.) major	HA2+S2 sKI	Mutant line		*ΔcDNA16::HYG/ΔcDNA16::PAC/ΔPAC::HASPA2+NEO/ ΔHYG::SHERP2+BSD*	√	√
L. (L.) major	HA1/2+ S2 sKI	Mutant line		*ΔcDNA16::HYG/ΔcDNA16::PAC/ΔPAC::HASPA1-HASPA2+NEO/ΔHYG::SHERP+BSD*	√	√
L. (L.) major	HA1+S2/HB sKI	Mutant line		*ΔcDNA16::HYG/ΔcDNA16::PAC/ ΔPAC::HASPA1+BSD/ ΔHYG:: SHERP2-HASPB+SAT*	√	√

Species	Strain	Type	Source	Mutation	Cultured	gDNA
L. (L.) major	HA2+ S2/HB sKI	Mutant line	Generated in this study	*ΔcDNA16::HYG/ΔcDNA16::PAC/ΔPAC::HASPA2+NEO/ ΔHYG:: SHERP2-HASPB+BSD*	√	√
L. (L.) major	HA1/2+ S2/HB sKI	Mutant line		*ΔcDNA16::HYG/ΔcDNA16::PAC/ΔPAC::HASPA1-HASPA2+NEO/ΔHYG:: SHERP2-HASPB+BSD*	√	√
Leishmania (Viannia) braziliensis	MHOM/BR/75/M2904	Reference strain	Smith lab cryobank		√	√
L. (V.) braziliensis	OHL sKO	Mutant line	Generated in this study	*ΔOHL::BSD/OHL*	√	√
L. (V.) braziliensis	MHOM/BR/84/LTB300	Clinical isolate	Smith lab cryobank		√	√
L. (V.) braziliensis	MHOM/BR/2004/EGS	Clinical isolate	Dr. S. Uliana USP, Brazil	Clinical isolate (*Lbr* 1)		√
L. (V.) braziliensis	MHOM/BR/2006/GDL	Clinical isolate		Clinical isolate (*Lbr* 2)		√
L. (V.) braziliensis	MHOM/BR/2006/HPV	Clinical isolate		Clinical isolate (*Lbr* 3)		√
L. (V.) braziliensis	MHOM/BR/2003/IMG	Clinical isolate		Clinical isolate (*Lbr* 4)		√
L. (V.) braziliensis	MHOM/BR/2006/PPS	Clinical isolate		Clinical isolate (*Lbr* 5)		√
L. (V.) braziliensis	MHOM/BR/2006/BES	Clinical isolate		Clinical isolate (*Lbr* 7)		√
L. (V.) braziliensis	MHOM/BR/2005/RPL	Clinical isolate		Clinical isolate (*Lbr* 8)		√
L. (V.) braziliensis	MHOM/BR/2006/UAF	Clinical isolate		Clinical isolate (*Lbr* 9)		√
L. (V.) braziliensis	MHOM/BR/2005/WSS	Clinical isolate		Clinical isolate (*Lbr* 10)		√
L. (V.) braziliensis	MHOM/BR/2006/EFSF	Clinical isolate		Clinical isolate (*Lbr* 11)		√

isolates (Table 2.2) provided by the Universidade de São Paulo were used, too. The OHL single and double deletion mutants were generated by homologous recombination in *Lbr*2904.

All newly generated mutants were passaged through BALB/c mice to restore their infectivity (see 2.2.5) before sand fly infections were undertaken.

2.2.2. Culture media and culture conditions

All *Leishmania* strains were routinely cultured in 1x medium 199 (M199) supplemented with 10% heat inactivated Foetal Calf Serum (FCS; Gibco) and penicillin-streptomycin (*332*). Medium was filter-sterilized and stored at 4 – 8 °C. FVI, *Lmj*cDNA16 dKO, *Lbr*2904 and *Lbr*300LBT were grown in M199 without antibiotics, *Lmj*cDNA16 sKI, *Lmj*HASPB sKI, *Lmj*SHERP2 sKI, *Lmj*HASPA2 sKI, *Lmj*HASPA1/2 sKI were grown in M199 + neomycin (NEO) (40 µg/ml), *Lmj*S2/HB sKI, *Lmj*HASPA1 sKI and *Lbr*OHL sKO M199 + blasticidin (BSD) (10 µg/ml), *Lmj*S2+HB sKI, *Lmj*HASPB dKI, *Lmj*HA1+HB sKI, *Lmj*HA1/2+HB sKI, *Lmj*HA2+HB sKI, *Lmj*HA1+S2 sKI, *Lmj*HA1/2+S2 sKI, *Lmj*HA2+S2 sKI, *Lmj*HA1/2+S2/HB sKI and *Lmj*HA2+S2/HB sKI in M199 + NEO (40 µg/ml) + BSD (10 µg/ml) and LmjHA1+S2/HB sKI in M199 + BSD (10 µg/ml) + streptophricin (SAT) (100 µg/ml). All strains were maintained either at 26°C in M199 cultures or at 23°C in M199 on biphasic rabbit blood-agar slopes (*329*).

2.2.3. Splitting and passaging Leishmania parasites *in vitro*

For M199 inoculation from cryo-samples, the cryo-samples were thawed quickly, checked by haemocytometer for vitality and either poured onto blood slopes and kept at 23°C for 24h prior to inoculation into M199 cultures or 500 µl of the cryo-sample were immediately inoculated into 10 ml M199 and kept at 26 °C.

For M199 inoculation with amastigotes, draining lymph nodes were removed from infected BALB/c mice and ground up in a sterile sieve or petri dish with sterile syringe plungers to release amastigotes, which were washed with pre-warmed (26°C) M199 into 10 ml culture flasks. The cultures were immediately used for a 1:10 dilution into 10 ml M199 as back-ups. Cultures were maintained at 26°C until motile promastigotes appeared.

All cultures were passaged at day 2-4 with 2-3 drops into fresh 10 ml M199

and kept at 26°C.

Parasites passaged multiple times (>10x) under axenic conditions become avirulent (*361*). To restore virulence *L. (L.) major* strains were passaged through BALB/c mice (see 2.2.5). No culture was passaged more than 10 times *in vitro* in this study.

2.2.4. Cryo-samples

850 µl of cultured parasites in early to mid-log phase (day 2-3 culture) were aliquoted with 136 µL 50% glycerol (filter-sterilized) (6.25:1) into labelled 1ml cryo-vials (NUNC) and placed into an isopropanol-filled double walled container (Mr. Frosty) for slow freezing at -80 °C for 24 – 48 h, before transferring samples into a liquid nitrogen tank for long term storage. For short term storage, cryo samples were kept at -80 °C.

2.2.5. Artificial mouse infection with *L. (L.) major*

2.2.5.1. *L. (L.) major* passage through BALB/c mice

Parasite passage through mice has been described elsewhere (*293*). Briefly, female BALB/c mice (6-8 weeks of age) were purchased from Harlan and kept and treated according to ethical standards enforced by the UK Home Office. For mouse infections, $3x10^8$ stationary-phase (day 6-7 culture) parasites were spun down for 10 min. at 3,200 rpm (~2,200 x *g*) in a Sorvall Legend RT centrifuge and washed once in 10 ml phosphate buffered saline solution (PBS) and once in 1 ml sterile PBS for 5 min. at 4,600 rpm (~2,000 x *g*) in a Sorvall Pico centrifuge. Washed parasites were re-suspended in PBS to a final concentration of $3x10^7$ parasites/30 µl. $3x10^7$ parasites (30 µl) were subcutaneously injected into the right footpad of female BALB/c mice, which were sacrificed after severe footpad swelling occurred. The right draining popliteal lymph node was dissected and amastigotes were harvested as described above (see 2.2.3).

2.2.5.2. Amastigote generation and isolation

The protocol was adapted from Paape *et al.* (2011) (*362*). BALB/c mice were infected with 10^7 stationary phase *L. (L.) major* cells from day 6 post inoculum (p.i.) cultures in 30 µl PBS by needle inoculation on both sides of the base of the tail. After lesions had developed (8-10 weeks p.i.), mice were sacrificed according to UK Home Office guidelines and lesion material

was excised with a scalpel, weighed and force through a 70 µm cell strainer into homogenization buffer (20 mM HEPES-KOH, pH 7.3, 0.25 M sucrose supplemented with cOmplete Mini proteinase inhibitor cocktail [Roche]). The suspension was centrifuged at 2,200x*g* for 10 min. and the cell pellets were suspended in 1 mL homogenization buffer. Amastigotes were released from amastigotes by forcing the cell suspension through a 25-gauge needle. Nuclei were removed by centrifugation at 100x*g* for 2 min. The supernatants were loaded onto a discontinuous sucrose gradient: 1 mL each of 20, 40, and 60% (w/w) sucrose in HEPES saline (30 mM HEPES-KOH, pH 7.3, 0.1 M NaCl, 0.5 mM $CaCl_2$, 0.5 mM $MgCl_2$) (*363*), centrifuged for 25 min. at 700x*g*. Amastigotes were isolated from the 40/60% sucrose interface, diluted in PBS and centrifuged for 10 min. at 2,200x*g*. Cells were suspended in an appropriate amount of PBS and Laemmli buffer (20 ml 0.5 M Tris-HCl [pH 6.8], 3.08 mg DTT, 40 ml SDS [10%], 50 mg Bromophenol Blue, 20 ml Glycerol [100%] and sterile Milli-Q water (MQH_2O) to 100 ml) and boiled for 10 min. at 95 °C before freezing the samples at -20 °C.

2.2.6. *Leishmania* homologous recombination mutant generation

2.2.6.1. Transfection

Parasite transfection using a Human T-Cell Nucleofection™ kit (Amaxa) was described previously (*364*). Briefly, recombinant DNA plasmids were amplified in *Escherichia coli*, the DNA was extracted as midi-preps (Qiagen) according to the supplier's protocol, restriction digested, the DNA was gel purified using a Qiagen kit according to the supplier's protocol, ethanol precipitated and suspended in sterile MQ H_2O to a final concentration of ~1 µg / µL. ~2 x 10^7 log-phase parasite cells were spun down in one falcon tube per transfection at 3,200 rpm (2,200x*g*) for 10 min. at room temperature in a Sorvall Legend RT centrifuge. The supernatant was removed and cells were washed once in 10 ml PBS and once in 1 ml PBS in Eppendorf tubes. Cell pellets were suspended in 100 µL Nucleofection™ solution (AMAXA) and 5 µl DNA (4-5 µg) were immediately added. Samples were then transferred into AMAXA cuvettes, which were placed into the Nucleotranfector™ II machine (AMAXA) and the program U-033 was applied for electroporation. Transfected cells were transferred into 10 ml pre-warmed M199 medium and incubated at 26 °C overnight.

2.2.6.2. *Leishmania* clone selection

The following morning, antibiotics for selection were added as required to the overnight cultures and the cultures were incubated for another 2 – 3 h at 26 ºC. Transfected parasites were spun down at 3,200 rpm (2,200x*g*) for 10 min. at 15 °C and the supernatant was removed. Cell pellets were suspended in the residual M199 medium (100 – 200 µl) in the tube and were then spread on M199-agar plates, which had been poured by mixing 25 ml of pre-warmed (37 °C) 2x M199 containing Biopterin (1.2 µl / ml) and the antibiotics for selection as required (control plates did not contain antibiotics) with 25 ml of 2% Agar kept at 56 °C. The cells were spread with sterile spreaders (not until dry) and the plates were left to dry at the surface just enough that the parasites could not swim freely, but the flagella could move. The plates were sealed and incubated at 26 °C for 10 – 21 days and checked regularly for colony growth.

2.2.6.3. Growing up parasite clones

Antibiotic resistant *Leishmania* colonies were observed as early as 7 days of incubation with *L. (L.) major*. They were picked and used to inoculate 100 µL 1x M199 medium + respective antibiotics on 96-well plates. Plates were sealed with parafilm and incubated at 26 °C until dense parasite growth was observed (usually 3 – 4 days). Cultures were diluted 1:2 with 100 µL fresh 1x M199 medium + respective antibiotics and 100 µL of culture were transferred into 1.5 ml fresh 1x M199 medium + respective antibiotics on 24-well plates, which were sealed and incubated at 26 °C until dense parasite growth was observed (usually 2-3 days). 500 µL were transferred into 5 ml fresh 1x M199 medium + respective antibiotics on 6-well plates, which were sealed and incubated at 26 °C until dense parasite growth was observed (usually 2-3 days). These cultures were then used to inoculate new cultures in flasks, to produce cryo-samples and for gDNA extraction for parasite clone screens.

2.2.7. Parasite measurements

Giemsa stained parasites from sand fly gut smears were measured with respect to flagellum length, cell body length and cell body width on microscope images using the software Image J. For each strain 60 images from parasites derived from the AMG and TMG, respectively, were taken per dissection day (day 5/6, 9 & 12 PBM) per glass slide. Three glass slides were imaged per condition totalling 180 measured parasites per midgut section per

dissection day. Four morphological forms were distinguished according to Walters *et al.* (1993) (*365*) and Cihakova & Volf (1997) (*157*): (i) procyclics: body width >4 mm and body length <7.5 mm; (ii) nectomonads: body length ≥14 mm; (iii) Leptomonads: body length < 14 mm and flagellar length < 2 times body length; (iv) metacyclic promastigotes: body length < 14 mm and flagellar length ≥2 times body length. Haptomonads could not be distinguished from Leptomonads by measurement.

2.2.8. Growth Assay

Parasites were inoculated into 10 ml 1x M199 to a final concentration of ~$5x10^5$ parasites / ml and incubated at 26 °C. Starting right after inoculation, 10 µl of culture were extracted every 24 h, mixed with 490 µl 1% Formaldehyde in saline solution for day 0, 1 & 2 p.i. or 990 µl 1% Formaldehyde in saline solution for day 3, 4, 5, 6 & 7 p.i. and parasites were counted on a haemocytometer. Concentrations per ml were calculated and plotted on a log-scale scatter graph.

2.2.9. Osmotaxis Assay

The osmotaxis assay was described previously (*366*, *367*). Briefly, plain glass capillary tubes (75 mm length, 0.8 inner / 1 mm outer diameter) were filled with wash and incubation (WIS) buffer (30 mM *β*-glycerophosphate disodium salt, 87 mM NaCl, 27 mM KCl, 2 mM $CaCl_2$, 2 mM $MgCl_2$, 0.004% enriched Bovine Serum Albumin [pH 7.1]) containing 1% agarose and 100 mM of sucrose, or not, leaving exactly 1 cm void (~5 µl). Once the agarose had settled, the remaining void was filled with WIS buffer and the filled glass capillary tubes were submerged horizontally in WIS buffer in a Petri dish, which was left for ~30 min. at room temperature on a rocking table to establish a sucrose gradient (control tubes were incubated on a separate Petri dish, too).

Parasites were grown to late log-phase / early stationary-phase in 1x M199 or 5% sucrose/PBS. They were harvested by centrifugation and washing the cell pellet twice in WIS buffer before suspending the cell pellet to a final concentration of ~$2.5x10^7$ cells / ml in WIS buffer. The cell suspension was transferred into a 7 ml universal tube.

The equilibrated glass capillary tubes were submerged into the parasite suspension at a slight angle with the WIS buffer filled end. 6 capillary tubes

with 100 mM sucrose and 6 without sucrose were used per strain. The tubes without sucrose were used as a negative control to normalize the data. The universals were sealed with parafilm and incubated at 26 °C for 1 h. The WIS buffer in the filled void of the capillary tubes (~5 µl) was removed, mixed with 195 µl 1% Formaldehyde in saline solution (1:40 dilution) for sample from tubes with attractant (+ve) or 95 µl 1% Formaldehyde in saline solution (1:20 dilution) for sample from control tubes without attractant (-ve) in 0.5 ml Eppendorf tubes and the parasites were counted on a haemocytometer.

The assay was repeated 3 times on different days and the attraction coefficient (AC) calculated in two distinct ways for each repeat:

1. $$AC = \frac{\left(\frac{a_1}{b_1}+\frac{a_2}{b_2}+\frac{a_3}{b_3}+\frac{a_4}{b_4}+\frac{a_5}{b_5}+\frac{a_6}{b_6}\right)}{6}$$

$\{\mathrm{a} = \text{parasite \# of postive samples};\ \mathrm{b} = \text{parasite \# of negative samples}\}$

2. $$AC = \frac{a_1+a_2+a_3+a_4+a_5+a_6}{b_1+b_2+b_3+b_4+b_5+b_6}$$

$\{\mathrm{a} = \text{parasite \# of postive samples};\ \mathrm{b} = \text{parasite \# of negative samples}\}$

The results were plotted and compared. The standard deviation for each strain and P-values were calculated.

2.3. DNA Manipulation Protocols

2.3.1. DNA sequencing and processing

All DNA sequencing was performed by technical staff in the Genomics lab of the Technology Facility (TF) of the University of York using an Applied Biosystems 3130 sequencer. Plasmid samples were submitted at 100-150 ng/µl concentrations and primers at 3.2 µM. A complete list of primers used in this study can be found in Appendices 2 – 6.

2.3.2. Genomic DNA extraction

Two different procedures were used for gDNA extraction depending on the quantity and quality of gDNA needed.

2.3.2.1. Phenol/chloroform gDNA extraction

Parasites were grown in 50 ml cultures to a density of 2-3x10^7 cells/ml and

spun down at 3,200 rpm (2,200x*g*) for 10 min. at 15 °C in a Sorvall Legend RT centrifuge. Cell pellets were washed twice in 40 ml sterile PBS and spun down as above before re-suspending cells thoroughly in 9 ml Net buffer (1 ml 0.5M Tris pH 8.0, 10 ml 0.5M EDTA, 1 ml 5M NaCl, 38 ml sterile MQ H_2O). 1 ml of 10% SDS was added and samples were mixed by inversions before incubation with 40 µl RNaseA (10 mg/ml) for 30 min. at 37 °C. 200 µL Proteinase K (600 mAU/ml) were added and samples were incubated at 55 °C overnight.

Samples were retrieved the following day and 1 volume Phenol/Chloroform (1:1) (Sigma/Fisher Scientific) was added, mixed by inversion and spun down at 3,200 rpm (~2,200x*g*) for 20 min. at room temperature in a Sorvall Legend RT centrifuge. The aqueous layer was transferred into a fresh tube and the steps repeated. 1 volume Chloroform was added to the aqueous layer and spun down as above. The transferred aqueous layer was ethanol precipitated (see 2.3.8.), the gDNA pellet dried and re-suspended in MQ H_2O for immediate use or 1x TE buffer for storage at 4 – 8 °C.

2.3.2.2. Genomic DNA extraction by blood and tissue kit (Qiagen)

Smaller amounts of gDNA were extracted by Blood and Tissue kit (Qiagen) according to the supplier's protocol. Genomic DNA extracts were eluted in 200 µl AE buffer (Qiagen). If concentration of gDNA was required, an ethanol precipitation step was used and the gDNA pellet was either re-suspended in MQ H_2O for immediate use or in 1x TE buffer for storage at 4 – 8 °C.

2.3.3. PCR amplifications

Examples of all PCR profiles are listed in **Table 2.3**.

2.3.3.1. Conventional PCR

For conventional PCR reactions, GO-Taq® polymerase (Promega) or Taq polymerase (NEB) were used according to the supplier's protocol. The high fidelity KOD hot start DNA polymerase (Novagen) was used for DNA amplifications for cloning. Reactions were set up and run according to the supplier's protocol (Table 2.3A&B).

2.3.3.2. Long range PCR

KOD XL DNA polymerase (Novagen), KAPA long range hot start DNA

Table 2.3 – PCR profile

A) Taq (NEB) and GO-Taq polymerase (Promega**)**

Stage	Temperature	Cycle #	Duration
Initial denaturing step	95 °C		3 min
Denaturing	95 °C	30 – 35 x	30 – 60 sec
Annealing	55 – 65 °C		30 – 60 sec
Extension	72 °C		1 min / Kb
Final Extension	72 °C		5 min
Final step	4 °C		∞

B) KOD hot start DNA polymerase (Inovagen**)**

Stage	Temperature	Cycle #	Duration
Initial denaturing step	95 °C		2 min
Denaturing	95 °C	30 – 35 x	30 – 60 sec
Annealing	55 – 65 °C		30 – 60 sec
Extension	72 °C		1 min / Kb
Final Extension	72 °C		5 min
Final step	4 °C		∞

C) KOD XL DNA polymerase (>12 Kb) (Inovagen**)**

Stage	Temperature	Cycle #	Duration
Initial denaturing step	94 °C		3 min
Denaturing	94 °C	25 – 30 x	30 sec
Annealing	55 – 65 °C		5 sec
Extension	70 – 74 °C		8-10 min / Kb
Final Extension	74 °C		10 min
Final step	4 °C		∞

D) KAPA long range hot start DNA pol. (>15 Kb) (KAPAbiosystems**)**

Stage	Temperature	Cycle #	Duration
Initial denaturing step	94 °C		2 min
Denaturing	95 °C	10 x	25 sec
Annealing	55 – 65 °C		15 sec
Extension	68 °C		1 min / Kb
Denaturing	95 °C	25 x	25 sec
Annealing	55 – 65 °C		15 sec
Extension	68 °C		1 min / Kb + 20 sec / cycle
Final Extension	72 °C		1 min / Kb
Final step	4 °C		∞

E) DyNAzyme™ EXT DNA Polymerase

Stage	Temperature	Cycle #	Duration
Initial denaturing step	94 ºC		2 min
Denaturing	94 ºC	10x	30 sec
Annealing	55 ºC		30 sec
Extension	70 ºC		40 sec / Kb
Denaturing	94 ºC	20x	30 sec
Annealing	55 ºC		30 sec
Extension	70 ºC		40 sec / Kb + 20 sec / cycle
Final Extension	70 ºC		5 – 10 min
Final step	4 ºC		∞

F) qPCR with Power SYBR® Green PCR master mix (AB)

Stage	Temperature	Cycle #	Duration
Initial denaturing step	95 ºC		20 sec
<u>Extension/Detection</u>:			
Denaturing	95 ºC	40x	3 sec
Annealing/Extension	60 ºC		30 sec
<u>Melting Curve</u>:			
Denaturing	95 ºC	1x	15 sec
Annealing/Extension	60 ºC		1 min
Denaturing	95 ºC		15 sec

polymerase (KAPA biosystems) and DyNAzyme™ EXT DNA Polymerase were used for high fidelity long range PCRs according to the supplier's protocols (Table 2.3C, D & E).

2.3.3.3. Reverse transcriptase – PCR

Prior to the reverse transcriptase-PCR (RT-PCR), RNA was treated for removal of DNA with the DNA-free kit (Ambion) according to the supplier's protocol. Briefly, 4 µg of RNA were added to 5 µL 10x rDNase I buffer and 0.5 µL rDNase I (2 units/µL) in a 50 µL reaction and left to incubate for 15 min. at 37 °C. 5 µL rDNase I inactivation buffer were added and left to incubate for 2 min. at room temperature to stop the reaction. The reaction mix was then spun down at 10,000 rpm (~9,500x*g*) for 1 min. in a Sorvall Pico centrifuge and the liquid phase was transferred to a fresh tube. RT-PCR was done with the Omniscript kit (Qiagen) according to the supplier's recommendations. Briefly, 10 µL DNase-treated RNA, 2 µL RNase-free water, 2 µL 10x Omniscript buffer, 2 µL dNTPs (5 mM), 2 µL oligo dT primers (10 µM), 1 µL RNase inhibitor (10 units/µL), 1 µL Omniscript reverse transcriptase were added to a single Eppendorf tube and incubated at 37 °C for 1 h. The sample was then either used in a qPCR or frozen at -20 °C.

RT-PCR was also performed with the Transcriptor First Strand cDNA Synthesis Kit (Roche), which later substituted the Omniscript kit, using random primers in 20 µl reactions according to the supplier's protocol. The reaction was performed at room temperature for 10 min. followed by incubation at 55 °C for 30 min. Samples were stored at -20 °C.

2.3.3.4. Quantitative Real Time – PCR

DNA template for the qPCR reactions was either complementary DNA (cDNA) from DNase-treated RNA or RNaseA treated gDNA. The starting material was prepared in serial-dilution to have between 1,000,000 – 10 molecules per µl. Molecule numbers per µl sample were calculated according to the following formula:

1) $$\#\ \text{of molecules}/_{\mu g} = \frac{1000\ \text{Kb}}{\text{Kb per molecule}}\ \text{x}\ 9.1\ \text{x}\ 10^{11}\ \ \text{molecules}/_{\mu g}$$

2) $$\#\ \text{of molecules}/_{\mu l} = \#\ \text{of molecules}/_{\mu g}\ \text{x sample concentration}\ (^{\mu g}/_{\mu l})$$

For the qPCR Power SYBR® Green PCR Master Mix (AB) was used and primers were diluted to a final concentration of 300 nM per reaction. Reactions were prepared in 25 µl aliquots in triplicate in optical 96-well plates (AB). The loaded plate was sealed and spun for 2 min. at 4,000 rpm (~2,100x*g*) in a Hettich Universal 32-A centrifuge to ensure that all liquid was at the bottom of the wells. The plates were then run in a Prism7000 machine (AB) (Table 2.3F) and the results analysed in Prism7000 system software (AB).

2.3.4. PCR product purification

PCR products were purified using a PCR purification kit (Qiagen) according to the supplier's protocol. PCR products were eluted in 30 – 50 µl elution buffer. Purified PCR products were either used right away or were frozen at -20 ºC in elution buffer.

2.3.5. Plasmid construction

A list of all plasmids constructs generated in this study can be found in Table 2.4. DNA constructs for homologous recombination in *Leishmania* were built in pCR®2.1-TOPO® vector plasmids (Invitrogen) by a multi-step protocol that incorporated the following steps:

2.3.5.1. 3' A-overhang addition

Proof-reading DNA polymerases, like KOD polymerase, have exonuclease activity and do not leave a 3'-mono-A-overhang like, for instance, conventional Taq-polymerase. The auto-ligating pCR®2.1-TOPO® vector (Invitrogen) exploits the 3'-mono(A)-overhang for initial DNA fragment integration. Since proof-reading DNA polymerase were used for DNA fragment generation for homologous recombination constructs, it was necessary to add 3'-mono-A-overhang after the PCR. Briefly, 0.7-1 unit of Taq polymerase (NED) was added to purify PCR products together with dATPs (final concentration 200 nM) and 10 x Taq polymerase buffer according to supplier's recommendations. Samples were incubated at 72°C for 8-10 min. and then transferred to ice. The product was immediately used for integration into the pCR®2.1-TOPO® vector (Invitrogen) according to supplier's protocol.

2.3.5.2. Restriction digests protocols

All restriction enzymes used in this study are listed in Table 2.5 and were

Table 2.4 – Plasmids

Plasmid Name	Vectors used	Inserted Gene	Antibiotic Marker	Restriction Sites	Comment
Leishmania (L.) major constructs:					
pHASPB(I)	pSP6-T3	HASPB	Neomycin (NEO)	*Hind*III / *Xma*I	Homologous recombination; gene replacement
pHASPB(II)	pCR2.1®-TOPO®	HASPB	Blasticidin (BSD)	*Apa*I / *Hind*III	Homologous recombination; gene replacement
pSHERP(I)	pSP6-T3	SHERP2	Neomycin (NEO)	*Hind*III / *Xma*I	Homologous recombination; gene replacement
pSHERP(II)	pCR2.1®-TOPO®	SHERP2	Blasticidin (BSD)	*Apa*I / *Hind*III	Homologous recombination; gene replacement
pS2+HB (I)	pCR2.1®-TOPO®	SHERP2 & HASPB	Blasticidin (BSD)	*Apa*I / *Hind*III	Homologous recombination; gene replacement
pS2+HB (II)	pCR2.1®-TOPO®	SHERP2 & HASPB	Streptothricin (SAT)	*Apa*I / *Hind*III	Homologous recombination; gene replacement
pHASPA1	pCR2.1®-TOPO®	HASPA1	Blasticidin (BSD)	*Apa*I / *Hind*III	Homologous recombination; gene replacement
pHASPA2	pCR2.1®-TOPO®	HASPA2	Neomycin (NEO)	*Apa*I / *Bam*HI	Homologous recombination; gene replacement
pHASPA1/2	pCR2.1®-TOPO®	HASPA1 & 2	Neomycin (NEO)	*Apa*I / *Bam*HI	Homologous recombination; gene replacement
pHASPA-HIS	pET-28a+	HASPA (ORF)			For Protein Production
pHASPB-GFP	pCR2.1®-TOPO® & pcDNA3.1/CT-GFP-TOPO®	HASPB – GFP fusion protein	Streptothricin (SAT)	*Apa*I / *Bgl*II	Homologous recombination; gene replacement
Leishmania (V.) braziliensis constructs:					
pOHL KOI	pCR2.1®-TOPO®	-	Blasticidin (BSD)		Homologous recombination; OHL deletion
pOHL KOII	pCR2.1®-TOPO®	-	Streptothricin (SAT)		Homologous recombination; OHL deletion
pOHL KOIII	pCR2.1®-TOPO®	-	Neomycin (NEO)		Homologous recombination; OHL deletion

either purchased from Promega or New England Biolabs (NEB). Reactions were performed according to the guidelines of the supplier. As a brief example, 2 – 3 µL of respective restriction enzyme (20 – 30 U) were added to a 100 µL reaction containing ~5 µg DNA, 10 µL of (10x) reaction buffer, 1 µL 25 mM BSF and rest MQ H_2O. The reaction was incubated at 37 °C for 3 h. All restriction digests were performed under the same conditions.

2.3.5.3. DNA ligation

Subsequent DNA fragment integration into plasmids required conventional DNA ligations, which were performed according to the supplier's protocol (NEB). Insert and vector DNA were used in a molecular 3:1 ratio. 40 – 50 ng of plasmid vector DNA were used per reaction and the required amount of insert DNA was calculated according to formula:

$$Insert\ DNA\ (ng) = \frac{\text{Insert DNA (Kb)}}{\text{Vector DNA (Kb)}}\ \text{x Vector DNA (ng) x}\ \frac{3}{1}$$

Briefly, a 20 µL reaction contained, apart from the insert and vector DNA, 2 µL T4 ligation buffer (NEB), 1 µL T4 DNA ligase (NEB) and MQ H_2O. The samples were incubated for 10 min. and 60 min. at room temperature for sticky and blunt-end ligation, respectively, and were then heated to 65°C for 10 min. for inactivation.

2.3.5.4. Transformation of chemically competent *E. coli* cells

Transformation of chemically competent *E. coli* DH5α, XL-1 cells or One Shot® TOP 10 cells (Invitrogen) was performed according to the supplier's protocol. Briefly, 2 – 10 µL ligated plasmid DNA were added to 50 µL One Shot® TOP 10 cells or 100 µL DH5α or XL-1 cells previously thawed slowly on ice. Cells were incubated with plasmid DNA for 30 min. on ice and then heat shocked at 42 °C in a heat block for 30 – 45 s for membrane-poration. Tubes were immediately replaced on ice for another 2 min. 250 µL SOC or SOB medium were added to the cells and they were incubated for 1 hour at 37 °C and 225 rpm in an Eppendorf tube placed in a horizontal position. Eventually, cells were spread in either 50 µl or 100 µl aliquots onto pre-warmed Luria-Bertani (LB) agar ampicillin or kanamycin plates (50 ml LB-agar + 50 µl (1000x) ampicillin or kanamycin) with sterile spreaders under sterile condition until dry. The plates were incubated at 37 °C overnight.

Table 2.5 – Restriction Enzymes used in this study

Restriction Enzyme	Supplier	Species in which applied	Reaction Conditions	Restriction Site
ApaI	Promega	*Leishmania (L.) major*	37 °C + BSA	GGGCCˇC
BamHI	Promega	*Leishmania (L.) major*	37 °C + BSA	GˇGATCC
EcoRV	Promega	*Leishmania (L.) major*	37 °C + BSA	GATˇATC
HindIII	Promega	*Leishmania (L.) major, Leishmania (V.) braziliensis*	37 °C + BSA	AˇAGCTT
NotI	Promega	*Leishmania (V.) braziliensis*	37 °C + BSA	GCˇGGCCGC
PstI	Promega	*Leishmania (V.) braziliensis*	37 °C + BSA	CTGCAˇG
PvuI	Promega	*Leishmania (L.) major*	37 °C + BSA	CGATˇCG
PvuII	Promega	*Leishmania (L.) major*	37 °C + BSA	CAGˇCTG
SacI	Promega	*Leishmania (L.) major, Leishmania (V.) braziliensis*	37 °C + BSA	GAGCTˇC
SalI	Promega	*Leishmania (L.) major*	37 °C + BSA	GˇTCGAC
XbaI	Promega	*Leishmania (L.) major*	37 °C + BSA	TˇCTAGA
XhoI	Promega	*Leishmania (L.) major, Leishmania (V.) braziliensis*	37 °C + BSA	CˇTCGAG
XmaI	NE Biolobs	*Leishmania (L.) major*	37 °C + BSA	CˇCCGGG

2.3.5.5. Plasmid extraction from cultured *E. coli* cells

Mini- and MidiPreps (Qiagen) were performed for plasmid isolation from transformed *E. coli* cells grown in LB medium according to the supplier's protocol. Plasmid DNA from MiniPreps was eluted in 30 – 50 µl elution buffer and either used right away or stored at -20°c. Plasmid DNA from MidiPreps was eluted in 5 ml of a special elution buffer (Buffer QF – Qiagen), to which 3.5 ml of isopropanol was added for plasmid DNA precipitation (see supplier's manual).

2.3.6. Gel electrophoresis

Ultrapure agarose (Invitrogen) or SeaKem® Gold agarose (Lonza) was dissolved in 1x Tris acetate EDTA (TAE) buffer to a 0.8% -1.2% or 0.3%-0.5% agarose content, respectively, by heating in a microwave (Panasonic) until the solution was clear. For UV visualization, SYBR® safer (Invitrogen; 1:13333) was added to hand warm liquid agarose. The agarose was poured into trays and left to set for >20 min. at room temperature. The solidified agarose gel was submerged in 1x TAE buffer in an electrophoresis tank (BioRad) and run at 80-100 V for 1-2 h or at 20 V overnight (for Southern blots). Agarose gels for UV visualization in a Syngene G:BOX either contained SYBR® safer or were submerged in Ethidium Bromide solution (1:10,000) for 15 min. on a shaker at room temperature. Gels to be visualized on a light table were submerged in 1x TAE buffer – Methylene Blue (Sigma) solution on a shaker for ~1 h at room temperature or overnight at 4 ºC.

2.3.7. DNA agarose gel extraction

DNA bands were extracted from Methylene blue stained agarose gels using a gel purification kit (Qiagen) according to the supplier's protocol. DNA was eluted by 30 – 50 µl elution buffer.

2.3.8. Ethanol precipitation

Ethanol precipitation has been described elsewhere (*368*). Briefly, sodium acetate, $6H_2O$ (BDH) was added from a 3 M stock (pH5.1) to DNA samples to a final-concentration of 0.3 M. 2 volumes of ice-cold 100% ethanol (Fisher Scientific) were added and mixed in by inverting the tube, and the samples left on ice for 20-30 min., prior to centrifugation in an Eppendorf 5415 R centrifuge at 4 °C for 10 min. at full speed. Supernatants were removed and DNA pellets washed twice in 550 ml 70% ethanol (Fisher Scientific) and spun down at room temperature for 2 min. at full speed in a Sorvall Pico centrifuge. DNA

pellets were dried either in a flowhood at room temperature (sterile) or in a vacuum centrifuge at 30 °C for 5-10 min. (not sterile) and were then re-suspended in sterile MQ H_2O. Genomic DNA suspensions were stored in 1x TE buffer at 4 °C, while DNA fragments and plasmids were frozen at -20 °C.

2.3.9. Southern blot

Genomic DNA (>10 µg) was *Sac*I-digested in 100 ml reactions overnight. The digested gDNA was purified and concentrated by ethanol precipitation and 4-5 µg of digested gDNA were run on a long agarose gel (0.35-0.8% agarose as required) at 20 V overnight (20-24 hours). The agarose gel was submerged in Ethidium Bromide solution (1:10,000) as described for UV light visualization. For DNA degeneration, gels were bathed in 0.5 M NaOH / 1.5 M NaCl solution on a shaker for 30-60 min. The gels were rinsed briefly in distilled H_2O before transfer to neutralising solution (0.5 Tris / 1.5 NaCl, pH 8) for 60 min. on a shaker. The gel was rinsed and mounted for DNA blotting on to an activated positively charged nylon membrane (Roche; activation for 10 min. in 20x SSC [3M NaCl, 0.3M sodium citrate, pH7.2]) (Fig.2.1).

The nylon membrane was briefly rinsed in 2x SSC solution the following day, left to dry on blotting paper and then cross-linked by UV light (160 Joules per cm^2). The cross-linked nylon membrane was then treated with the nonradioactive digoxigenin (DIG) system from Roche according to the supplier's guidelines. Briefly, membranes were pre-hybridised in 20 ml DIG-easy-hyb solution (Roche) at 42 °C in a rolling tube for at least 4 hours. Hybridization with a DIG-labelled probe in DIG-easy hub solution occurred at 42 °C in a rolling tube overnight. All DIG-probes used in this study (5' UTR, HASP, HASPA, SHERP, NEO, BSD, HYG, PAC, BLE, OHL1, OHL2 and OHL3) were generated with the DIG-DNA labelling kit (Roche) according to the supplier's instructions. The following day, the nylon membrane was washed twice for 15 min. in 2x SSC + 0.01% SDS at room temperature and twice for 15 min. in pre-heated 0.5x SSC + 0.01% SDS at 68 °C. The membrane was then rinsed in 1x Washing buffer (Roche), blocked for at least 30 min. with 25 ml blocking buffer (Roche) and for 30 min. with 25 ml blocking buffer containing Anti-digoxigenin antibody conjugated with alkaline phosphatase (AP) (1:12,500) (Roche) on a shaker. The membrane was washed twice for 15 min. in washing buffer on a shaker before applying detection buffer (Roche) for 2-5 min. CDP-star ready-to-use solution (Roche) was applied to the membrane and incubated in the dark for 5 min. Excess

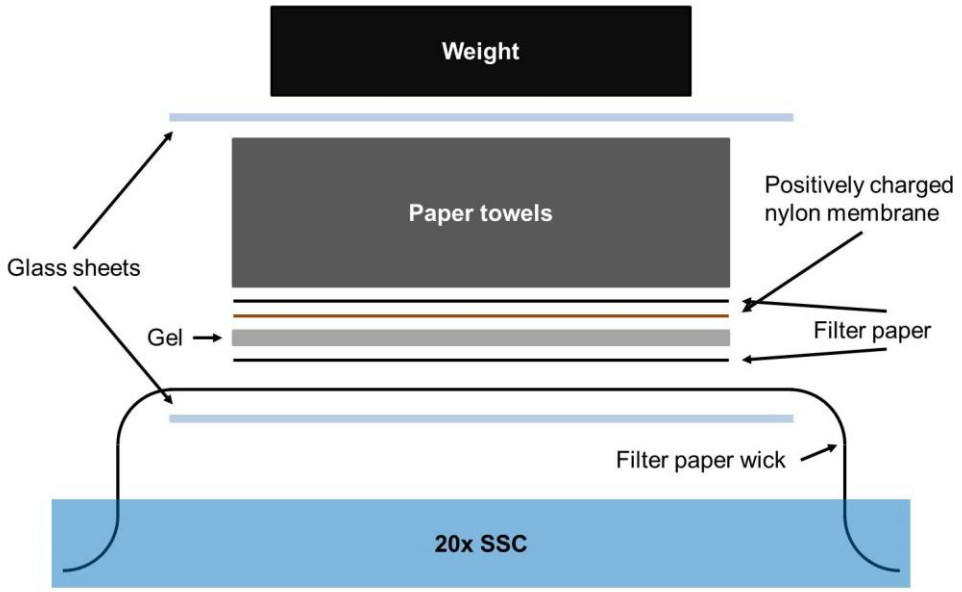


Fig.2.1 – Southern blot set-up

A wick cut from filter paper was placed over a glass slide resting on an open Tupperware box with its end hanging into 20x SSC filling the box. The gel was placed on top of a cut to size piece of filter paper on top of the wick. On top of the gel, a nylon membrane, a second sheet of filter paper cut to size and at least 5 cm of paper towels were stacked finished by a weight of ~600 g. The blot was left overnight for DNA transfer onto the nylon membrane.

CDP-star was removed and the membrane exposed to x-ray films (Amersham) for varying amounts of time. The films were developed in an automated developer (Konica Minolta SRX-101A).

2.3.10. Amplification of genomic DNA extracts

The illustra™ GenomiPhi™ V2 DNA amplification kit was used for amplification of gDNA samples from *L. (V.) braziliensis* clinical isolates for Southern blot analysis according to the supplier's protocol. Briefly, at least 10 ng/µl of template gDNA were mixed with 9 µl sample buffer and heated to 95°C for 3 min. Samples were transferred onto ice before incubation at 30 °C for 1½-2 h with 9 µl reaction buffer and 1 µl enzyme mix. Samples were then heated to 65 °C for 10 min. for enzyme inactivation and transferred back on ice before further use.

2.4. mRNA Manipulation Protocol

2.4.1. mRNA extraction from cultured parasites

1×10^7 parasites were pelleted by centrifugation and were suspended in residual culture before lysis in 1 ml TRIZOL® Reagent by repetitive pipetting. Samples were left for 5 min. to incubate at room temperature before either freezing the sample at -80 °C or immediate extraction. 0.2 ml chloroform were added to samples per 1 ml of TRIZOL® Reagent used. Tubes were shaken vigorously by hand for 15 seconds and then incubated for 2 – 3 min. at room temperature. The samples were centrifuged at ~12,000×*g* for 15 minutes at 4 °C. The aqueous phase was transferred to a fresh tube. RNA was precipitated by mixing 0.5 ml of isopropanol per 1 ml of TRIZOL® Reagent used for the initial lysis to the sample and the mix was incubated at room temperature for 10 minutes followed by centrifugation ~12,000×*g* for 10 min. at 4 °C. The RNA pellet was washed once with 1 ml 75% ethanol per 1 ml of TRIZOL® Reagent used per sample by vortexing followed by centrifuge at ~7,500×*g* for 5 min. at 4 °C. The supernatant was removed and the RNA pellet was air dried for 5-10 min. at room temperature. The RNA was dissolved in RNase-free water and incubated for 10 minutes at 55 – 60 °C.

2.4.2. mRNA extraction from midgut derived parasites

The Magnetic mRNA Isolation Kit (NEB) was used. 20 infected sand fly midguts were dissected into 50 µl IP buffer (150 mM NaCl, 10 mM Tris [pH 7.5], 0.05 mM EDTA, 1 mM DDT in MQ H_2O) with 1µl RNasin (2U; Promega)

in an Eppendorf kept on ice. The samples were homogenized, snap frozen on dry ice with 70% ethanol and thawed on ice. 250 µl Lysis buffer were added prior to DNase treatment (AB) according to the supplier's protocol. mRNA extraction was done with the Magnetic mRNA Isolation Kit (NEB) according to the supplier's protocol. The mRNA was eluted from the magnetic beads with 50 µl elution buffer for 2 min. at 50 °C. Samples were either used immediately for RT-PCR or stored at -80 °C.

2.5. Protein studies

2.5.1. Protein extraction

Leishmania parasites were collected by centrifugation in amounts as required from day 2 to day 7 cultures. Cell pellets were washed once in 1 ml PBS and re-suspended in 25 µL PBS before adding 25 µl Laemmli buffer (20 ml 0.5 M Tris-HCl [pH 6.8], 3.08 mg DTT, 40 ml SDS [10%], 50 mg Bromophenol Blue, 20 ml Glycerol [100%] and sterile MQ H_2O to 100 ml). Samples were heated to 95 °C for at least 10 min. Whole lysates were either directly loaded onto SDS-Polyacrylamide Gel Electrophoresis (PAGE) or stored at -20ºC (*332*).

2.5.2. Western / Immunoblotting

Varying amount of whole cell lysates were separated by SDS-PAGE as described (*332*) and blotted by electroblot onto Immobilon® transfer membranes (Millipore). Membranes were blocked overnight in PBS / 0.05% TWEEN / 5% milk and then probed with polyclonal antisera from rabbits in PBS / 0.05% TWEEN / 5% milk using either HASPA (non-affinity purified anti-HASP ab) and HASPB (non-affinity purified anti-HASP ab or ab336 (*333*)) or SHERP (anti-SHERP ab (*341*)) for 2 h. Membranes were washed 3x for 10 min. in PBS / 0.05% TWEEN and probed with a goat anti-rabbit HRP antibody (Sigma). Membranes were washed again 3x for 10 min. in PBS / 0.05% TWEEN, treated with ECL plus or ECL prime (Amersham) according to the supplier's guidelines and were exposed to x-ray films (Amersham). As a protein loading control, membranes were re-probed with rabbit polyclonal antiserum against *N*-myristoyl transferase (NMT) (*369*).

2.5.3. Ponceau S stain of immunoblot membranes

For the visualization of protein bands on an immunoblot, the membrane was transferred directly from the electroblotter into a Ponceau S solution (1:10 dilution of 2% 3-hydroxy-4-[2sulfo-4-(sulfophenylazo)-phenylazo]-2,7-

naphthalene disulfonic acid in 30% trichloroacetic acid) prior to blocking. The membrane was stained for 10 min. at room temperature on a shaker and then gradually destained in PBS until protein bands became visible.

2.5.4. Promastigote secretory gel (PSG) extraction

PSG extraction was adapted from Rogers *et al.* (2009) (*254*). Briefly, ten sand fly midguts were dissected in a drop of sterile PBS and the TMGs were transferred into an Eppendorf tube containing 50 µL PBS. The samples were centrifuged 6x at 10,000x*g* for 5 min. and the supernatant was transferred into fresh Eppendorf tubes every time. The PSG suspensions were stored at -20 °C until use.

2.5.5. PSG detection by Dot-blot

Nitrocellulose membrane was activated in 100% methanol for 30 – 45 sec. and was then washed in distilled H_2O for 2 min. prior to washing it in PBS for 5 – 10 min. and then drying briefly. Drops of 2 – 5 µL of PSG extract were applied to a moist nitrocellulose membrane placed on a stack of moist blotting paper and left to soak in for a few minutes. The membrane was briefly transferred onto a stack of dry blotting paper and was then blocked with PBS / 0.05% TWEEN / 5% milk overnight, prior to further treatment as specified in 2.5.2.

2.5.6. Biotinylation assay

The biotinylation assay was preceded by cell sorting of AMCA-Sulfo-NHS (sulfosuccinimidyl-7-amino-4-methylcoumarin-3-acetate) (Pierce, Perbio Rockford) – an amine-reactive fluorophore – labelled (live/dead labelled) cells in a modular flow cytometer (MoFlo) to distinguish between intact and damaged cells. The live/dead labelling method is described elsewhere ((*336*); see also 2.7.3.2.). The cell sorting by MoFlo was performed in the Cytometry lab of the Technology Facility of the University of York.

The biotinylation assay is described elsewhere (*335*). Briefly, sorted live cells were washed 3x in 1 ml ice cold PBS (pH 8) and re-suspended in ice cold PBS (pH 8) to an approximate concentration of 10^8 cells / ml. Biotin solution was freshly prepared from a solid Sulfo-NHS-SS-biotin (sulfosuccinimidyl-2-[biotinamido]ethyl-1,3-dithiopropionate) stock (Thermo Scientific) in ultrapure water to a final concentration of 10 mM (e.g. 2.2 mg / 360 µL) and was added immediately to a final concentration of 1 mM to the cell suspension. Cells

were incubated for 5 min. on ice before quenching the reaction by three washes in 1 ml ice cold Tris Buffered Saline (TBS; pH 8) + 50 mM NH_4Cl + 50 mM glycine. Parasites were washed one more time in 1 ml ice cold TBS (pH 8) to remove traces of quenching solution. Cells were lysed in 200 µl PBS (pH 7) + 1% SDS containing one cOmplete Mini, EDTA-free tablet (Roche) / 10 ml) and boiled for 3-5 min. The volume was adjusted to 1 ml by addition of PBS (pH 7) + 2% Triton X-100 (Sigma). Streptavidin resin was warmed to room temperature, shaken and 100 µl transferred to an Eppendorf tube per extraction reaction. The resin was spun briefly and the supernatant was removed. The cell lysate was poured onto the streptavidin beads and incubated for 1 h on a wheel shaker at room temperature. The streptavidin beads were washed at least 5x in 1 ml PBS (pH 7) + 0.1% SDS and spun briefly at 750 rpm. 1 volume of 2x Laemmli buffer was added to the beads and the sample boiled for 10 min. before loading the supernatant onto a 12% SDS-PAGE for a Western blot. Alternatively, samples were frozen at -20 °C and re-boiled before use.

2.6. Sand fly manipulation

The establishment and maintenance of sand fly colonies were done by the staff at the Charles University, Prague, CZ. Procedures can be viewed in Volf & Volfova 2011 (*370*).

2.6.1. Sand fly strains

For sand fly infection studies, the specific sand fly vector, *Ph. (Ph.) papatasi*, and the permissive vector, *Ph. (Ph.) duboscqi*, were used (**Table 2.6**). The adult sand flies were maintained in net cages (**Fig.2.2 B**) at ~26 °C in humid conditions on a 50% sucrose solution with 14 hours of light and 10 hours of dark photoperiods per day as described by Benkova and Volf (2007) (*371*). Cotton wool was drenched in the 50% sucrose solution and small portions of it were offered to the sand flies on small glass dishes, which were placed inside the cages. These were replaced every other day. The net cages were kept in transparent plastic sacks, which also contained an open Petri dish containing moist cotton wool for humidity.

2.6.2. Artificial sand fly infections

Methods for artificial sand fly infections are described elsewhere (*146*). In detail, parasites were grown to day 3 p.i. in 2 ml M199 cultures at 23 °C and were collected by centrifugation at 5,000 rpm (~2350x*g*) for 5 min. at room

Table 2.6 – Sand Fly Vector Strains used in this study

Genus	Subgenus	Species	Origin
Phlebotomus	*Phlebotomus*	*papatasi*	Turkey
Specific vector for *L. (L.) major*; belongs to the same subgenus and is closely related to *Ph. (Ph.) duboscqi* (females are morphologically indistinguishable) (*372*)			
Phlebotomus	*Phlebotomus*	*duboscqi*	Senegal
More permissive vector for *L. (L.) major*; belongs to the same subgenus and is closely related to *Ph. (Ph.) papatasi* (females are morphologically indistinguishable) (*372*)			

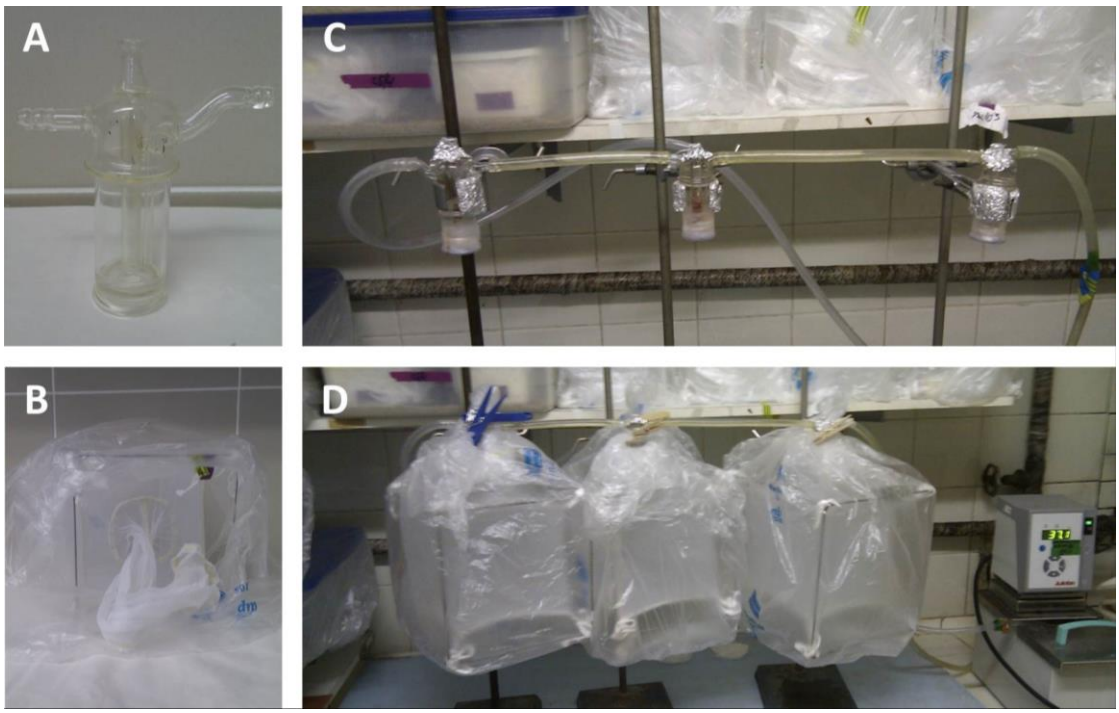


Fig.2.2 – Artificial sand fly feeding method

A) Double walled glass feeder. B) Sand fly net cage attached to a metal frame kept in a plastic bag with moist cotton wool on a Petri dish for humidity. C) Infected blood loaded lass feeders attached to 37 °C water bath by tubing. D) Net cages attaché to loaded glass feeder for sand fly infection.

temperature in a Jouan BR4*i* centrifuge. Cell pellets were washed once in 1 ml sterile saline solution and were re-suspended in 1 ml fresh sterile saline solution. 10 µl were diluted in 990 µL 1% Formaldehyde in saline solution (1:100) and 10 µl of this dilution were applied per side of a haemocytometer (Brückers). Parasites were counted on 5x 16 tiles (= 0.5 µl) on each half of the haemocytometer under the light microscope. Parasite counts per ml were established by summing counts from both halves (parasites / µl) and multiplying by 10^2 (1:100 dilution in 1% Formaldehyde solution) and by 10^3 (to get to one ml).

For experimental infection, $1x10^6$ parasite cells / ml were used: for 3 ml infected blood, 300 µl parasite suspension in saline solution ($1x10^7$ parasites in total) were diluted (1:10) in 2700 µl heat-inactivated rabbit blood (35 min. incubation at 56°C). The required volumes of available parasite suspensions were calculated according to the formula:

$$\text{volume of parasite suspension (µl)} = \frac{1\text{x } 10^7 \text{ cells/ml}}{\text{counted cells x } 10^2 \text{ x } 10^3 \text{ cells/ml}} \text{ x } 300 \text{ µl}$$

Sand flies were fed with infected blood through sterile glass feeders (Fig.2.2 A) covered with a chick-skin membrane. Chick-skins were prepared by removal of the dorsal and ventral skin of plugged chicks with sterile scissors and forceps. The skins were washed once in sterile saline solution in a Petri dish and the attached adipose tissue was removed. The skins were transferred into 70% ethanol and left for a few minutes before transferring them into fresh sterile saline solution. Skins were spread, exterior face down, in individual Petri dishes, sealed and were stored at -20 °C. Skins were defrosted before attaching them to the glass feeders with parafilm under sterile conditions. Infected blood was loaded onto the feeder and it was verified that none leaked out. Loaded feeders were attached by tubing to a 37 °C water bath to keep the blood at physiological temperature (Fig.2.2 C). Net cages containing female sand flies were attached to the feeders and were left wrapped in plastic bags in the dark for 1 – 2 h to feed (Fig.2.2 D). CO_2 was exhaled into the cages to stimulate blood feeding.

2.6.3. Sand fly midgut dissection and analysis

This procedure is described elsewhere (*146*). Briefly, 10 – 15 sand flies per infection were collected by aspirator on day 2, 5 or 6, 9 and 12 PBM (later only on day 6 and 12 PBM). Sand flies were stunned by cold in a collection

vessel and their midguts were dissected in sterile saline solution by first removing head, legs and wings with a small needle with 90° angle and then pulling the midgut out by gently tearing off the rear two abdominal segments with the needle. Dissected midguts were split into AMG and TMG and analysed separately under the light microscope for parasite localization and infection load. AMG and TMG were smeared on the slides by pressing a cover slip onto them. The gut smear glass slides were air-dried and specimens were fixed with 100% methanol at room temperature.

Parasite loads were established (1) by light microscopy, scoring as either uninfected, light (<100 parasites/gut), moderate (100-1000 parasites/gut), heavy (>1000 parasites/gut) (according to Myšková *et al.* (2008) (*373*)) or very heavy (>>1000 parasites/gut) infection and (2) by qPCR as described by Sádlová *et al.* (2010) (*146*).

2.6.4. Gene regulation in culture

Sand flies were fed with heat inactivated blood and were allowed to live for up to 12 days PBM. 50 blood fed midguts were dissected at day 6 and 12 PBM into 200 µL M199 + Amikin (250 µg/ml) + penicillin (60 µg/ml) + fluorocytosin (1.5 mg/ml). The midguts were homogenised and filter through a 0.22 µm filter spinning column (Ultrafree – MC, GV dutapore ®). The midgut extract was added to 4 ml of M199 and 1 ml was aliquoted into culture tubes. FVI, *Lmj*cDNA16 dKO, *Lmj*cDNA16 sKI and *Lmj*HASPB sKI were inoculated into the 1 ml medium, respectively, and were left to grow for 6 days. Cell were pelleted and washed twice in PBS before suspending in 25 µl PBS. 25 µl of 2x SDS-loading dye was added and the samples were boiled for 10 min. at 95 °C before storage at -20 °C.

2.7. Microscopy

2.7.1. Giemsa stained gut slide analysis by light microscopy

For light microscopic imaging of parasite cells at day 5, 6, 9 and 12 PBM, gut smears on glass slides were stained for 20 min. with a 1:20 dilution of Giemsa (Sigma-Aldrich) and briefly rinsed with water before analysis under the Olympia BX51 upright light microscope at a 1000x magnification (100x oil-immersion). 130 images per slide (65 of AMG smear, 65 of TMG smear) were taken with an Olympus DP70 camera using the DP controller software (Olympus) and parasite cells were measured for morphological analysis using

Image J software.

2.7.2. Confocal microscopy

For the confocal analysis of immunostained samples on glass slides, an upright Zeiss LSM 510 and invert LSM 710 META confocal microscope was used.

2.7.2.1. Fixed parasite antibody staining

The procedure is described elsewhere (*374*). $1x10^7$ parasites were spun down, washed and suspended in PBS. 1 volume 4% formaldehyde was added to the cell suspension and incubated for 15 min. at room temperature before washing twice in PBS. For the first protocol, fixed parasite cells were suspended in 200 µL PBS. Wells (1.5 cm x 1.5 cm) were marked on poly-lysine glass slide with a Pap pen (Sigma) and ~$5x10^6$ and ~$5x10^5$ cells were applied per well per sample. Slides were left for 30 min. for parasites to settle and attach. The liquid was removed and samples were incubated for 15 min. with 100 µL Triton-X in PBS (0.2%) per well for plasma membrane permeabilization. Liquid was removed and wells were washed once with 100 µL PBS per well. Samples were blocked with Image-iT FX signal enhancer (Invitrogen) for 30 min. Liquid was removed again and samples were washed once with 100 µL PBS per well. Samples were incubated for 1 hour with the anti-HASPB 336 or anti-SHERP antibody and the slides were washed at least 3x for 5min. in PBS. Samples were incubated in the dark for 1 hour with the Alexa Fluor® 488 Dye (Invitrogen) secondary anti-rabbit antibody diluted 1:250 in PBS. Slides were washed at least 3x for 5 min. in PBS in the dark before drying and sample mounting with 10 – 15 µL of Vectashield® or Mowviol® with 4', 6-diamidino-2-phenylindole (DAPI; Vector). Mowviol was prepared by heating 6 g Glycerol and 2.4 g Mowviol 4-88 [Hoechst. Calbiochem.] in 6 ml MQ H_2O and 12 ml 0.2 M Tris [pH 8.5] to 50 °C and repeatedly inverting the mixture at regular intervals until everything had dissolved. The mixture was centrifuged at 5000x*g* for 15 min. and aliquoted for storage at -20 °C or 4 °C (for aliquots in use). Samples were covered with coverslips and sealed with nail polish and were either immediately analysed or stored at 4 °C for next day analysis.

2.7.2.2. Live / dead staining

This protocol was described previously (*294*, *336*). $1x10^7$ washed parasites

were suspended in 95 µL PBS and 5µl AMCA-Sulfo-NHS (sulfosuccinimidyl-7-amino-4-methylcoumarin-3-acetate) (Pierce, Perbio Rockford) – an amine-reactive fluorophore – and incubated for 10 min. on ice to stain dead cells bright blue. The AMCA reaction was quenched by addition of 10 µL of 100 mM Tris (pH 8.5) and incubation for 5 min. on ice. Samples were washed 3x in 1% cold fatty acid free BSA/PBS. Washed parasites were suspended in 100 µL 1% BSA/PBS for 20 min. at RT. Triton-X100 was added to 0.1% final concentration for cell permeabilization when required. Cells were washed once in 1% BSA/PBS and suspended in 100 µL 1% BSA/PBS with the anti-HASPB antibody (336) (1:200) for a 30 min. incubation. Cells were washed at least 3x for 5 min. in PBS and suspended in 4% formaldehyde in PBS for a 15 min. incubation on ice for fixation. Cells were washed once for 5 min. in PBS and were suspended in 100 µL 1% BSA/PBS with the Alexa Fluor® 488 Dye (Invitrogen) secondary anti-rabbit antibody for 1 hour incubation at RT in the dark. Cells were washed 3x for 5 min. in PBS and were suspended in 200 µL PBS. Wells (1.5 cm x 1.5 cm) were marked on poly-lysine glass slides with a Pap pen (Sigma) and 100 µL of cell suspension were loaded per well and incubated for 30 min. in the dark. Parasites were mounted as described in 2.7.2.1.

3. Chapter III. – Generating *Leishmania* HASP and SHERP replacement mutants

3.1. Introduction

The diploid *L. (L.) major* cDNA16 locus contains four distinct contiguous genes, of which one, SHERP, occurs in two copies (Fig.1.25A). These genes – unusual for *Leishmania* parasites – were shown to be stage specifically regulated (Fig.1.24) (*329*, *330*). Previously published work on the *L. (L.) major* cDNA16 locus had shown that full deletion of this locus did not show any significant phenotype in *in vitro* culture or mice (*332*). Conversely, when the *L. (L.) major* cDNA16 null mutant (*Lmj*cDNA16 dKO) was introduced into the alimentary tract of its specific sand fly vector, *Ph. (Ph.) papatasi*, metacyclogenesis was stalled primarily in the nectomonad stage. This phenotype was rescued by the replacement of the full cDNA16 locus back into its original position on chromosome 23 (*Lmj*cDNA16 sKI) (*146*). Another mutant line containing only a single HASPB gene on an episome in the null mutant background was tested in the same study, with the data suggesting that HASPB on its own might be sufficient to rescue the parental strain (*L. (L.) major* Friedlin VI [FVI]) phenotype. However, it was shown that HASPB was overexpressed in this episomal replacement mutant. Episomes are circularized strands of independently replicating DNA, not unlike bacterial plasmids. Unusually for eukaryotic cells, *Leishmania* parasites are able to express genes on episomes and amplify them independently from the genome, although with the drawback of being unregulated and often overexpressed. Episomal genes have been shown to generate overexpression phenotypes that may differ strongly from the wild type phenotype, when the gene(s) in question are regulated and expressed at physiological levels (*332*).

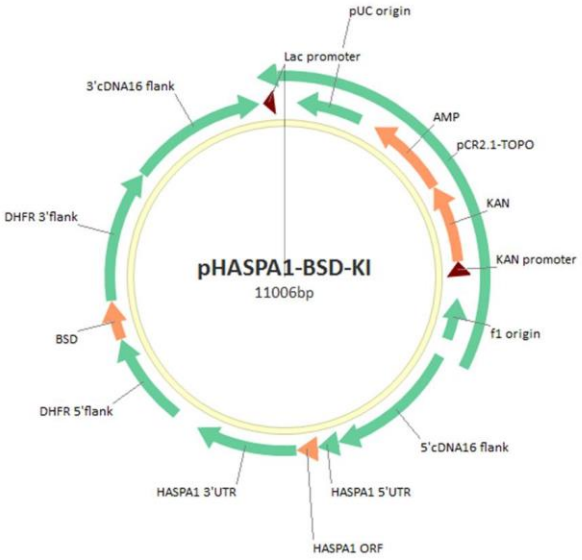
To confirm that the observations from Sádlová *et al.* (2010) were also true for regulated HASP and SHERP genes expressed at physiological levels and not an overexpressor phenotype, wild type gene regulation of these target genes had to be re-established. It had been shown previously that gene replacement by homologous recombination back into the original cDNA16 locus re-established wild type gene regulation (*146*, *332*). In this study, individual gene deletion from the cDNA16 locus was technically not possible due to high levels of sequence similarity in the intergenic regions between the HASP and SHERP genes. Therefore, gene replacement into the null background of *Lmj*cDNA16 dKO was chosen. Nine genes constructs containing HASPs and/or SHERP ORFs plus their native flanking sequences were synthesised to generate 17 mutant lines with

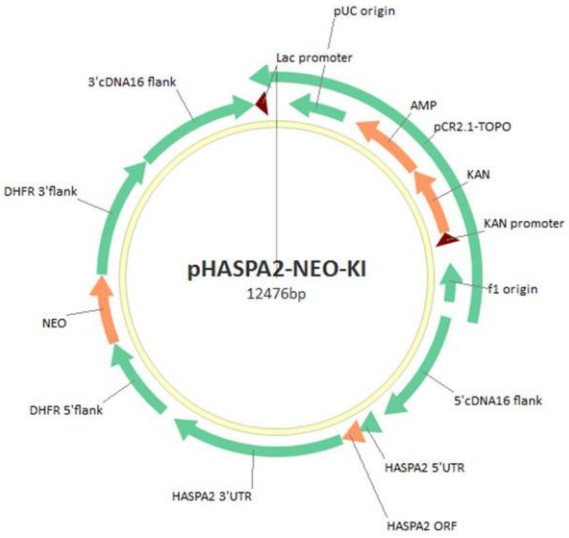
either one or several HASP and SHERP genes replaced into the cDNA16 locus in the null background.

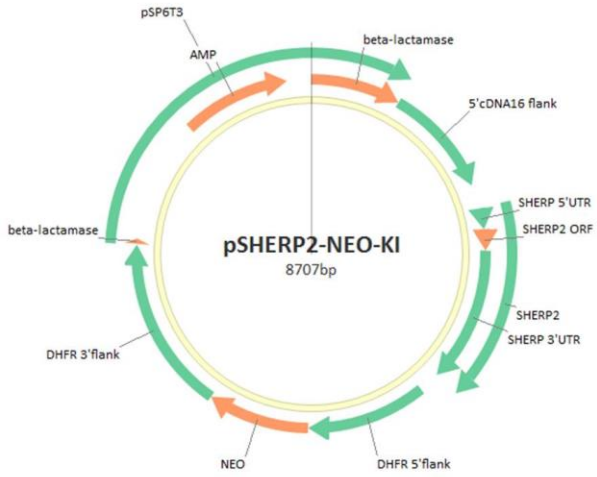
3.2. Recombinant construct generation

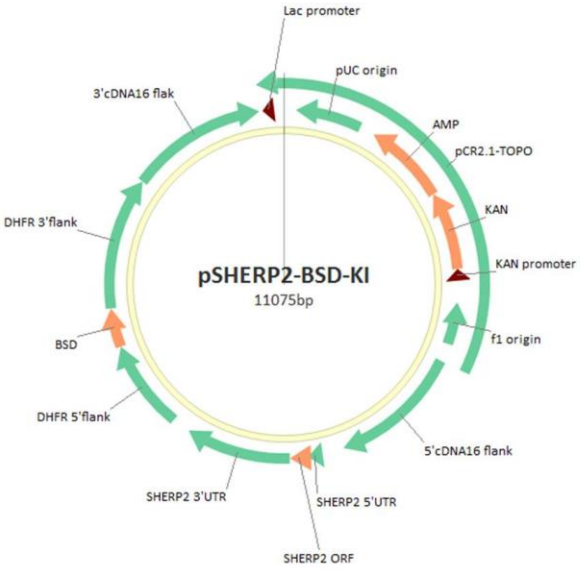
Nine gene constructs were generated containing HASP and/or SHERP gene(s) and either a neomycin (NEO), blasticidin (BSD) or streptophricin (SAT) selectable antibiotic resistance gene flanked by the flanking sequences of the *L. (L.) major* endogenous DHFR gene for constitutive marker expression (Table 2.4). Two HASPB (HASP–NEO-KI, HASPB–BSD-KI), two SHERP (SHERP–NEO-KI, SHERP–BSD-KI), two HASPB+SHERP (S2/HB–BSD-KI, S2/HB–SAT-KI), one HASPA1 (HASPA1–BSD-KI), one HASPA2 (HASPA2–NEO-KI) and one HASPA1+HASPA2 (HASPA1/2–NEO-KI) construct were synthesized. Due to high levels of sequence similarity, it was not possible to amplify SHERP1 individually and only SHERP2 was used in the construct generation. Since both SHERP ORFs have 100% identity, while the gene copies have 98.8% identity (*341*), it was considered sufficient to use only the SHERP2 copy in the construct and mutant generation.

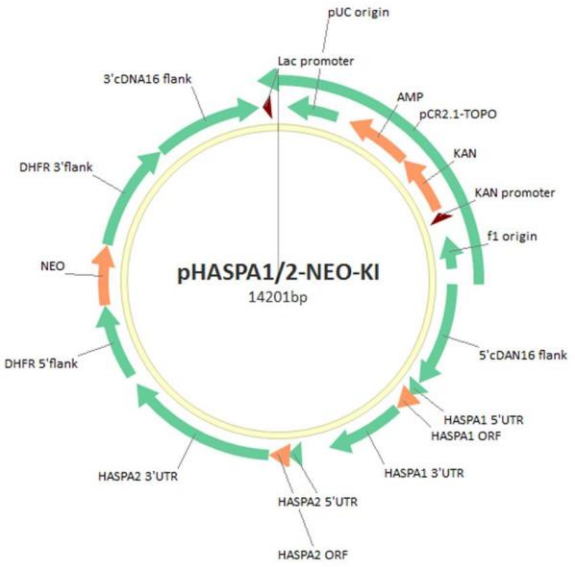
Apart from the HASPB–NEO-KI and the SHERP–NEO-KI constructs, which were generated by recycling the existing cDNA16 locus–NEO construct plasmid by construct substitution, all other gene constructs were generated by assembly in the pCR2.1®-TOPO® vector (Invitrogen). Generally, an antibiotic resistance marker (NEO, BSD or SAT) with the DHFR flanking regions, a 3'cDNA16 locus flank, HASPs and/or SHERP gene(s) and a 5'cDNA16 flank were cloned in that order step-by-step into the pCR2.1®-TOPO® vector to generate the constructs (Fig.3.1). The individual construct components were amplified by high fidelity PCR from either *L. (L.) major* gDNA or from other pre-existing plasmids, purified and verified by sequencing. The first construct component was integrated by 3' mono-(A) overhangs, while other construct components were integrated by restriction enzyme digestion of plasmid and construct component followed by a DNA ligation step. After each integration step, plasmids were transformed into chemically competent *E. coli* XL-1 cells, which were plated out on ampicillin or kanamycin agar plates and incubate at 37 ºC overnight. The *E. coli* colonies were picked the following day and checked by PCR screens for correct component integration. Selected clones were cultured overnight, the plasmids extracted by Mini- or MidiPreps (Qiagen) and verified by DNA sequencing for correct construct assembly. These steps were repeated until the full construct was assembled.

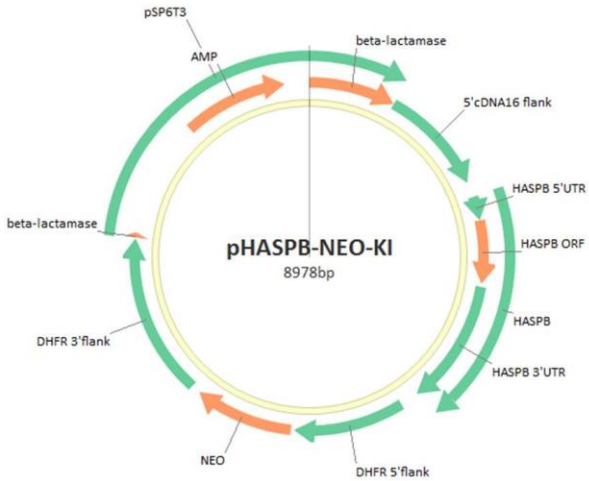

pUC origin
Lac promoter
3'cDNA16 flank
AMP
pCR2.1-TOPO
KAN
KAN promoter
DHFR 3'flank
pHASPA1-BSD-KI
11006bp
f1 origin
BSD
DHFR 5'flank
5'cDNA16 flank
HASPA1 3'UTR
HASPA1 5'UTR
HASPA1 ORF


pUC origin
Lac promoter
3'cDNA16 flank
AMP
pCR2.1-TOPO
KAN
DHFR 3'flank
KAN promoter
pHASPA2-NEO-KI
12476bp
f1 origin
NEO
5'cDNA16 flank
DHFR 5'flank
HASPA2 5'UTR
HASPA2 3'UTR
HASPA2 ORF


pSP6T3
AMP
beta-lactamase
5'cDNA16 flank
SHERP 5'UTR
SHERP2 ORF
beta-lactamase
pSHERP2-NEO-KI
8707bp
SHERP2
SHERP 3'UTR
DHFR 3'flank
NEO
DHFR 5'flank


Lac promoter
pUC origin
3'cDNA16 flak
AMP
pCR2.1-TOPO
KAN
DHFR 3'flank
KAN promoter
pSHERP2-BSD-KI
11075bp
f1 origin
BSD
DHFR 5'flank
5'cDNA16 flank
SHERP2 3'UTR
SHERP2 5'UTR
SHERP2 ORF


pUC origin
Lac promoter
3'cDNA16 flank
AMP
pCR2.1-TOPO
KAN
DHFR 3'flank
KAN promoter
f1 origin
NEO
pHASPA1/2-NEO-KI
14201bp
5'cDAN16 flank
DHFR 5'flank
HASPA1 5'UTR
HASPA1 ORF
HASPA2 3'UTR
HASPA1 3'UTR
HASPA2 5'UTR
HASPA2 ORF


pSP6T3
AMP
beta-lactamase
5'cDNA16 flank
HASPB 5'UTR
beta-lactamase
HASPB ORF
pHASPB-NEO-KI
8978bp
HASPB
DHFR 3'flank
HASPB 3'UTR
NEO
DHFR 5'flank

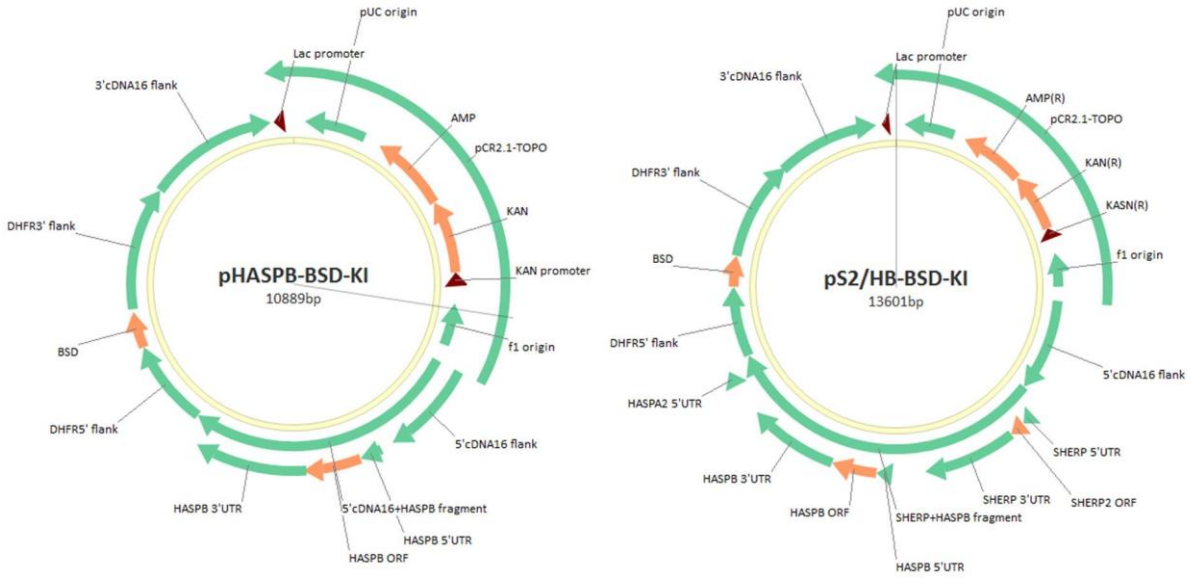


Fig.3.1 – Plasmids constructed for homologous recombination

Schematic representation of the plasmids generated containing the HASP and SHERP gene constructs for homologous recombination into the original cDNA16 locus in the *Lmj*cDNA16 dKO null background. ORFs, orange; other elements (flanking region, UTRs, origins), green.

Completed construct-containing plasmids were amplified in 20 ml LB cultures of transformed *E. coli* cells overnight at 37 ºC and were extracted by mini- or MidiPreps (Qiagen). 20 – 30 µg of purified plasmid were digested by required restriction enzymes for construct excision (Table 2.4) and the digests were then run on 1% agarose gels for gel purification and extraction by Qiagen kit. The purified constructs were then ethanol precipitated and re-suspended in sterile MQ H_2O to a final concentration of ~1 µg / µl.

3.3. *L. (L.) major* HASP and SHERP gene(s) mutant generation

The purified gene constructs were used for transfection by the AMAXA system into *L. (L.) major* mutant lines generating a series of different genotypes (Fig.3.2). 17 new mutant strains were generated; these are summarized in Table 2.1 together with all other *Leishmania* lines used in this study. *Lmj*HASPB sKI, *Lmj*SHERP sKI, *Lmj*S2/HB sKI, *Lmj*HASPA1 sKI, *Lmj*HASPA1/2 sKI and *Lmj*HASPA2 sKI were based on the *Lmj*cDNA16 dKO mutant; *Lmj*S2+HB sKI and *Lmj*HA1+S2 sKI were based on the *Lmj*SHERP sKI mutant; *Lmj*HASPB dKI and *Lmj*HA1+HB sKI were based on the *Lmj*HASPB sKI mutant; *Lmj*HA1+S2/HB sKI was based on the *Lmj*HASPA1 sKI mutant; *Lmj*HA1/2+HB sKI, *Lmj*HA1/2+S2 sKI and *Lmj*HA1/2+S2/HB sKI were based on the *Lmj*HASPA1/2 sKI mutant; *Lmj*HA2+HB sKI, *Lmj*HA2+S2 sKI and *Lmj*HA2+S2/HB sKI were based on the *Lmj*HASPA2 sKI mutant (Fig.3.3). Newly transfected mutant parasite lines were grown on Medium 199 (M199)-agar plates containing the appropriate antibiotic for up to 14 days for selection. Clones were picked as early as day 7 post plating and were grown up in sequentially increased amounts of M199 (100 µl → 1.5 ml → 5 ml → 10 ml M199) before extensive clone screening.

All SHERP containing mutants only contained a single copy of SHERP2. *Lmj*HASPB dKI contained two copies of HASPB to ensure that HASPB was expressed at parental strain levels, distinct from the *Lmj*HASPB sKI strain, which only contained a single HASPB gene copy. *Lmj*S2+HB sKI and *Lmj*S2/HB sKI had the same genotype, but are distinct, because *Lmj*S2+HB sKI has the SHERP and HASPB gene in separate constructs in the cDNA16 allele on neighbouring alleles, while *Lmj*S2/HB sKI has both genes in one construct on the same allele in the former cDNA16 locus. Since it was known from Flinn *et al.* (1992) and Keen *et al.* (unpublished) that polycistronic gene transcription and post-transcriptional gene regulation were important for HASP and SHERP gene expression and regulation from the original cDNA16 locus, we wanted to know if HASPB and SHERP gene regulation was different if genes were replaced individually or polycistronically.

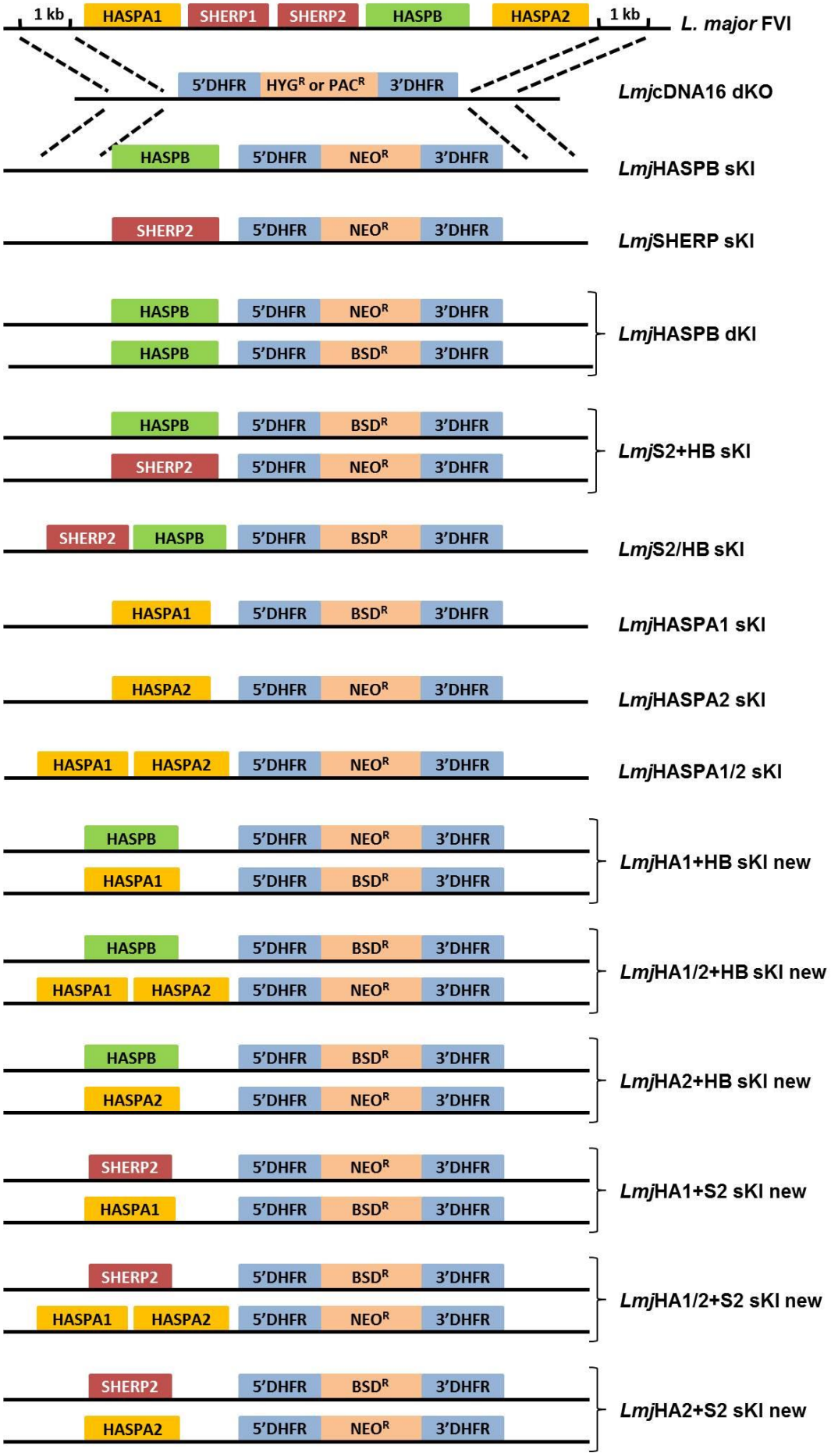

1 kb
HASPA1
SHERP1
SHERP2
HASPB
HASPA2
1 kb
L. major FVI
5'DHFR
HYGR or PACR
3'DHFR
LmjcDNA16 dKO
HASPB
5'DHFR
NEOR
3'DHFR
LmjHASPB sKI
SHERP2
5'DHFR
NEOR
3'DHFR
LmjSHERP sKI
HASPB
5'DHFR
NEOR
3'DHFR
HASPB
5'DHFR
BSDR
3'DHFR
LmjHASPB dKI
HASPB
5'DHFR
BSDR
3'DHFR
SHERP2
5'DHFR
NEOR
3'DHFR
LmjS2+HB sKI
SHERP2
HASPB
5'DHFR
BSDR
3'DHFR
LmjS2/HB sKI
HASPA1
5'DHFR
BSDR
3'DHFR
LmjHASPA1 sKI
HASPA2
5'DHFR
NEOR
3'DHFR
LmjHASPA2 sKI
HASPA1
HASPA2
5'DHFR
NEOR
3'DHFR
LmjHASPA1/2 sKI
HASPB
5'DHFR
NEOR
3'DHFR
HASPA1
5'DHFR
BSDR
3'DHFR
LmjHA1+HB sKI new
HASPB
5'DHFR
BSDR
3'DHFR
HASPA1
HASPA2
5'DHFR
NEOR
3'DHFR
LmjHA1/2+HB sKI new
HASPB
5'DHFR
BSDR
3'DHFR
HASPA2
5'DHFR
NEOR
3'DHFR
LmjHA2+HB sKI new
SHERP2
5'DHFR
NEOR
3'DHFR
HASPA1
5'DHFR
BSDR
3'DHFR
LmjHA1+S2 sKI new
SHERP2
5'DHFR
BSDR
3'DHFR
HASPA1
HASPA2
5'DHFR
NEOR
3'DHFR
LmjHA1/2+S2 sKI new
SHERP2
5'DHFR
BSDR
3'DHFR
HASPA2
5'DHFR
NEOR
3'DHFR
LmjHA2+S2 sKI new

Fig.3.2 – Schematic representation of mutant genotypes

The schematic shows the linear constructs used for homologous recombination into the *L. (L.) major* cDNA16 locus, alignment with the mutant genotypes found in Table 2.1. For each construct, the structure of the 2 diploid alleles is shown. The wild type cDNA16 locus is shown at the top of the figure.

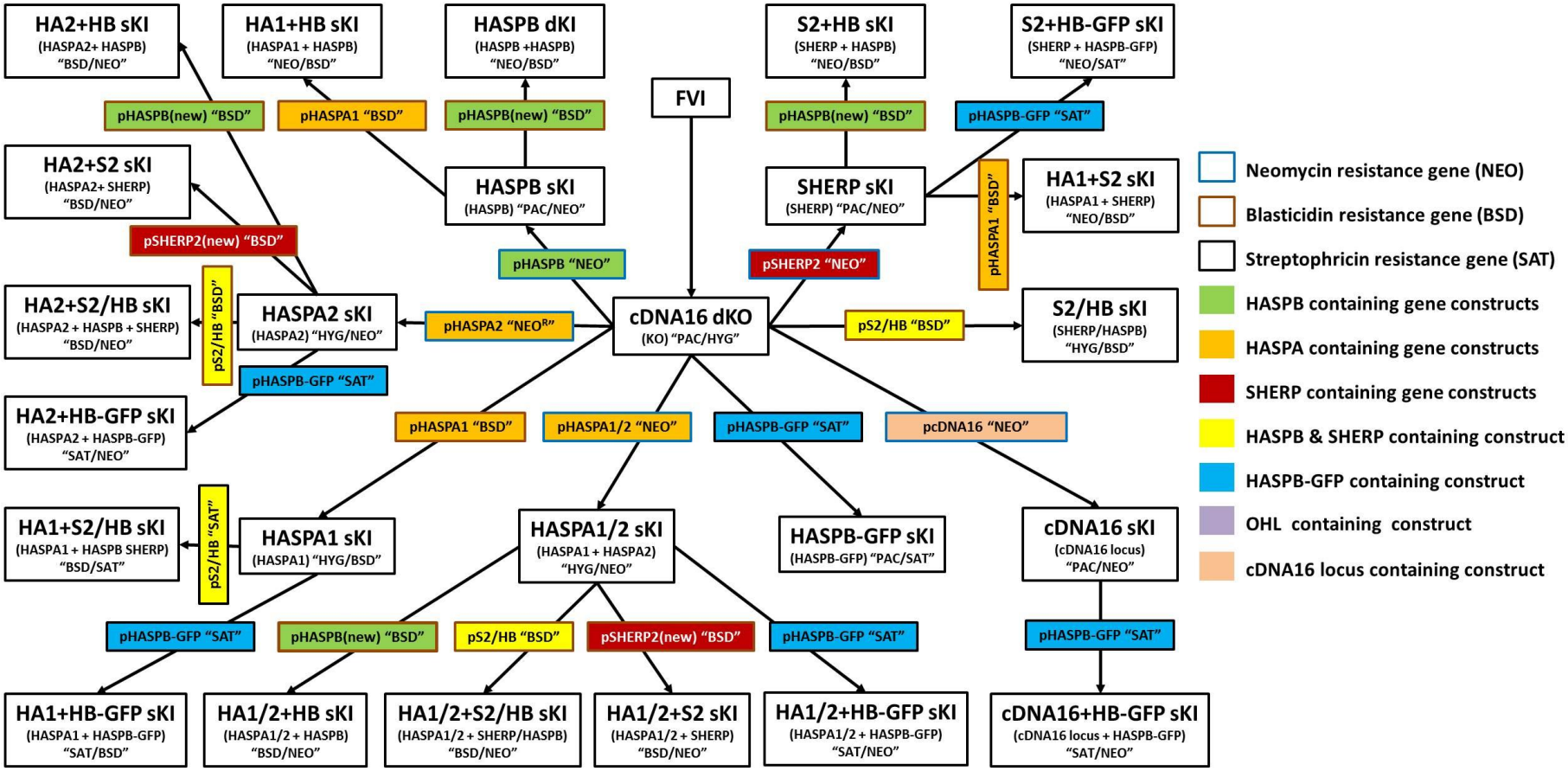


Fig.3.3 – Relationships of *L. (L.) major* wild type and mutant lines used in this study

The *L. (L.) major* Friedlin I (FVI) strain was used as the wild type. The *Lmj*cDNA16 dKO and *Lmj*cDNA16 sKI strains are described in McKean *et al.* (2001) and Sádlová *et al.* (2010), respectively. All other mutant strains were generated in this study. The diagram shows the relationships between the strains (large boxes) and which construct was used to generate them (coloured small boxes).

The *Lmj*cDNA16 dKO and *Lmj*cDNA16 sKI mutants had already been generated and used in previous work (*146*, *332*).

3.3.1. Screening the *L. (L.) major* cDNA16 mutant genes

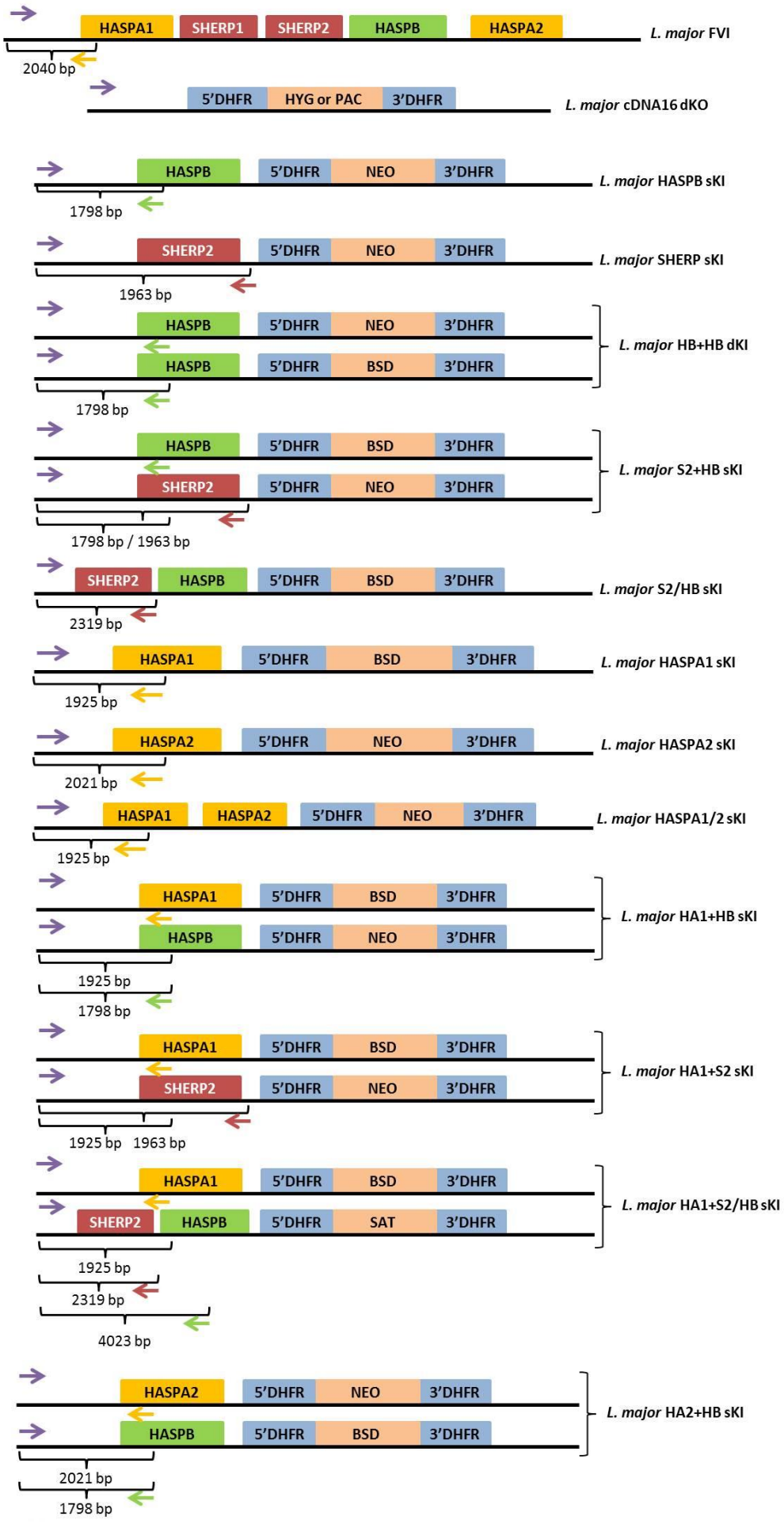
Given the complexities of *Leishmania* genome structure and regulation, it was essential to check each construct for correct integration of their transgenes. Transfected parasite clones were therefore screened for correct construct integration by conventional PCR and Southern blot. Integrated construct copy numbers were assessed by qPCR and protein expression and regulation by time course Western blots.

3.3.1.1. PCR screen of *L. (L.) major* cDNA16 mutant genes

Selecting mutants solely by PCR is error prone in aneuploid organisms of considerable genomic complexity. Therefore, PCR was only used as a simple positive/negative screen for construct integration. The forward primer (Lmj-H/S-F) was chosen upstream of the integration side on chromosome 23, while the reverse primers (Lmj-HASPB-R, Lmj-SHERP2-R or S2+HB-R) were within the construct (Fig.3.4). Primers Lmj-HASPB-R, Lmj-SHERP2-R and S2+HB-R were used when HASPB, SHERP or one of the HASPAs followed the 5'flanking region, respectively. Single gene constructs, SHERP-NEO-KI, SHERP-BSD-KI, HASPB-NEO-KI, HASPB-BSD-KI, HASPA1-BSD-KI, HASPA2-NEO-KI and double gene constructs S2/HB-BSD-KI and HASPA1/2-NEO-KI were expected to generate single bands of ~1.95 Kb, ~2.2 Kb ~1.8 Kb, ~1.8 Kb ~1.9 Kb, ~2 Kb, ~2.3 Kb and ~1.9 Kb, respectively, after amplification with the required primers. Mutant strains containing two constructs in the cDNA16 locus allele were checked for both. If the reverse primer Lmj-HASPB-R was used in mutants containing the S2/HB sKI construct, a ~4 Kb fragment was generated. Fig.3.5 shows gel images of selected *L. (L.) major* mutant clones that showed correct construct integration. These were further analysed by Southern blotting.

3.3.1.2. Southern blots of *L. (L.) major* cDNA16 genes mutant

Southern blotting and hybridization is the only method that can unequivocally demonstrate correct gene construct integration and also show the presence of random construct integration or episomes in addition to the correctly integrated construct. Therefore, Southern blots were employed as a refined screen of PCR positive clones. For this analysis,


HASPA1
SHERP1
SHERP2
HASPB
HASPA2
L. major FVI
2040 bp
5'DHFR
HYG or PAC
3'DHFR
L. major cDNA16 dKO
HASPB
5'DHFR
NEO
3'DHFR
L. major HASPB sKI
1798 bp
SHERP2
5'DHFR
NEO
3'DHFR
L. major SHERP sKI
1963 bp
HASPB
5'DHFR
NEO
3'DHFR
HASPB
5'DHFR
BSD
3'DHFR
L. major HB+HB dKI
1798 bp
HASPB
5'DHFR
BSD
3'DHFR
SHERP2
5'DHFR
NEO
3'DHFR
L. major S2+HB sKI
1798 bp / 1963 bp
SHERP2
HASPB
5'DHFR
BSD
3'DHFR
L. major S2/HB sKI
2319 bp
HASPA1
5'DHFR
BSD
3'DHFR
L. major HASPA1 sKI
1925 bp
HASPA2
5'DHFR
NEO
3'DHFR
L. major HASPA2 sKI
2021 bp
HASPA1
HASPA2
5'DHFR
NEO
3'DHFR
L. major HASPA1/2 sKI
1925 bp
HASPA1
5'DHFR
BSD
3'DHFR
HASPB
5'DHFR
NEO
3'DHFR
L. major HA1+HB sKI
1925 bp
1798 bp
HASPA1
5'DHFR
BSD
3'DHFR
SHERP2
5'DHFR
NEO
3'DHFR
L. major HA1+S2 sKI
1925 bp
1963 bp
HASPA1
5'DHFR
BSD
3'DHFR
SHERP2
HASPB
5'DHFR
SAT
3'DHFR
L. major HA1+S2/HB sKI
1925 bp
2319 bp
4023 bp
HASPA2
5'DHFR
NEO
3'DHFR
HASPB
5'DHFR
BSD
3'DHFR
L. major HA2+HB sKI
2021 bp
1798 bp

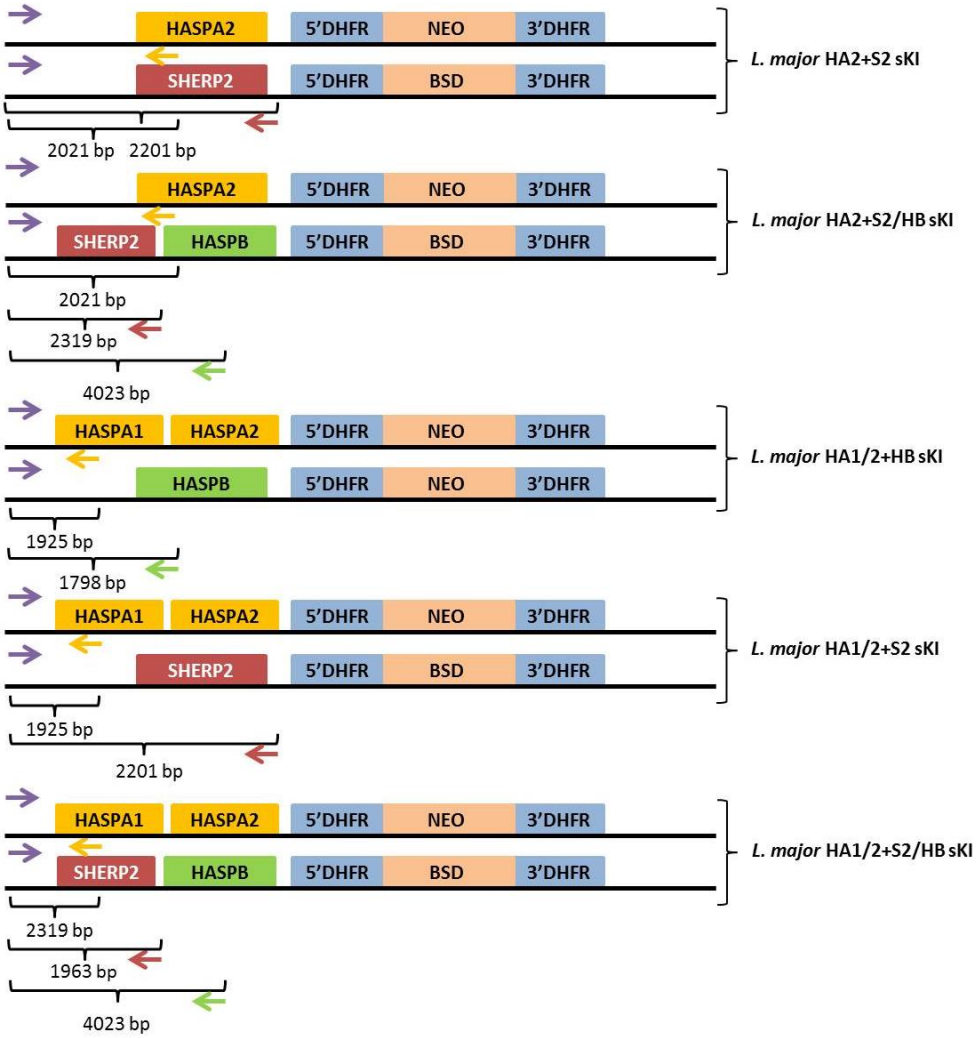


Fig.3.4 – Schematic representation of PCR screen

A forward primer (**Lmj-H/S-F**, purple arrow) binding up-stream of the construct integration site within the wild type genomic DNA and a reverse primer (**S2+HB-R**, orange arrow [FVI, *Lmj*HASPA1 sKI, *Lmj*HASPA2 sKI, *Lmj*HA1+HB sKI, *Lmj*HA1+S2 sKI, *Lmj*HA1+S2/HB sKI, *Lmj*HA2+HB sKI, *Lmj*HA2+S2 sKI, *Lmj*HA2+S2/HB sKI, *Lmj*HA1/2+HB sKI, *Lmj*HA1/2+S2 sKI, *Lmj*HA1/2+S2/HB sKI], ***Lmj*HASPB-R**, green arrow [*Lmj*HASPB sKI, *Lmj*HASPB dKI, *Lmj*S2+HB sKI, *Lmj*HA1+HB sKI, *Lmj*HA2+HB sKI, *Lmj*HA1/2+HB sKI], ***Lmj*SHERP2-R**, burgundy arrow [*Lmj*SHERP sKI, *Lmj*S2+HB sKI, *Lmj*S2/HB sKI, *Lmj*HA1+S2 sKI, *Lmj*HA2+S2 sKI, *Lmj*HA1/2+S2 sKI *Lmj*HA1+S2/HB sKI, *Lmj*HA2+S2/HB sKI, *Lmj*HA1/2+S2/HB sKI]) binding down-stream of the integration site within the construct, were used for PCR amplification. In each case, a fragment could only be amplified, if the respective gene construct had been integrated into the former cDNA16 locus. Fragments of 1,925 bp, 2,021 bp, 1,798 bp and 1,963 bp were expected for HASPA1, HASPA2 HASPB and SHERP, respectively. The wild type control (FVI) generated a 2,040 bp fragment.

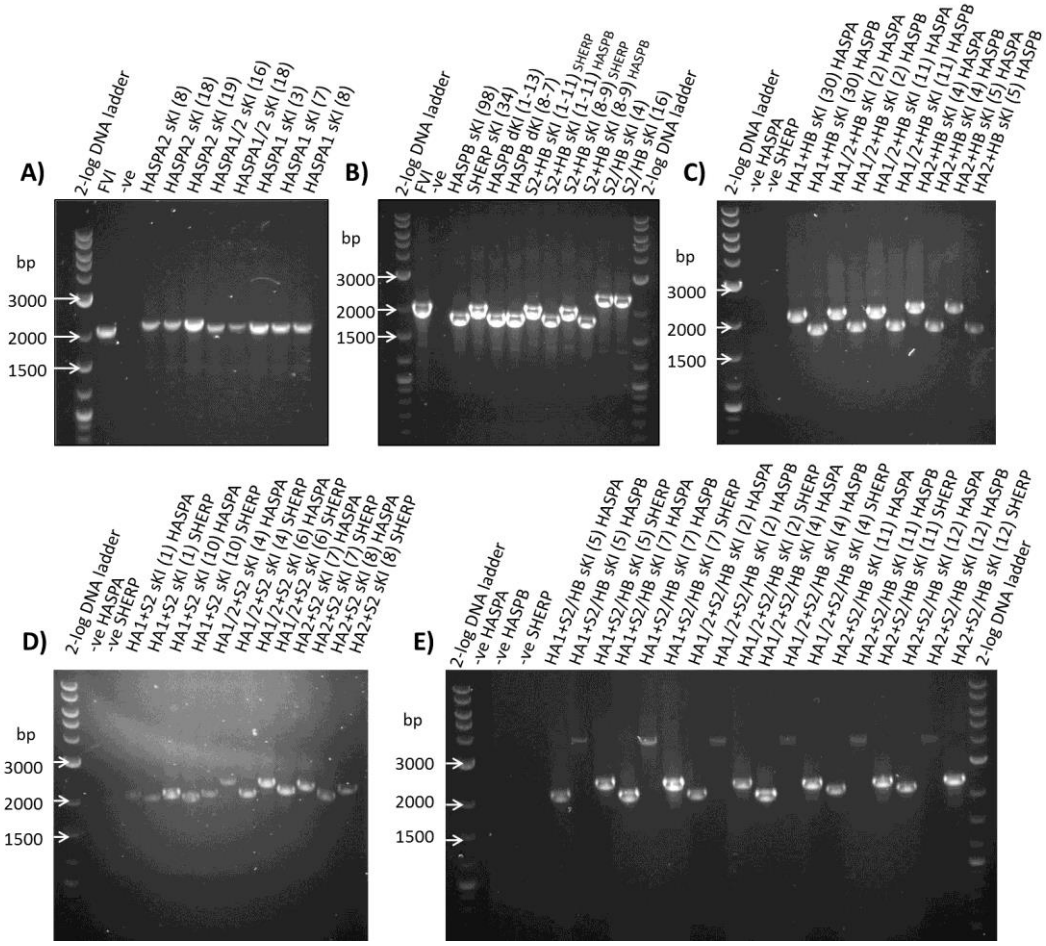


Fig.3.5 – PCR screen for construct integration in *Leishmania* mutant strains

The images show selected mutant clones screened positively for construct integration. Reverse primers were commonly chosen depending on the gene following the 5' flank in the construct (*Lmj*SHERP2-R for SHERP; *Lmj*HASPB-R for HASPB; S2/HB-R for HASPA1 & HASPA2). In E) S2/HB-R was also used in the S2/HB-NEO/BSD-sKI construct. Expected band sizes are listed in the Table. In the case of clones that were probed with more than one set of primers, the gene probed for is written behind the given clone name.

Table 3.1 – Expected band sizes according to primer pair used

Primer Pair	Construct	Approx. size (Kb)
Lmj-H/S-F & *Lmj*SHERP2-R	SHERP-NEO/BSD-sKI	1.95
	S2/HB-NEO/BSD-sKI	2.3
Lmj-H/S-F & *Lmj*HASPB-R	HASPB-NEO/BSD-sKI	1.8
	S2/HB-NEO/BSD-sKI	4
Lmj-H/S-F & S2/HB-R	HASPA1-BSD-sKI	1.9
	HASPA1/2-NEO-sKI	1.9
	HASPA2-NEO-sKI	2

gDNA was extracted from up to 10 ml day 5 cultures, which generally had a density of 3 – 3.6 x 10^7 cells / ml, with the Blood and Tissue kit from Qiagen, which yielded approximately 10 – 20 µg of gDNA. Where larger amounts of gDNA were required, the gDNA was extracted by a phenol-chloroform protocol from 50 ml day 5 cultures.

Extracted gDNA was digested by *Sac*I, run on a long 0.8% agarose gel overnight and transferred onto a positively charged nylon membrane (Roche) and analysed by specifically generated DIG-labelled probes (Roche) of between 300 – 500 bp in length (Fig.3.6), generated by high-fidelity PCR with a DIG-labelling kit (Roche). Each blot was probed with several probes as required. Fig.3.7 shows examples of the differently probed Southern blots with the PCR-selected mutant strains. All mutants shown in Fig.3.7 had bands of expected sizes. The 5' UTR probe, which hybridised to a section of the 5' flanking region used in the targeting for homologous recombination, proved particularly useful. Every construct and FVI contained it and with only two expected bands per mutant strain, this helped to distinguish quickly between correct integration and random or episomal integration. A list of expected band sizes that the respective probes were expected to detect on the Southern blots can be found in Table 3.2. The expected band sizes were based on the construct maps that were based on the verified *L. (L.) major* cDNA16 locus sequence submitted by the Smith lab (AJ237587). This has not been corrected on any online data base, where the locus is misassembled due to the difficulties of dealing with repetitive regions of the genome during automated assembly. Episomes generally produce very strong bands by hybridization since they are amplified independently from the genome. Selected mutants were further analysed by qPCR for construct copy number.

3.3.1.3. qPCR screen for replacement gene copy number

Using gDNA extracted from mutant strains and FVI, a qPCR was employed to verify the integrated construct copy number in the mutant genomes. Genomic DNA from FVI was used to generate a standard curve by serial dilution (1:10) of gDNA (10 ng → 1 ng → 0.1 ng → 0.01 ng → 0.001 ng). Mutant strain gDNA samples were used at 1 ng per well; initially 0.1 ng samples were used, too, but these were later abandoned as 1 ng amounts gave clearer results. Each sample was set up in 2x triplicates on 96 well plates (Fig.3.8). The first 48 wells were probed with HASPB (qPCR-H-F1 /

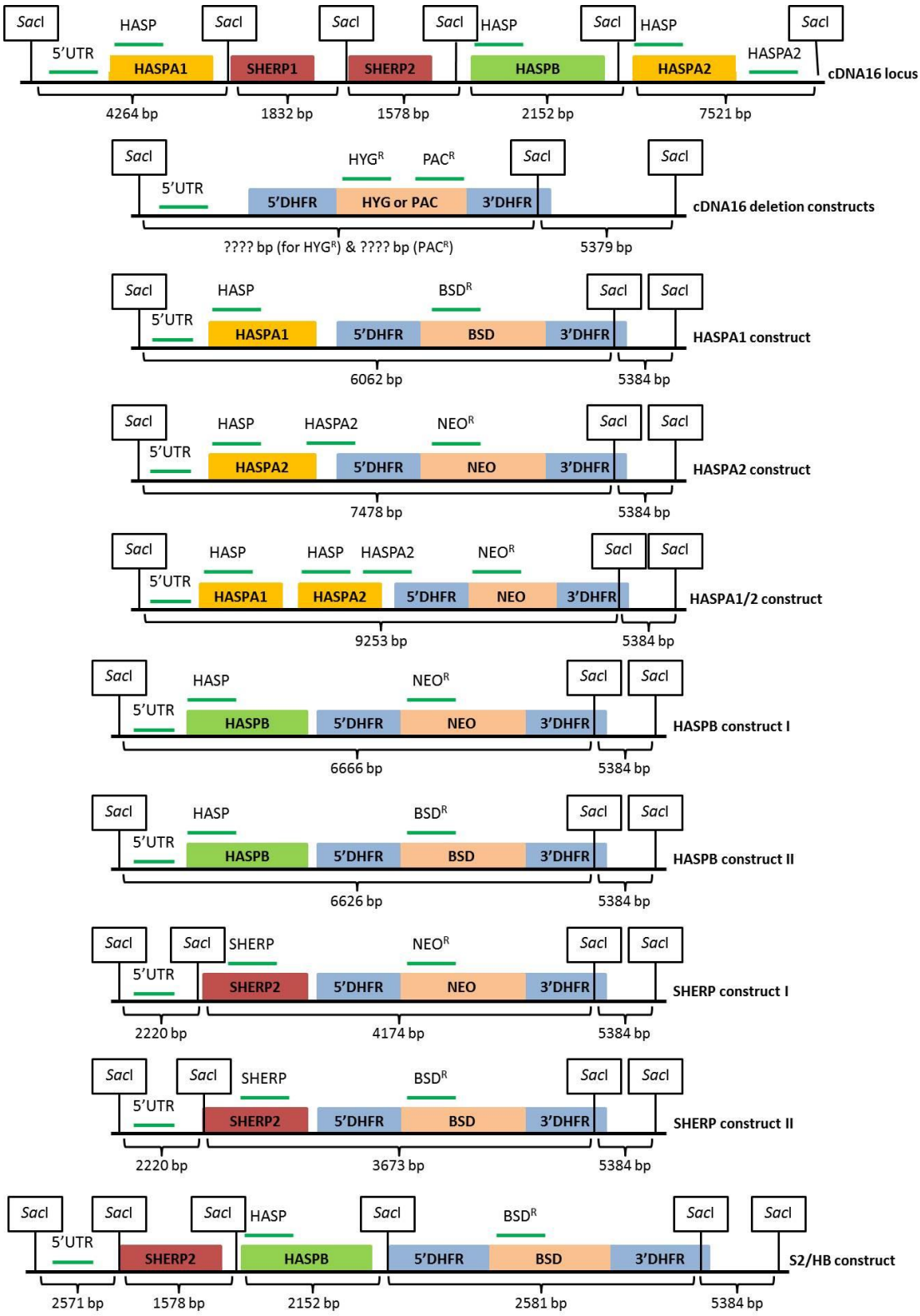
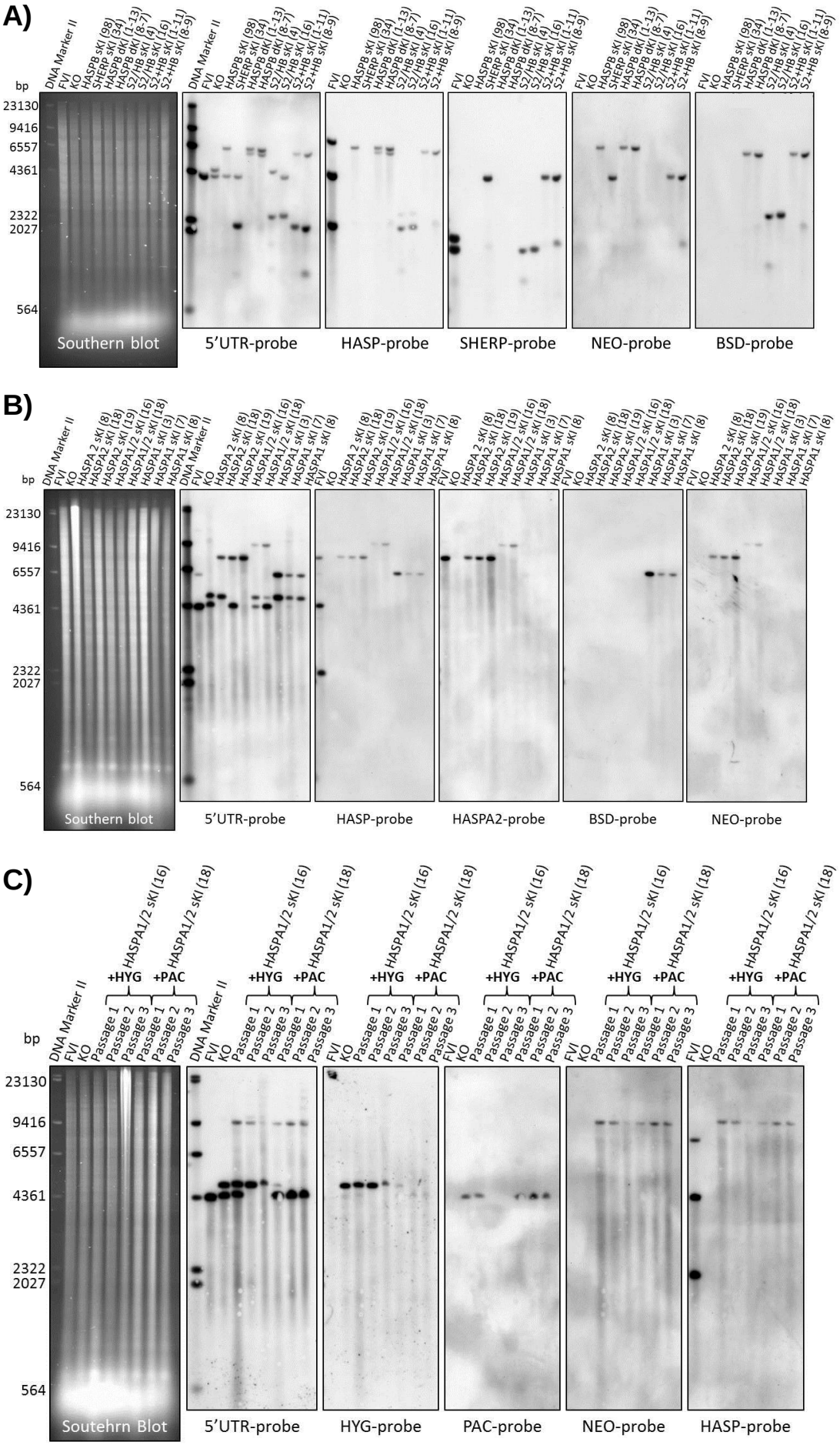


Fig.3.6 – Schematic of *Sac*I restriction sites and DIG-probe binding sites

The *Sac*I restriction sites following construct integration into the genomic DNA of mutant strains are shown. Binding sites of DIG-labelled probes (green bars) and the size of the fragments generated by *Sac*I digestion are shown. These correspond directly with the expected band sizes on the Southern blot, detected with the corresponding DIG-labelled probes. The schematic is not to scale.


A)
bp
23130
9416
6557
4361
2322
2027
564
Southern blot
5'UTR-probe
HASP-probe
SHERP-probe
NEO-probe
BSD-probe
B)
Southern blot
5'UTR-probe
HASP-probe
HASPA2-probe
BSD-probe
NEO-probe
C)
+HYG
+PAC
Soutehrn Blot
5'UTR-probe
HYG-probe
PAC-probe
NEO-probe
HASP-probe

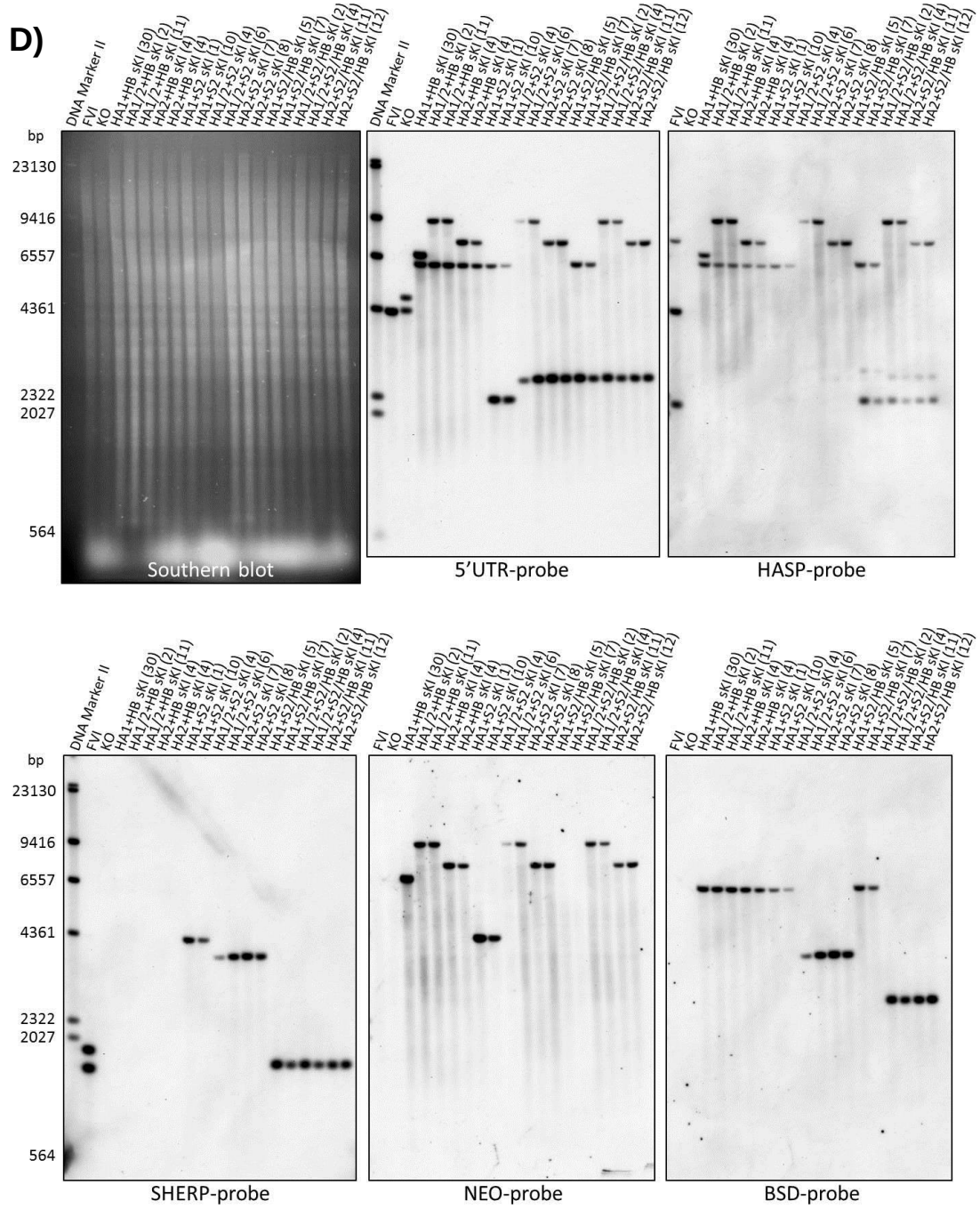
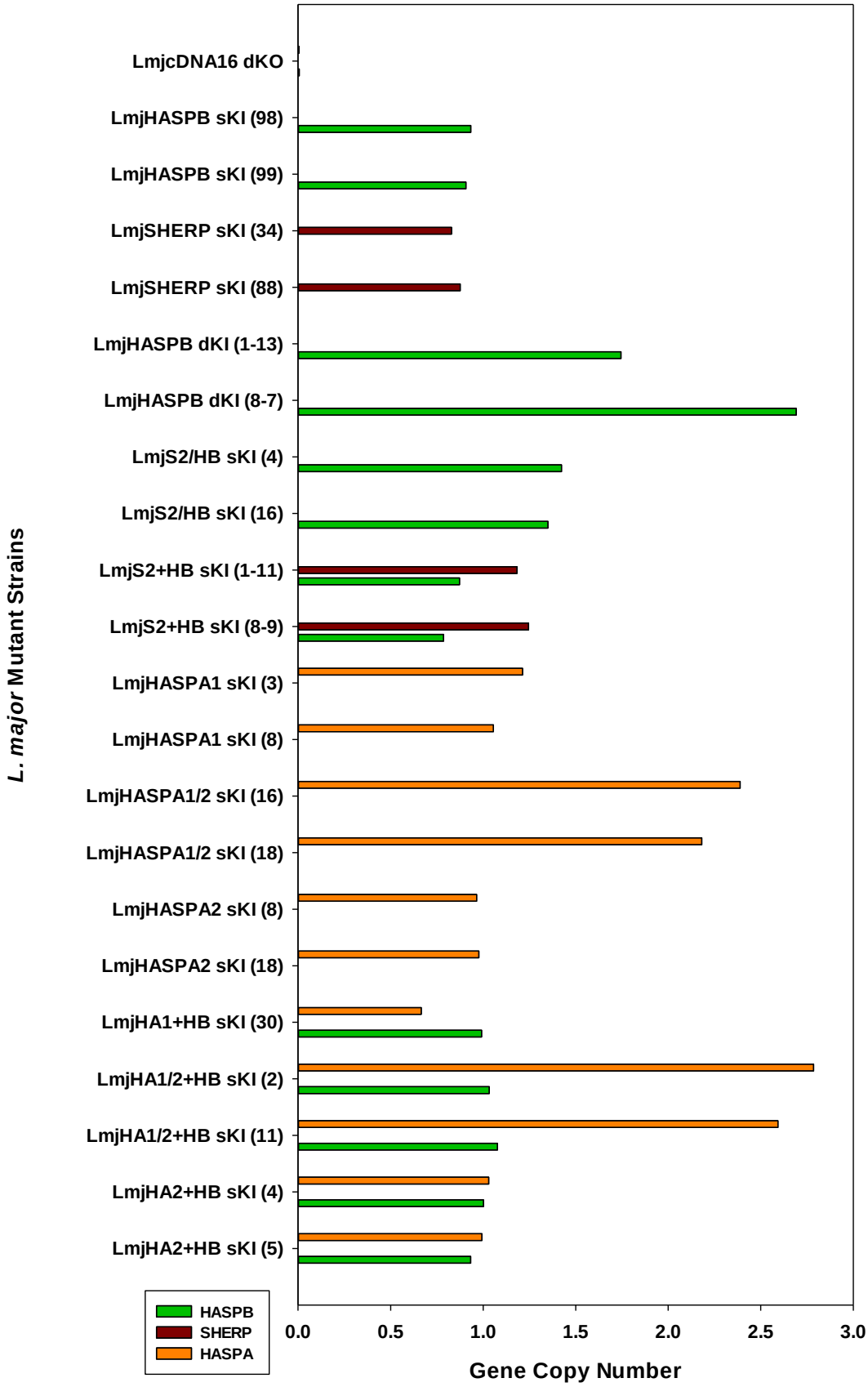


Fig.3.7 – Southern blot verification of construct integration

A), B) and D) show the Southern blots of *Sac*I digested genomic DNA samples probed with the relevant DIG-labelled probes for HASPA, SHERP, HASPB and the antibiotic resistance markers as a control for construct integration. The 5' UTR probe y bound in the 5' flanking region of the constructs and gave two signals per mutant line. All mutants represented here tested positive for correct integration. C) The 5' UTR probe was used as quick screen in the case of the HASPA1/2 sKI mutants which had shown three bands in B), to verify that the third band was not due to trisomy of chromosome 23. Growth in Hygromycin or Puromycin showed that one band was lost over time and proved that an inoculation of two mutants had occurred, which perhaps had grown over on another on the agar plates.

Table 3.2 – Expected band sizes for probes per integrated construct

Construct	DIG-labelled Probes	Nucleotides (bp)
SHERP-NEO-KI	5' UTR	2220
	SHERP	4172
	NEO	
SHERP-BSD-KI	5' UTR	2464
	SHERP	3673
	BSD	
HASPB-NEO-KI	5' UTR	6666
	HASP	
	NEO	
HASPB-BSD-KI	5' UTR	6206
	HASP	
	BSD	
HASPA1-BSD-KI	5' UTR	6062
	HASP	
	BSD	
HASPA2-NEO-KI	5' UTR	7478
	HASP	
	NEO	
HASPA1/2-NEO-KI	5' UTR	9253
	HASP	
	NEO	
S2/HB-BSD-KI	5' UTR	2464
	SHERP	1578
	HASP	2152
	BSD	2581


L. major Mutant Strains
LmjcDNA16 dKO
LmjHASPB sKI (98)
LmjHASPB sKI (99)
LmjSHERP sKI (34)
LmjSHERP sKI (88)
LmjHASPB dKI (1-13)
LmjHASPB dKI (8-7)
LmjS2/HB sKI (4)
LmjS2/HB sKI (16)
LmjS2+HB sKI (1-11)
LmjS2+HB sKI (8-9)
LmjHASPA1 sKI (3)
LmjHASPA1 sKI (8)
LmjHASPA1/2 sKI (16)
LmjHASPA1/2 sKI (18)
LmjHASPA2 sKI (8)
LmjHASPA2 sKI (18)
LmjHA1+HB sKI (30)
LmjHA1/2+HB sKI (2)
LmjHA1/2+HB sKI (11)
LmjHA2+HB sKI (4)
LmjHA2+HB sKI (5)
HASPB
SHERP
HASPA
0.0
0.5
1.0
1.5
2.0
2.5
3.0
Gene Copy Number

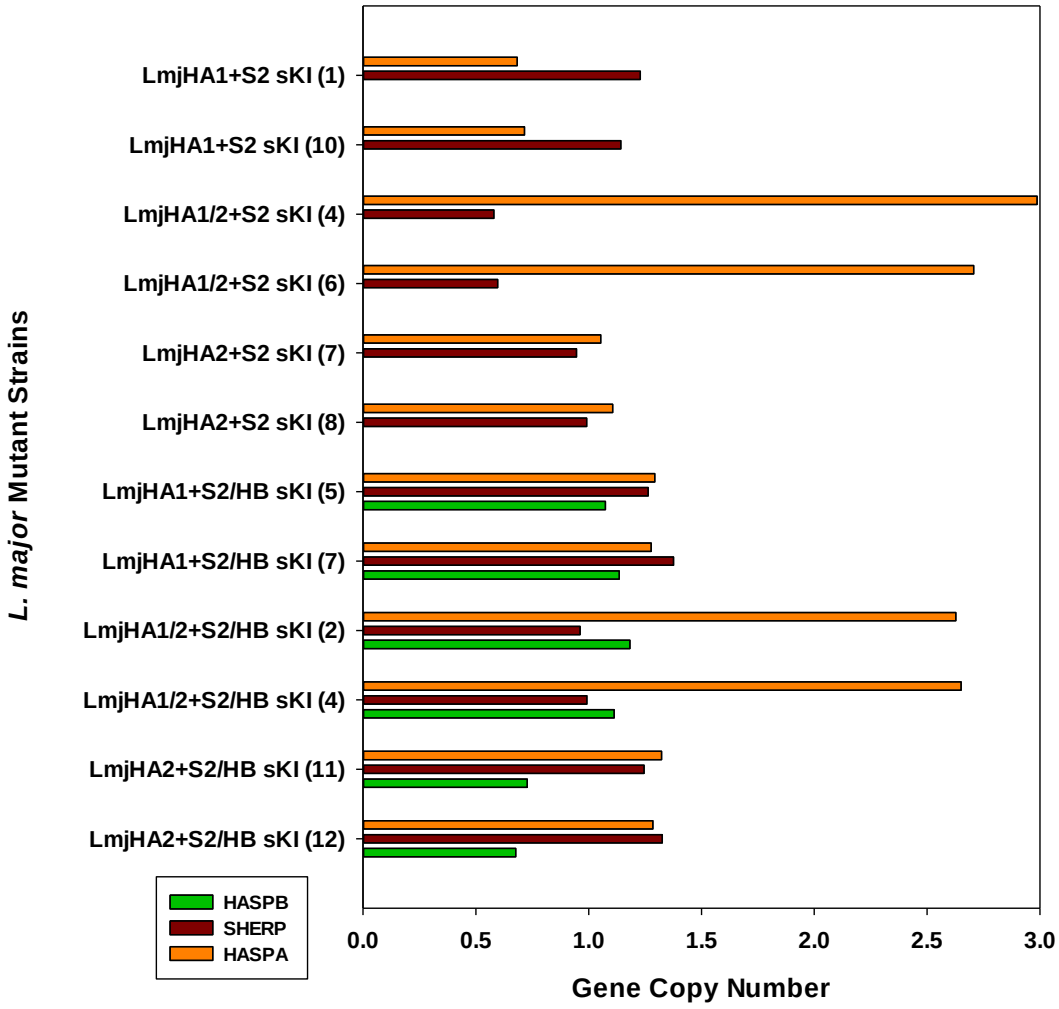


Fig.3.8 – qPCR for gene copy number verification

Genomic DNA preps were probed in a qPCR with gene-specific primers for HASA, HASPB, SHERP and a control gene on chromosome 23 (LmjF.23.0830). The parental line (FVI) was used to establish a standard curve. Mean quantities for HASPA, HASPB and SHERP were divided by the mean quantities of LmjF.23.0830 and adjusted for wild type gene copy number. With the exception of the *Lmj*HASPB dKI and HASPA1/2 construct containing lines, for which the expected value was 2, the expected value was 1 for a single integrated gene copy. The qPCR is error prone due to its sensitivity to small variations in pipetted gDNA volumes and 30% deviations were expected from the ideal values (1 or 2). Clones within this error range were accepted to contain only 1 or 2 gene copies depending on the line. The graph shows all strains selected from the Southern blot.

qPCR-H-R1), HASPA (qPCR-HA-F / qPCR-HA-R) or SHERP (qPCR-S-F1 / qPCR-S-R1) specific qPCR primers depending on the gene probed for, while the second 48 wells containing the same sample array were probed for the *Lmj*F.23.0830 gene (primers: qPCR-F23.0830-F1 / qPCR-F23.0830-R1), present as two copies on chromosome 23, as control to normalise the data. The mean results per strain were then divided by the mean results of its control and multiplied by the difference in gene copy number compared to FVI. The expected results were a ratio of 1 for all mutants containing single copy genes and 2 where two identical ORFs existed as in *Lmj*HASPB dKI and all mutants containing the HASPA1/2 construct. Single copy mutants varying >0.4 gene copies and double copy mutants varying >0.55 gene copies from the expected value were discarded from further experiments, where possible. Two clones from each strain, each having an approximate copy number of 1 or 2 as appropriate, were picked for analysis by Western blotting.

3.3.1.4. Western blot time courses to assess the expression and stage-regulation of the replacement genes

Western blot time courses were essential to assess not only gene expression at the protein level, but also gene regulation at the protein level. To analyse protein expression patterns in culture, 10^7, $2x10^7$ and $4x10^7$ parasites were collected from day 2-7 cultures, lysed and the extracts probed for HASPB, HASPA or SHERP with the available anti-HASPB (336), non-affinity purified anti-HASP and anti-SHERP antibodies, respectively. All antibodies were polyclonal antibodies raised in rabbits and 336 and anti-SHERP were affinity purified with an affinity column. Samples for HASP screening were usually run on 12% polyacrylamide gels, while samples for SHERP detection were run on 15% polyacrylamide gels, because of SHERP's small size (6.5 KDa), which also makes this protein difficult to blot. The use of 0.2 µm pore membranes from Millipore on a semi dry electroblotter gave the best results for SHERP detection, while HASPB and HASPA from 12% gels were blotted successfully onto Immobul membranes (Roche). All blots were probed first with the appropriate antibody for the respective protein of interest and then with an anti-rabbit immunoglobulin G (IgG) HRP antibody for signal detection by ECL plus or ECL prime.

The Western blot time courses were designed to show re-establishment of

parental line gene regulation and expression by homologous recombination of the HASP and SHERP genes back into the former cDNA16 locus. This was a critical demand on these new mutant lines, which would distinguish them from previously used episomal replacement lines (*146*, *332*). The results showed that in selected mutant lines, the replacement of the genes back into the cDNA16 locus was sufficient to re-establish the previously observed parental line (FVI) gene regulation (*334*, *341*). HASPB and SHERP showed increased expression from day 2-7 as expected (Fig.3.9). HASPA2 proved difficult to detect, because HASPA2 expression from the integrated construct was lower than anticipated in comparison to HASPA expression in FVI, which may be due to having only one HASPA2 copy present in the mutant lines compared to the two copies in FVI. In FVI, however, HASPA1 and HASPA2 cannot be distinguished, because both genes have an identical ORF. This could have added to the differences observed in HASPA expression levels between FVI and HASPA2-containing mutant lines. In some replacement mutants, HASPA2 was detected only at day 6 and 7 p.i. (Fig.3.9 I & R). HASPA1 was not detected at all with the non-affinity purified anti-HASPB antibody in mutant lines containing only HASPA1, but not HASPA2 (Fig3.9 G, J, M & P). This could mean that HASPA1 is amastigote specific, although their previously observed mRNA profiles suggested that upregulation at the transcription level occurs already in the metacyclic stage (*329*, *330*). Surprisingly, the HASPA1/2 construct containing replacement mutant lines (Fig.3.9 H, K, N & Q) expressed HASPA at higher levels than in the single HASPA2 replacement mutant lines (Fig.3.9 I, L, O, & R), although the HASPA2 containing DNA fragment, which had been used for construct generation, was the same in both constructs. All Western blots were also subsequently probed with an antibody against the constitutively expressed *N*-myristoyl transferase (NMT) as a loading control, to validate increased expression of HASP and SHERP over time. This loading control confirmed the increases in HASP and SHERP expression as real rather than artefactual. Selected mutant lines were inoculated into BALB/c mice for re-establishment of parasite virulence after prolonged parasite culturing.

3.3.1.5. Assessing HASPA1 expression in amastigotes

Total HASPA expression had previously been analysed in FVI and shown to be gradually upregulated from the procyclic to the metacyclic stage, with continued expression in amastigotes (*334*). However, since HASPA1 and

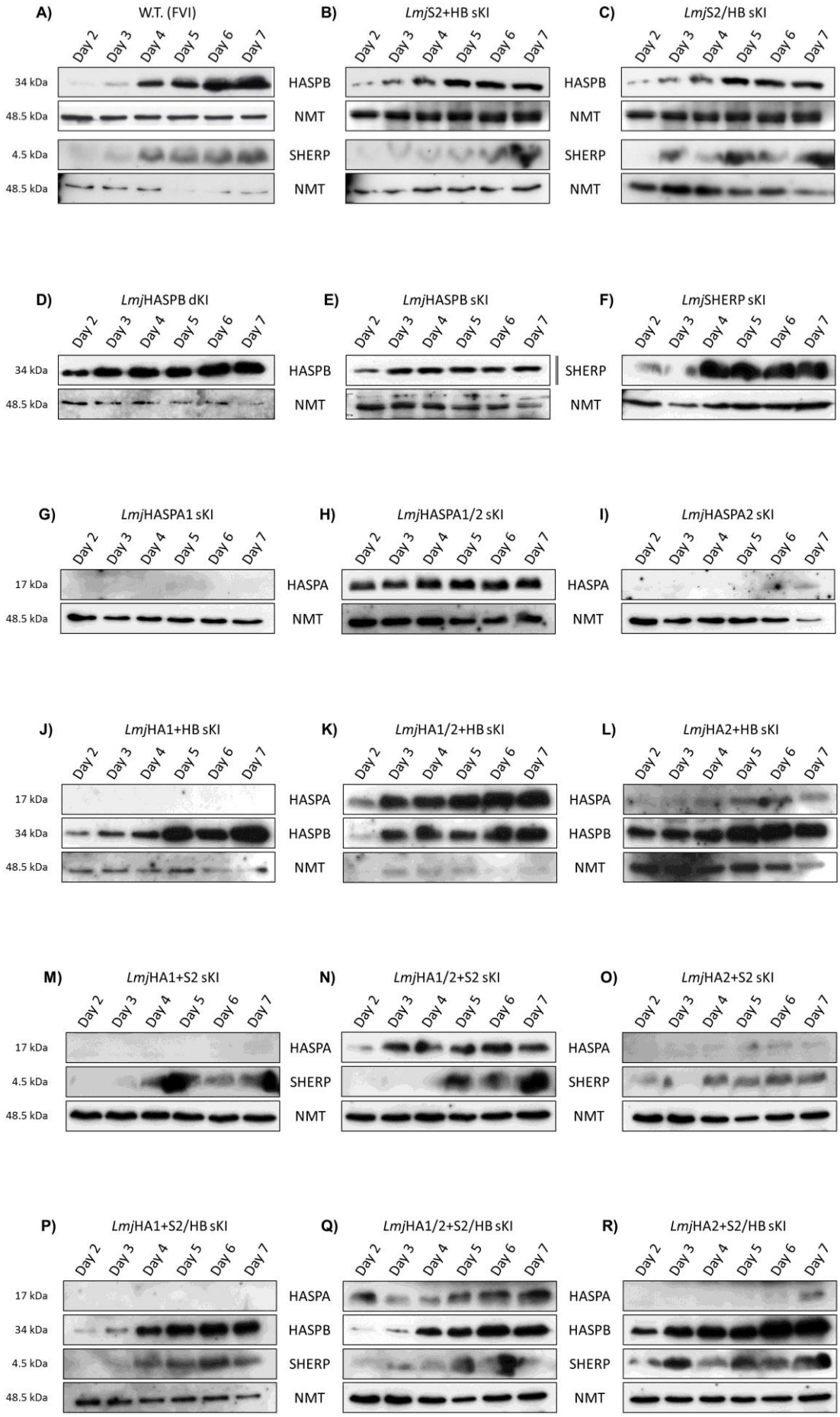

A)
W.T. (FVI)
B)
LmjS2+HB sKI
C)
LmjS2/HB sKI
D)
LmjHASPB dKI
E)
LmjHASPB sKI
F)
LmjSHERP sKI
G)
LmjHASPA1 sKI
H)
LmjHASPA1/2 sKI
I)
LmjHASPA2 sKI
J)
LmjHA1+HB sKI
K)
LmjHA1/2+HB sKI
L)
LmjHA2+HB sKI
M)
LmjHA1+S2 sKI
N)
LmjHA1/2+S2 sKI
O)
LmjHA2+S2 sKI
P)
LmjHA1+S2/HB sKI
Q)
LmjHA1/2+S2/HB sKI
R)
LmjHA2+S2/HB sKI
Day 2
Day 3
Day 4
Day 5
Day 6
Day 7
17 kDa
34 kDa
4.5 kDa
48.5 kDa
HASPA
HASPB
SHERP
NMT

Fig.3.9 –Western blot time-courses of mutant lines

The time-course Western blots were run for every mutant line to verify gene regulation and expression *in vitro*. HASPB was detected with the ab336, which was affinity purified on the central repeats, while HASPA was detected with a non-affinity purified HASP antibody, which recognized the conserved *N*-terminal region. SHERP was detected with the abSHERP. *N*-myristoyl transferase (NMT) was used as a constitutive loading control. The only inconsistencies found were in the HASPA1 construct, which did not express the protein at all, which could mean that HASPA1 expression is amastigote specific, and in the HASPA1/2 construct (H, K, N & Q), which had a much stronger expression HASPA than the HASPA2 construct (I, L, O & R). This suggested differential regulation for HASPA expression from the HASPA2 construct compared to the HASPA1/2 construct.

HASPA2 have the same ORF, they are indistinguishable at the protein level in FVI. mRNA expression data suggested HASPA2 upregulation early in procyclic stage with peak expression in metacyclics, while HASPA1 was upregulated only in metacyclic stage and continued to be expressed in amastigotes (*329*). These distinct expression patterns for HASPA1 and HASPA2 had not been previously verified at the protein level.

The generation of HASPA1 and HASPA2 replacement mutant lines, expressing only one of the two genes, allowed discrimination HASPA1 and HASPA2 expression individually. In the Western blot time courses described in section 3.3.1.4 HASPA2 was detected in promastigotes in HASPA2 only containing mutant lines as expected, but HASPA1 was not detected in mutant line promastigotes, although the mRNA analysis had suggested upregulation in metacyclics. It was possible that HASPA1 mRNA might be upregulated in metacyclics, but that the protein was only expressed in amastigotes. To address this hypothesis, FVI, *Lmj*cDNA16 dKO, *Lmj*cDNA16 sKI, *Lmj*HASPA1 sKI and *Lmj*HASPA2 sKI were infected into BALB/c mice and amastigotes were isolated from 9 weeks old lesions. The Western blot run from the amastigote lysates showed that HASPA1 was expressed in *Lmj*HASPA1 sKI amastigotes, which confirmed that it was amastigotes specific (Fig.3.10). HASPA1 and HASPA2 expression appears to alternate with one another between life-cycle stages with HASPA2 being promastigote-specific and HASPA1 amastigote-specific, suggesting stage specific function for both proteins. This hypothesis would require analysis of *Lmj*HASPA2 sKI amastigote lysates to ensure that no HASPA2 is expressed in amastigotes. However, due to the lack of lesion development in *Lmj*HASPA2 sKI, it was not possible to address this question. A detailed analysis of the capacity of HASPA1 and HASPA2 mutant lines for infection/lesion development is currently still under way.

3.3.1.6. Assessing HASPB surface localization *in vitro* by biotinylation assay

To confirm that the observed surface localization of HASPB in Fig.5.1 in culture derived mutant parasites was not a false positive, since HASPB also localizes to the inner leaflet of the plasma membrane, a biotinylation assay was performed to label *Leishmania* surface proteins for extraction by streptavidin beads. Water soluble EZ-Link Sulfo-NHS-SS-Biotin from Thermo Scientific, which did not require any solvents, was used as the

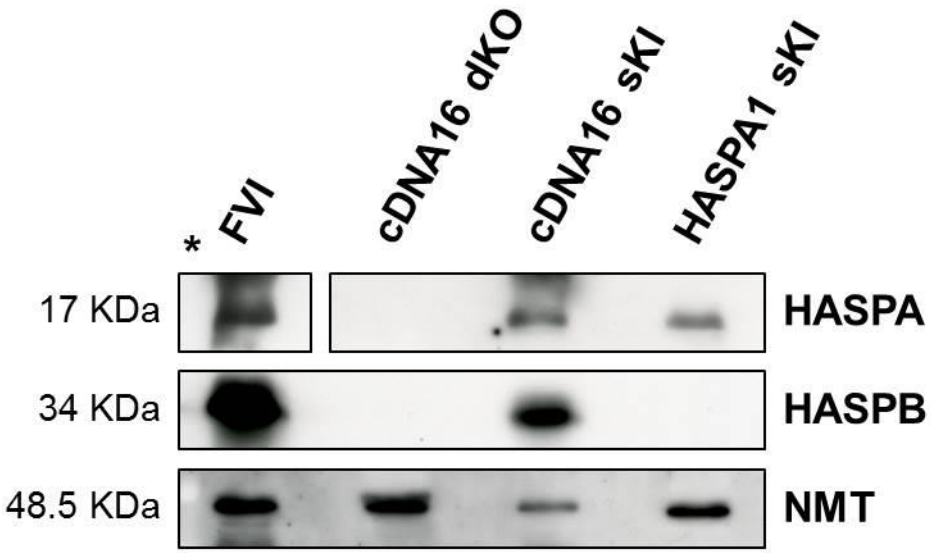


Figure 3.10 – Western Blot of Amastigote Lysates

Amastigotes isolated from 9 weeks old BALB/c mice lesions were lysed and their lysates run on an SDS-PAGE for a Western blot. The Western blot shows that HASPA1 is expressed in the *Lmj*HASPA1 sKI mutant line in amastigotes, although it was not detectable at the promastigote level (Fig.3.9). FVI and *Lmj*cDNA16 sKI also show HASPA detection, but also expression of HASPB, which is absent from the *Lmj*HASPA1 sKI mutant line as expected.

* The HASPA band of FVI was exposed for a shorter time (30 sec) compared to the three mutant lines (1.5 min), because the HASPB signal bled into the HASPA signal at longer exposures, obliterating the band

reagent for protein labelling (*335*). The *N*-hydroxysulfosuccinimide (NHS) ester group on this reagent reacts stably with the ε-amine of lysine residues in all accessible proteins leaving the biotin group exposed to react with streptavidin fixed to beads, allowing for labelled protein extraction by low speed centrifugation or gravity precipitation of the beads. Since HASPB is also present in the cytosol during its synthesis and trafficking to the cell surface, it was necessary to ensure that intracellular HASPB was not labelled by EZ-Link Sulfo-NHS-SS-Biotin. In particular, dead and compromised cells represented a source of cytosolic HASPB contamination. For this purpose, ~5x10^8 parasite derived from culture were stained with Sulfo-NHS-AMCA for a live/dead cell sort by Modular Flow Cytometer (MoFlo) prior to Sulfo-NHS-SS-Biotin labelling (*336*). The NHS group on Sulfo-NHS-AMCA reacts with the ε-amine of lysine residues, like the NHS group of Sulfo-NHS-SS-Biotin. While intact cells are impermeable for Sulfo-NHS-AMCA, compromised and dead cells have permeable plasma membranes and Sulfo-NHS-AMCA can enter these cells to label all the cytosolic protein. This causes compromised and dead cells to fluorescent very brightly, while intact cells only fluorescent weakly from the Sulfo-NHS-AMCA surface staining. This difference in fluorescence intensity can be exploited by a MoFlo separating brightly fluorescent dead cells from weakly fluorescent weakly fluorescent cells. The cell suspensions highly enriched in intact alive parasites were used in the biotinylation assay. Since Sulfo-NHS-AMCA had the same mode of binding to protein as Sulfo-NHS-SS-Biotin, there was a chance that labelling with Sulfo-NHS-SS-Biotin after Sulfo-NHS-AMCA staining would be inefficient. However, labelling with Sulfo-NHS-SS-Biotin remained efficient enough to isolate surface HASPB by streptavidin beads. Fig.3.11 shows that HASPB was detected in all sorted and Sulfo-NHS-SS-Biotin-labelled parasite sample, while detection of poly(A)-binding protein 1 (PAB1), which was used as a control for cytosolic contaminant, was almost undetectable in all sorted parasite samples. Only in the parasite lysate control, PAB1 was detected strongly. This suggested that the detected HASPB was from the cell surface of all examined parasite lines. This confirmed that HASPB targeting to the parasite cell surface was functioning as expected in culture derived mutant parasites. It is not clear, if the weaker HASPB signal in the *Lmj*S2+HB sKI and *Lmj*S2/HB sKI samples is suggestive of lower HASPB exposure on the cell surface of these lines or only an artefact. But it is suggestive that that both mutant lines containing SHERP should have a lower HASPB signal

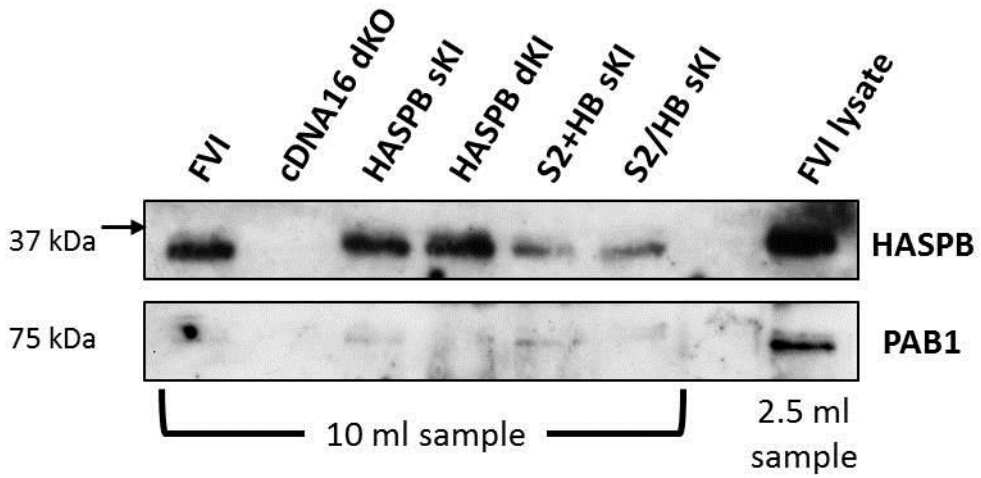


Figure 3.11 – Biotinylation Assay

The Western blots of the lysates of MoFlo sorted intact parasite cells confirm the presence of HASPB on the surface of live parasites, while the control poly(A)-binding protein (PAB1), which indicated cytosolic contamination, was barely detectable in the sorted samples. This proved that the HASPB signal was due to Sulfo-NHS-SS-Biotin-labelled HASPB from the cell surface and not from the cytosol. The lysate control showed that the antibodies used against HASPB and PAB1 were working.

compared to the HASPB only containing mutant lines, *Lmj*HASPB sKI and *Lmj*HASPB dKI, whose HASPB signal is comparable to the FVI one. However, since the antibody against GP63, which was used as a loading control, did not detect any protein (data not shown), it cannot be excluded that the difference in HASPB signal is due to differences in loaded material. Unfortunately, the assay could not be repeated to clarify this matter due to constraints in this study. However, it is clear that HASPB is trafficked normally to the cell surface in all tested mutant lines.

3.4. Parasite passage through mice

It was observed that *Leishmania* parasites lose their virulence and in particular, lose a clear HASPB expression pattern, over the course of excessive passaging through culture. Usually past-passage ten, HASPB becomes difficult to detect by antibodies on the cell surface and is increasingly shed from the parasite's cell surface (MacLean, L., personal communication). In order to restore virulence and controlled HASPB upregulation and expression, *Leishmania* parasites from day 6 p.i. were inoculated into BALB/c mice by subcutaneous injection into the right foot pad. Parasites were harvested after ~8 weeks by dissection of draining lymph nodes, if excessive foot pad swelling did not require earlier killing. Lymph nodes were ground up and inoculated into fresh M199 with antibiotics as required and incubated at 26 °C. Promastigotes began to show as early as day 2 and usually cultures were dense enough by day 7 for inoculation into fresh medium and for the setup of cryo-samples.

An interesting observation was made with respect to all mutants containing HASPA2 with or without HASPB and/or SHERP, but not with HASPA1. All mutants containing HASPA2 without HASPA1 had only minor lymph node swelling and no obvious foot pad swelling by the end of week 8 of mouse infection. They were also markedly slower (3 – 5 days) to produce promastigotes in M199 after lymph node harvest, if they grew up at all, compared to all other mutant strains. It had previously been shown that the expression patterns of HASPA1 (metacyclics and amastigotes) and HASPA2 (procyclics to metacyclics) are distinct at the mRNA level (*329*, *330*). This may suggest a function for HASPA1 in amastigotes that supports survival and/or differentiation into amastigotes and/or procyclics. Conversely, HASPA2 may have a negative effect on the transformation of metacyclics into amastigotes in the absence of HASPA1, impacting on the survival of amastigotes and/or their virulence in BALB/c mice, although HASPA2 expression is down-regulated in amastigotes. On the other

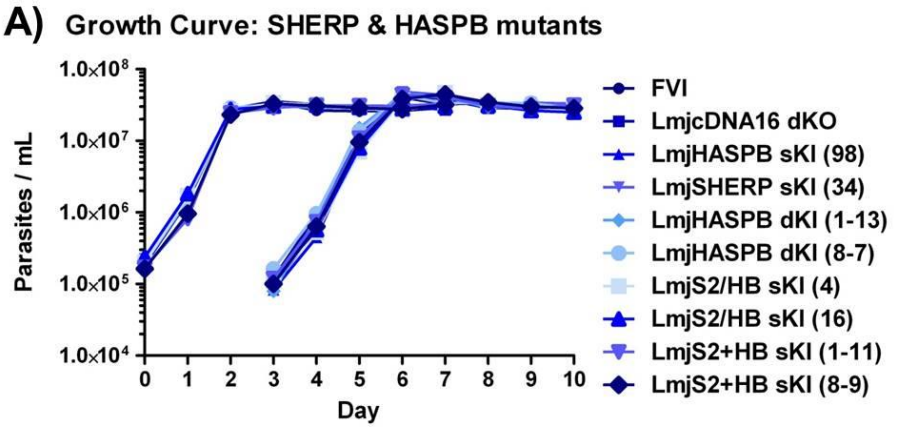
hand this seems puzzling, since the ORFs of HASPA1 and HASPA2 are identical. The major difference between HASPA1 and HASPA2 is their 3' UTR, which plays a role in their distinct expression patterns (Keen *et al*., unpublished). At this point, it is not clear whether HASPA1 and/or HASPA2 are post-translationally modified, which could explain a difference in function. However, since both ORFs have no known consensus motifs for protein modification and are translated in the cytosol never entering the ER and Golgi, where commonly protein modifications, like glycosylation, take place, it does not seem likely that HASPA1 and HASPA2 are post-translationally modified. HASPB, which contains two potential *N*-linked glycosylation sites and which is non-classically transported to the cell surface, is not modified *in vitro* by *N*-glycosylation or phosphorylation (*333*, *336*).

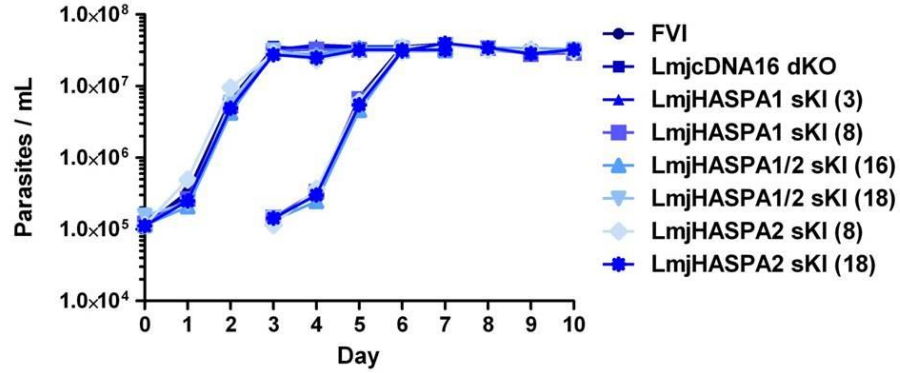
3.5. Growth Assay

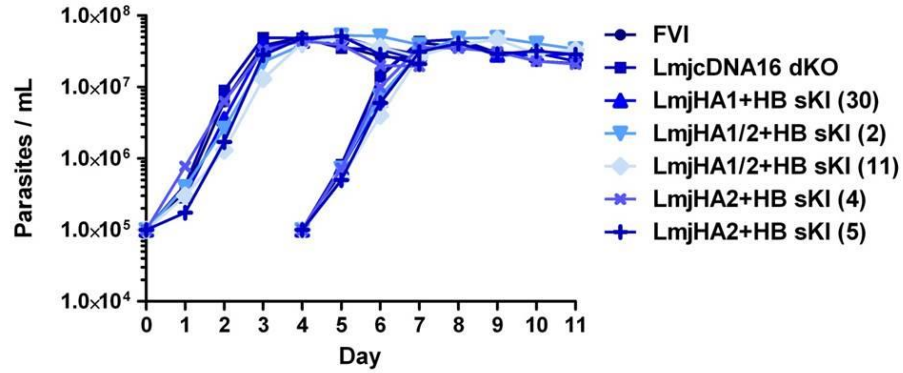
After mouse passage, parasite viability was checked by a growth assay. Parasites were inoculated into 10 ml of fresh M199 to a final concentration of 10^5 parasites / ml. Parasites were counted every ~24 h on a haemocytometer by making appropriate dilutions of small samples of culture in 1% formaldehyde in saline solution. Due to the quantity of sample tested in each growth assay, cultures were re-inoculated twice into fresh M199 by day 3 or 4 and parasites were counted again, instead of preparing parallel triplicate cultures (Fig.3.12). There was no obvious growth defect in culture for any of the generated mutant lines compared to FVI and, in general, mutant growth corresponded best to the wild type curve in the second and third repetition. Statistical analysis of growth rates between day 0 – 3 p.i. by t-test revealed, however, statistically significant differences ($P<0.001$) in log phase (Table 3.3; raw data in Appendix 7), while there were no significant differences between replicates or in stationary phase. Differences in growth rates were small, however, and the parental line (FVI) was not the strongest grower in all replicates. In general, growth rates varied for the individual lines between replicates, although never to statistical significance, and all lines reached similar parasite levels before entering into stationary phase between days 3 and 4 p.i., which suggested no significant fitness defect due to genetic manipulation in the generated mutant lines.

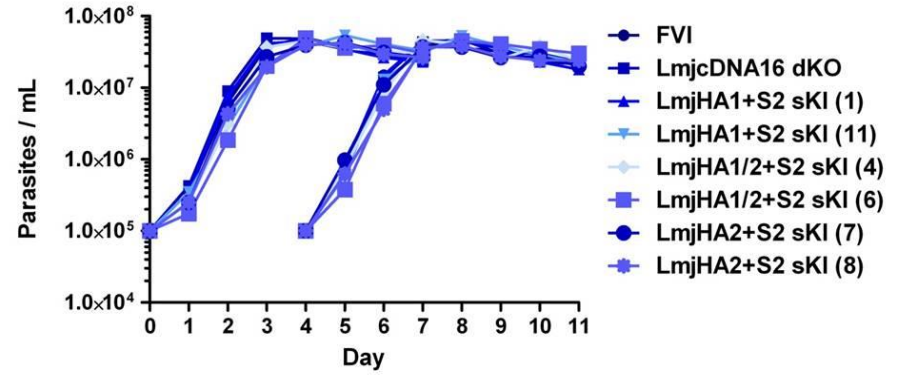
3.6. Conclusions

The aim of this chapter was to produce two clones for each of the 17 new mutant lines for sand fly infection assays that were thoroughly checked for correct gene integration by PCR, Southern blot and qPCR and correct expression by Western blot. While the results for PCR, Southern blot, qPCR and Western blots for


A) Growth Curve: SHERP & HASPB mutants
Parasites / mL
1.0x10^8
1.0x10^7
1.0x10^6
1.0x10^5
1.0x10^4
0 1 2 3 4 5 6 7 8 9 10
Day
FVI
LmjcDNA16 dKO
LmjHASPB sKI (98)
LmjSHERP sKI (34)
LmjHASPB dKI (1-13)
LmjHASPB dKI (8-7)
LmjS2/HB sKI (4)
LmjS2/HB sKI (16)
LmjS2+HB sKI (1-11)
LmjS2+HB sKI (8-9)


B) Growth Curve: HASPA mutants
Parasites / mL
1.0x10^8
1.0x10^7
1.0x10^6
1.0x10^5
1.0x10^4
0 1 2 3 4 5 6 7 8 9 10
Day
FVI
LmjcDNA16 dKO
LmjHASPA1 sKI (3)
LmjHASPA1 sKI (8)
LmjHASPA1/2 sKI (16)
LmjHASPA1/2 sKI (18)
LmjHASPA2 sKI (8)
LmjHASPA2 sKI (18)


C) Growth Curve: HASPA + HASPB mutants
Parasites / mL
1.0x10^8
1.0x10^7
1.0x10^6
1.0x10^5
1.0x10^4
0 1 2 3 4 5 6 7 8 9 10 11
Day
FVI
LmjcDNA16 dKO
LmjHA1+HB sKI (30)
LmjHA1/2+HB sKI (2)
LmjHA1/2+HB sKI (11)
LmjHA2+HB sKI (4)
LmjHA2+HB sKI (5)


D) Growth Curve: HASPA + SHERP mutants
Parasites / mL
1.0x10^8
1.0x10^7
1.0x10^6
1.0x10^5
1.0x10^4
0 1 2 3 4 5 6 7 8 9 10 11
Day
FVI
LmjcDNA16 dKO
LmjHA1+S2 sKI (1)
LmjHA1+S2 sKI (11)
LmjHA1/2+S2 sKI (4)
LmjHA1/2+S2 sKI (6)
LmjHA2+S2 sKI (7)
LmjHA2+S2 sKI (8)

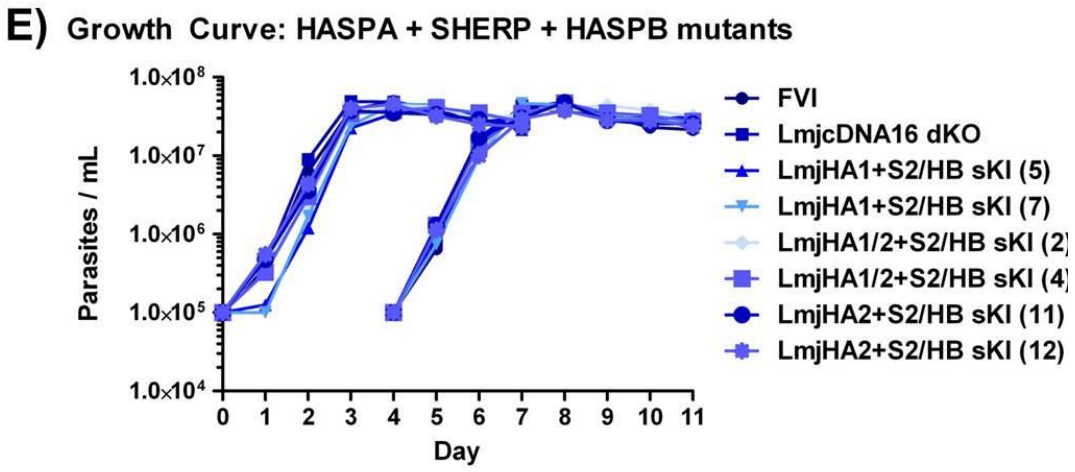


Figure 3.12 – Growth Assays

10 ml 1xM199 cultures were inoculated with *Leishmania* parasites to a final concentration of 10^5 cell/10 mL and were grown at 26 °C for seven days. The growth assay was repeated twice by splitting at day 3 or 4 into fresh 10 mL 1x M199 to 10^5 cell/10 mL. The results generated from the second and third round growth assays are shown here. FVI and *Lmj*cDNA16 dKO were used as controls. Although there are significant differences ($P<0.001$) in growth rates (days 0-3 p.i) according to t-test, differences were small (Table 3.2) and all strains reached similar parasite number by days 3-4 p.i., which suggested no fitness defect in culture due to genetic manipulation in any of the mutant lines compared to the parental line (FVI). Numbers in brackets are the number of the respective clone.

Table 3.3 – Statistics on growth rates (day 0 – 3)

Graph	Column	Samples	Mean	Std Dev	Std. Error	Median	Max	Min	Range	Norm. test	t-test
A)	G.R. I (D0-3)	10	1.733	0.055	0.017	1.721	1.861	1.649	0.212	Passed	P<0.001
	G.R. II (D0-3)	10	1.695	0.046	0.015	1.703	1.770	1.619	0.151	Passed	P<0.001
B)	G.R. I (D0-3)	8	1.841	0.042	0.015	1.844	1.900	1.771	0.130	Passed	P<0.001
	G.R. II (D0-3)	8	1.837	0.039	0.014	1.833	1.900	1.792	0.108	Passed	P<0.001
C)	G.R. I (D0-3)	7	1.866	0.119	0.045	1.901	1.957	1.623	0.335	Failed	P<0.001
	G.R. II (D0-3)	7	1.943	0.058	0.022	1.918	2.025	1.872	0.153	Passed	P<0.001
D)	G.R. I (D0-3)	8	1.881	0.111	0.039	1.881	2.078	1.758	0.320	Passed	P<0.001
	G.R. II (D0-3)	8	1.989	0.048	0.017	1.992	2.044	1.912	0.132	Passed	P<0.001
E)	G.R. I (D0-3)	8	1.937	0.082	0.029	1.934	2.065	1.805	0.259	Passed	P<0.001
	G.R. II (D0-3)	8	1.974	0.057	0.020	1.981	2.047	1.896	0.152	Passed	P<0.001

HASPB and SHERP indicated at least two clones per strain as fulfilling these criteria, Western blots for HASPA1 and HASPA2 expression in replacement mutants raised some questions. HASPA2 expression from the single HASPA2 construct in the replacement mutants was lower than expected as compared to HASPA expression in FVI as shown by McKean *et al.* (1997) (*334*). HASPA1 and HASPA2, however, cannot be distinguished at the protein level, because both genes have an identical ORF. HASPA1 was not detected at all in promastigotes in the single replacement mutants. This suggested exclusive amastigote expression for this protein and this was assessed by amastigote generation in BALB/c mice (Fig.3.10). The distinct 3' UTRs of HASPA1 and HASPA2 could play a role in HASPA2 being promastigote specific, while HASPA1 is amastigote specific. The expression levels of HASPA in the HASPA1 and HASPA2 replacement mutants , however, was comparable to the observation made by McKean *et al.* 1997 (*334*). It is not clear why the expression levels of HASPA were so different between the single HASPA1 or HASPA2 replacement mutant and the HASPA1 and HASPA2 containing mutants. Since HASPA1 was not expressed in promastigotes from the HASPA1 construct, it can be proposed that HASPA expression in the HASPA1/2 construct containing mutants is from the HASPA2 gene only. Since it is the same HASPA2 gene containing DNA fragment in the HASPA2 and HASPA1/2 constructs, this could mean that HASPA2 is overexpressed due to the unexpressed HASPA1 gene containing DNA fragment placed ahead of the HASPA2 gene in the HASPA1/2 construct, or, conversely, that HASPA1 is expressed in the presence of the HASPA2 gene immediately downstream of it, in the absence of the intervening HASPB and SHERP genes found in the parental line (FVI) locus. Since it is impossible to distinguish between HASPA1 and HASPA2 at the protein level, mRNA detection by targeting the distinct 3' UTRs would be required, which has the drawback that mRNA levels do not necessarily correspond to protein levels. Unfortunately, it was not possible to investigate this due to time constrains in this project. In culture, none of the 17 replacement mutants showed any significant growth defect compared to FVI, although significant differences were observed in the growth rates by t-test. All strains reached peak growth between days 3-4 p.i., however, which suggested that the genetic manipulations had no fitness disadvantage *in vitro*. Differences in growth rates could be attributed to inevitable variations in culture inoculations. With the exception of *Lmj*HASPB sKI, *Lmj*HA1+HB sKI and *Lmj*HA2+S2 sKI, where only one clone each survived the mouse passage, two clones for each mutant line have been successfully passaged through BALB/c mice. Parasite cryo-samples were sent to the Charles University in Prague for sand fly infection.

4. Chapter IV. – Investigating metacyclogenesis in the sand fly

4.1. Introduction

The work described in Sádlová *et al.* (2010) had shown that the mutant phenotype of *Lmj*cDNA16 dKO was only measureable in the sand fly vector rather than in *in vitro* culture. Therefore, it was necessary to analyse parasite metacyclogenesis in the natural context of the sand fly midgut (*146*). As the University of York did not have facilities for maintenance of appropriate sand fly vector colonies at that time, the mutants were taken to the laboratory of Prof. Petr Volf at the Charles University in Prague for analysis. This work done there was performed with the help of Dr. Jovana Sádlová.

The aim of this part of the study was the investigation of metacyclogenesis of the generated mutant lines within the natural context of the sand fly midgut and the identification of the key elements require for completion of this process. Identifying the particular contribution of the HASP and SHERP genes to metacyclogenesis completion was considered to give new insights into the yet unknown functions of these genes and their regulation. For the investigation, experimentally infected natural vector species of *L. (L.) major* were analysed with respect to the infection establishment in the midgut and the progression of parasite development over the course of 12 days post blood meal (PBM), which is detailed in the following section.

4.2. HASP and SHERP mutant development in the sand fly midgut

4.2.1. Artificial sand fly infections

For sand fly infections, the specific vector of *L. (L.) major* in the Middle-East, *Ph. (Ph.) papatasi*, was used initially. However, this was substituted by the more permissive sand fly vector of *L. (L.) major* in Western Africa, *Ph. (Ph.) duboscqi* – closely related to *Ph. (Ph.) papatasi* – because the *Ph. (Ph.) papatasi* colony became unstable and was unfit for further use during this project. There are important differences between these two sand fly species, which had to be considered in the comparison of results between experiments. *Ph. (Ph.) papatasi* (~2 mm in length) is considerably smaller than *Ph. (Ph.) duboscqi* (~3 mm in length), resulting in *Ph. (Ph.) duboscqi* having a larger midgut volume, known to be able to harbour more parasites than *Ph. (Ph.) papatasi* (personal communication from Petr Volf and Jovana Sádlová (2012)). This observation was reflected in the infection intensity

results of this study (see 4.2.4.). Another notable difference between these two sand fly species is that *Ph. (Ph.) papatasi* is known to be specific to *L. (L.) major* parasites with particular side chain galactosyl-modifications of the LPG, while *Ph. (Ph.) duboscqi* is permissive to all known *L. (L.) major* strains (*375*). As a consequence, intraspecific differences in the parasite's ability to survive within the sand fly were much more pronounced within *Ph. (Ph.) papatasi* (*157*). In general, however, this did not make a significant impact on the results in this study. In fact, our two positive controls, FVI and *Lmj*cDNA16 sKI, and the negative control, *Lmj*cDNA16 dKO, behaved very similarly in both sand fly species with no notable differences in their development. The only exception was in the parasite loads, which were higher in *Ph. (Ph.) duboscqi*, as mentioned above.

For artificial sand fly infections, early passage parasites were inoculated into 2 ml M199 with the appropriate antibiotics and incubated at 23 °C until late mid- to late log-phase (2 – 3 days p.i.; see 2.6.2. for protocol). Parasite growth was checked daily by light microscopy. On the day of sand fly infection, parasites were recovered from culture by centrifugation, washed and suspended in 1 ml saline solution. Parasites were counted by removal of 10 µl from the 1 ml parasite suspension followed by dilution (1:100) in 1% formaldehyde in saline solution and applying 10 µl of that dilution onto a haemocytometer. Parasites were counted on a light microscope at 400x magnification. Parasite density in saline solution was adjusted to 10^7 parasites / 300 µl saline solution, which was then diluted 1:10 in heat inactivated rabbit blood to a final concentration of 10^6 parasites / ml blood. 3 ml of infected blood, kept at 37 °C in glass feeders, were offered to individual sand fly colonies (150 – 250 individuals) in net-cages for 90 – 120 min. in the dark (Fig.2.2). Blood-fed sand flies were separated from unfed sand flies and were allowed to live for up to 12 days PBM. *Ph. (Ph.) duboscqi* proved to be a more efficient feeder on the artificial blood feeding system resulting in more well fed female sand flies than had been achieved initially with *Ph. (Ph.) papatasi*. This may have supported better establishment of infection in *Ph. (Ph.) duboscqi* by increased parasite ingestion, with *Ph. (Ph.) duboscqi* showing less infection clearing than in *Ph. (Ph.) papatasi*. As a result, infection loads were only compared between parasite strains infected into the same sand fly species (see 4.2.3.).

4.2.2. Sand fly dissections

Midgut dissections were initially performed at day 2, 5, 9 and 12 PBM to

correlate the results obtained, in particular for the control stains FVI, *Lmj*cDNA16 dKO and *Lmj*cDNA16 sKI, with the findings of Sádlová *et al.* (2010) both in *Ph. (Ph.) papatasi* and in *Ph. (Ph.) duboscqi*. After verification of developmental trends in the control strains, dissections were performed only at day 6 and 12 PBM, because these two time points were sufficient to observe the previously established parasite developmental trends.

Sand flies were collected for dissection by aspiration with a glass aspirator and were stunned by chilling in a small cups kept on ice. Whole midguts were dissected on glass slides in drops of sterile saline solution with needles and fine forceps under a magnifying glass. The sand fly midgut was divided into TMG and AMG and the parts placed separately in drops of saline solution on a glass slide and then covered with glass cover slips for light microscopic evaluation. A total of 19 *L. (L.) major* strains (FVI, *Lmj*cDNA16 dKO, *Lmj*cDNA16 sKI, *Lmj*HASPB sKI, *Lmj*HASPB dKI, *Lmj*SHERP sKI, *Lmj*S2+HB sKI, *Lmj*S2/HB sKI, *Lmj*HASPA1 sKI, *Lmj*HASPA1/2 sKI, *Lmj*HASPA2 sKI, *Lmj*HA1+HB sKI, *Lmj*HA1/2+HB sKI, *Lmj*HA2+HB sKI, *Lmj*HA1+S2 sKI, *Lmj*HA1/2+S2 sKI, *Lmj*HA2+S2 sKI, *Lmj*HA1+S2/HB sKI & *Lmj*HA1/2+S2/HB sKI) have been analysed in this fashion and a total of 2736 sand fly midguts have been dissected for this study. Only *Lmj*HA2+S2/HB sKI was not passaged through sand flies, because the cryo-samples did not survive transportation to Prague and no substitutes could be arranged in time.

4.2.3. Assessing *Leishmania* forward migration in the sand fly vector

Under the light microscope, parasite localization was firstly estimated by eye. Since the sand fly midgut was translucent, assessment of parasite localization was easily done at 400x magnification using a light polarizer in the set up for better contrast and resolution. Parasite localization was distinguished between localization in the EnS, the midgut lumen without stomodeal valve (SV) colonization (AMG – Cardia), weak SV colonization (weak SV col.) and heavy SV colonization (SV col.) (see Fig.1.9 for midgut anatomy). Parasites were mostly observed within the EnS at day 2 PBM before blood meal excretion and rarely at day 5 PBM, when blood meals were not completely defecated. In general, parasites were observed in the midgut lumen up to the cardia at day 5 PBM, but infrequently with SV involvement even for the positive controls, FVI and *Lmj*cDNA16 sKI. At day 6 PBM, SV colonization by FVI and *Lmj*cDNA16 sKI were significantly higher and by days 9 and 12 PBM it became clear that all mutant strains were not able to colonize the SV

efficiently, as observed for FVI, with the exception of *Lmj*cDNA16 sKI (Fig.4.1; raw data in Appendix 8). Occasionally, a few parasites appeared attached to the SV for some mutant strains, which was then classified as weak SV colonization. However, in general, parasites only migrated as far as the cardia (AMG – Cardia). Another observation was that in almost all mutant strains, parasites were found in the entire AMG up to the HG and very rarely even within the HG, which was distinct from the observation in FVI, where parasites accumulated at the anterior of the AMG. Conversely, TMG infections were visibly weaker for all mutant strains even at day 12 PBM, than in FVI and *Lmj*cDNA16 sKI, where TMGs appeared enlarged by day 12 PBM. It was further observed that when the TMGs of FVI and *Lmj*cDNA16 sKI infected midguts were analysed by day 12 PBM, using a cover slip to squeeze out the contents, the immobilized parasites were fanning out from the TMG in what was assumed to be the promastigote secretory gel (PSG) (Fig.4.2). This was never observed with any other mutant strain tested.

Leishmania leptomonads secrete filamentous proteophosphoglycans (fPPG) within the TMG (*161*). At sufficient fPPG concentration, the PSG gel is spontaneously formed to which nectomonads and leptomonads adhere (*255*). Since most mutant strains did not show TMG enlargement and did not show gel-immobilized parasites after squeezing on the TMG, it was reasonable to question whether the tested mutant lines were efficiently secreting PSG. It had previously been shown that *Lmj*cDNA16 dKO secreted PSG in culture (*146*), but it was not clear, if the same was true in the sand fly midgut. Since the PSG is considered to be required for transmission, it was important to test for PSG secretion in the midgut. Data addressing this question are presented in section 5.5.

4.2.4. Assessing *Leishmania* infection loads in the sand fly vector

Parasite infection loads were assessed for all dissection time points by estimating parasite loads under the light microscope; day 12 PBM samples were also quantitated by qPCR targeting kinetoplastid minicircle DNA to verify light microscopic estimates (Fig.4.3). Infection loads were scored as either uninfected, light (<100 parasites/gut), moderate (100-1000 parasites/gut), heavy (>1000 parasites/gut) or very heavy (>>1000 parasites/gut) infection (see 2.6.3 for details).

By light microscopy, significant differences were observed between strains

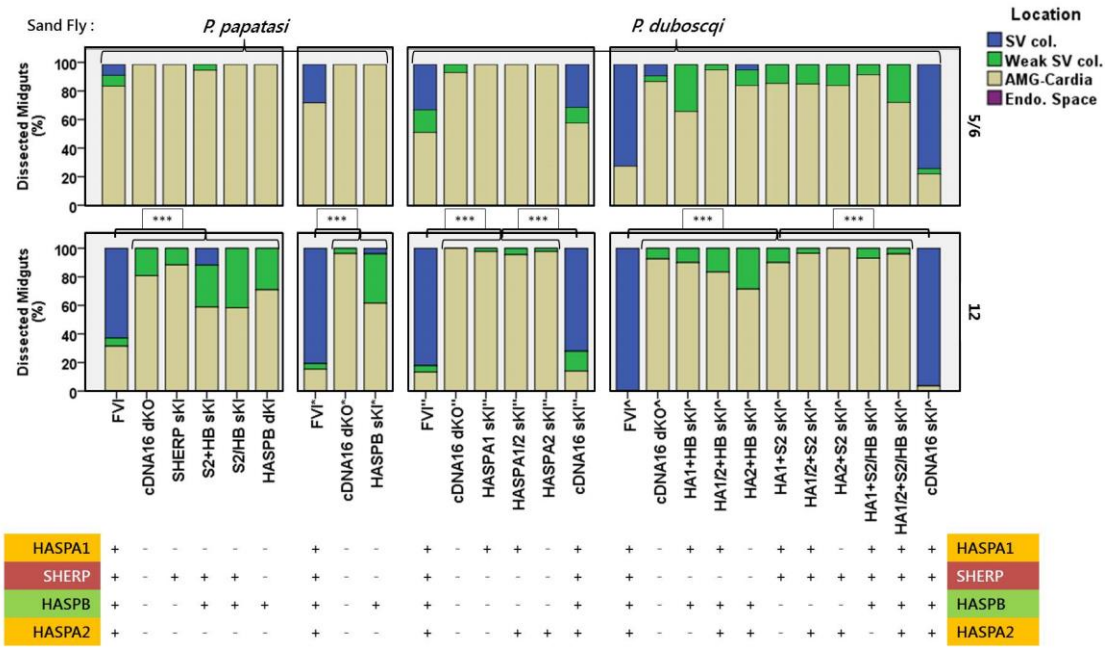


Fig.4.1 – Parasite localization in the sand fly midgut

The graphs show a steady increase of SV colonization (blue) for the positive controls, FVI and *Lmj*cDNA16 sKI, from day 5/6 to 12 PBM, while all other mutant stains show a significant incapacity (P<0.001) for efficient SV colonization by day 12 PBM in both, *Ph. (Ph.) papatasi* and *Ph. (Ph.) duboscqi*.

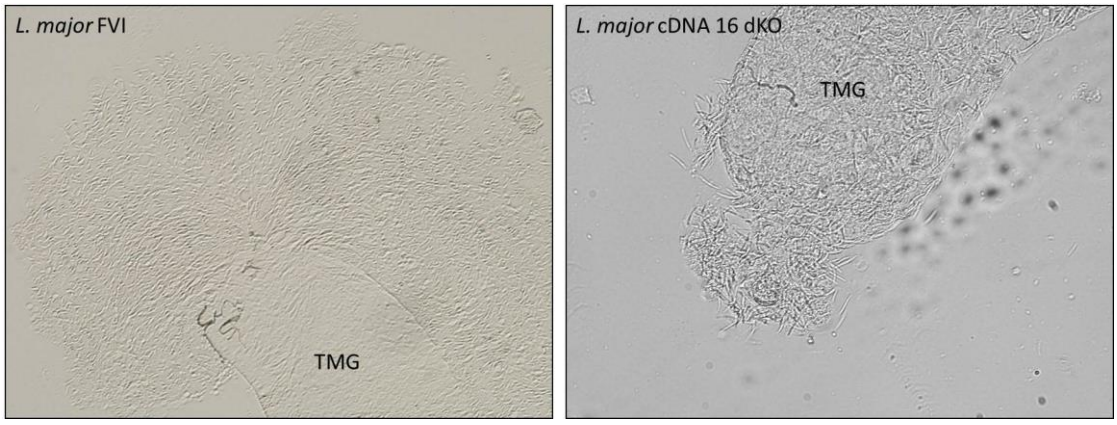


Fig.4.2 – Apparent presence and absence of PSG in the TMG in different L. (L.) major strains

Wet mounts of FVI and *Lmj*cDNA16 dKO showed that there was an apparent lack of PSG secretion in the null mutant. While immobilized parasites fanned out from a day 12 PBM FVI midgut, parasites from *Lmj*cDNA16 dKO were immediately free swimming and the fan shape of immobilized parasites was never observed. *Lmj*cDNA16 dKO is representative for all tested mutant strain with the exception of *Lmj*cDNA16 sKI (data not shown).

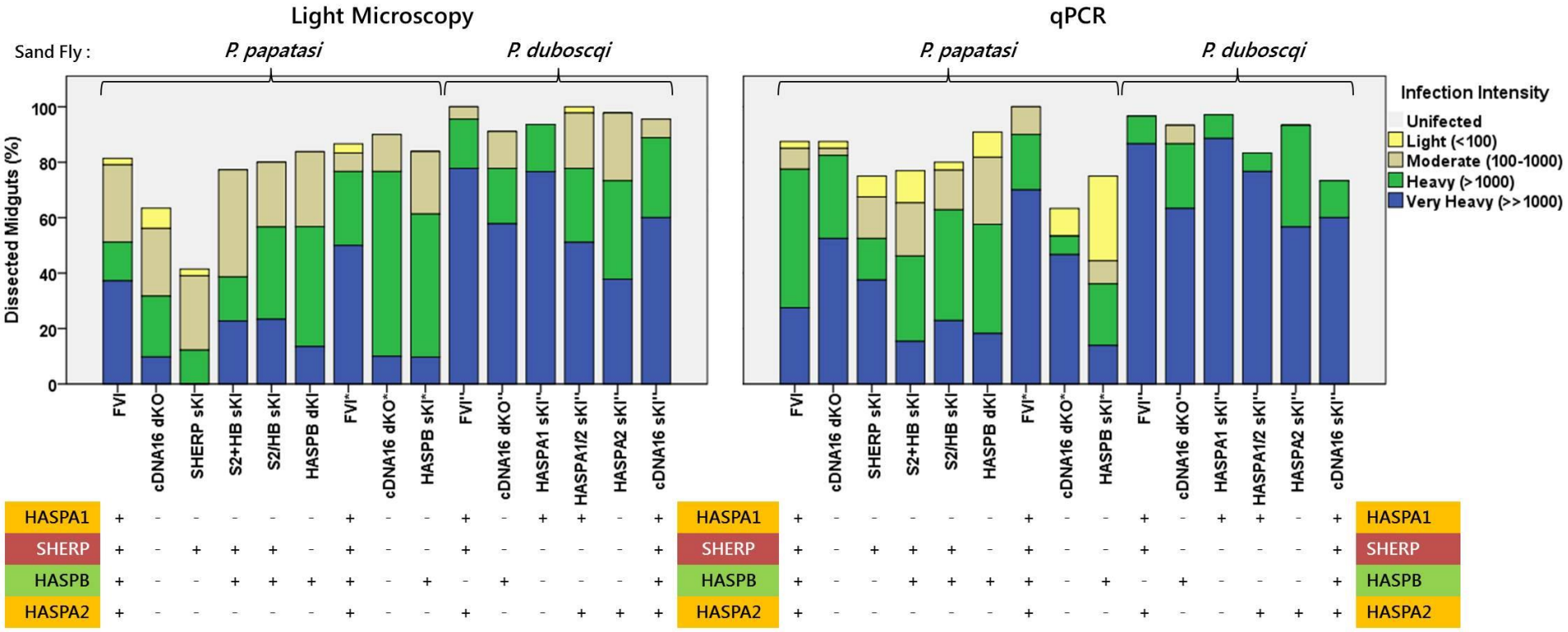


Fig.4.3 – Comparing light microscopic infection load results with qPCR results for day 12 PBM

With a few exceptions (in particular, *Lmj*cDNA16 dKO & *Lmj*SHERP sKI) there was a good correlation of trends in infection intensity between light microscopic and qPCR analysis. In general, light microscopic analysis underestimated true parasite loads with the exception in *Lmj*HASPB sKI. The qPCRs to assess infection loads were performed by Jan Votýpka at the Charles University, Prague, CZ. (note: the nine mutants are still missing from the qPCR analysis)

tested in *Ph. (Ph.) papatasi* and *Ph. (Ph.) duboscqi* (Fig.4.4). FVI and *Lmj*cDNA16 dKO both survived better and achieved higher infection loads in *Ph. (Ph.) duboscqi* compared to *Ph. (Ph.) papatasi* (P<0.001 for both), which was expected from previous observations in the Volf lab. Therefore, experiments done in *Ph. (Ph.) papatasi* were considered separately from those done in *Ph. (Ph.) duboscqi* for statistical analysis of infection intensity.

All tested parasite strains showed efficient survival in both vectors. For most strains, there were steady increases on average in parasite load from day 5/6 PBM to day 12 PBM, of which some reached statistical significance (FVI [P=0.022]; *Lmj*HASPB dKI [P=0.011] & *Lmj*HASPB sKI [P=0.016] in *Ph. (Ph.) papatasi* and FVI [P<0.001]; *Lmj*cDNA16 sKI [P=0.038]; *Lmj*HASPA1 sKI [P<0.001]; *Lmj*HASPA1/2 sKI [P=0.001] & *Lmj*HASPA2 sKI [P=0.007] in *Ph. (Ph.) duboscqi*). A few exceptions (*Lmj*cDNA16 dKO in *Ph. (Ph.) papatasi* and *Lmj*HA1/2+S2 sKI, *Lmj*HA2+S2 sKI and *Lmj*HA1+S2/HB sKI in *Ph. (Ph.) duboscqi*) showed non-significant decreases from day 5/6 PBM to day 12 PBM These never reached statistical significance with the exception of *Lmj*HA1+S2/HB sKI (P=0.033) (raw data in Appendix 9). These decreases in parasite numbers are not considered to confer true fitness disadvantage, however, since these experiments are prone to strong variations and the data are cumulative of three repeats that were initiated at weekly intervals (Fig.4.5).

The only mutant strain which survived poorly in *Ph. (Ph.) papatasi* by light microscopic analysis was *Lmj*SHERP sKI. This line was observed in only ~40% of all dissected midguts by day 12 PBM in *Ph. (Ph.) papatasi*, while in *Ph. (Ph.) duboscqi*, *Lmj*SHERP sKI did not show any lack of survival (>80% by day 12 PBM) compared to other mutant strains. Whether this is a vector species-specific phenotype, or an artefact caused by the *Ph. (Ph.) papatasi* colony becoming unreliable, is not clear, but the qPCR data in contrast showed a ~70% survival rate for *Lmj*SHERP sKI in *Ph. (Ph.) papatasi* (Fig4.3). Although the light microscopic approach was shown to be prone to underestimation compared to qPCR, which is reflected in the data in Fig.4.3, it is not likely that parasites were completely missed in the light microscopic analysis in ~30% of dissected midguts for *Lmj*SHERP sKI. A re-run in *Ph. (Ph.) papatasi* would be required to address this issue, but due to the lack of an available colony this could not be done in this study.

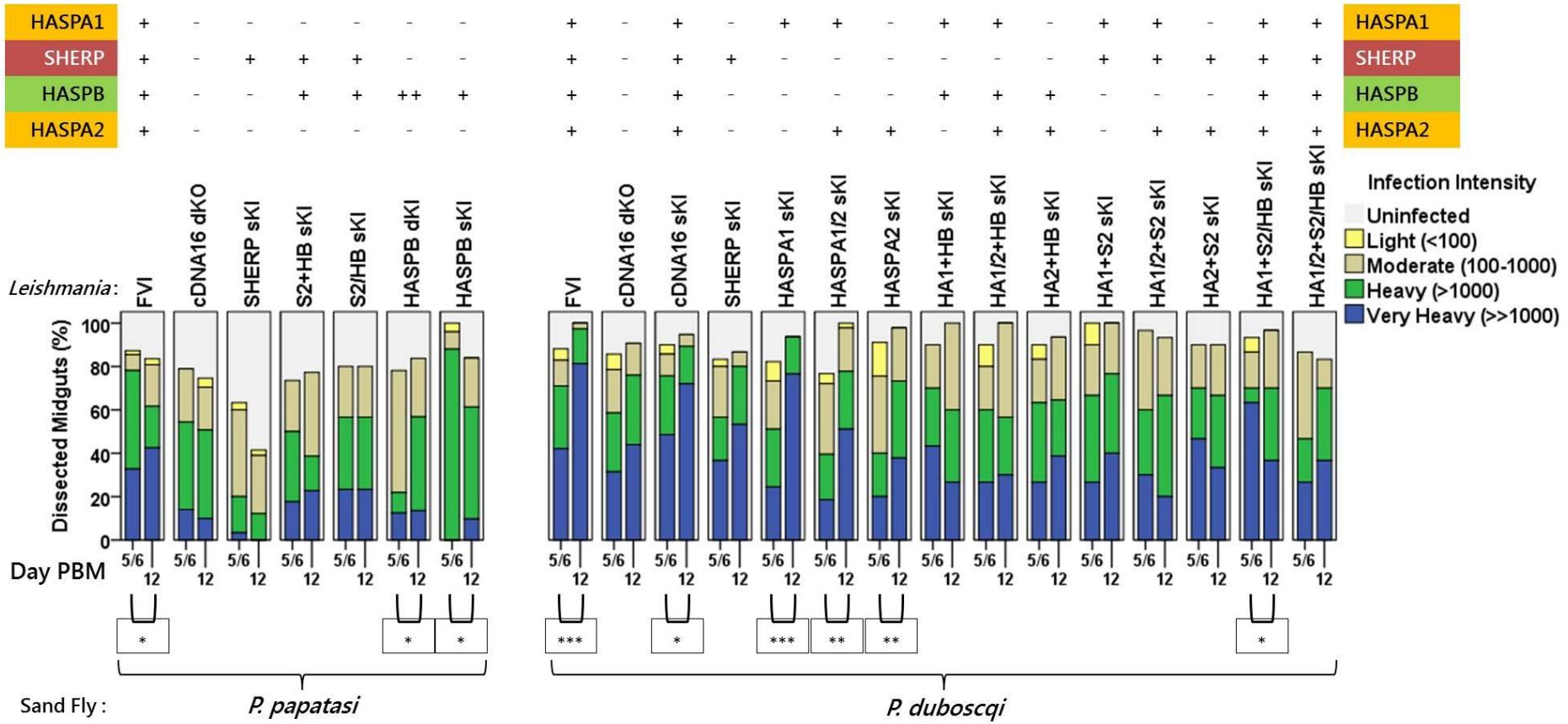


Fig.4.4 – Parasite infection load by light microscopy in the sand fly midgut

In most cases, parasite loads increased over time as expected, which was more obvious in *Ph. (Ph.) duboscqi* than in *Ph. (Ph.) papatasi*, although infection load increases were strongest in the two positive controls, FVI and *Lmj*cDNA16 sKI ($P<0.001$). The few cases where in particular very heavy infections (blue) cases reduced over time usually did not show a significant reduction with the exception in *Lmj*HA1+S2/HB sKI. The dramatic difference of *Lmj*SHERP sKI survival in *Ph. (Ph.) papatasi* compared to *Ph. (Ph.) duboscqi* was not supported by the qPCR analysis.

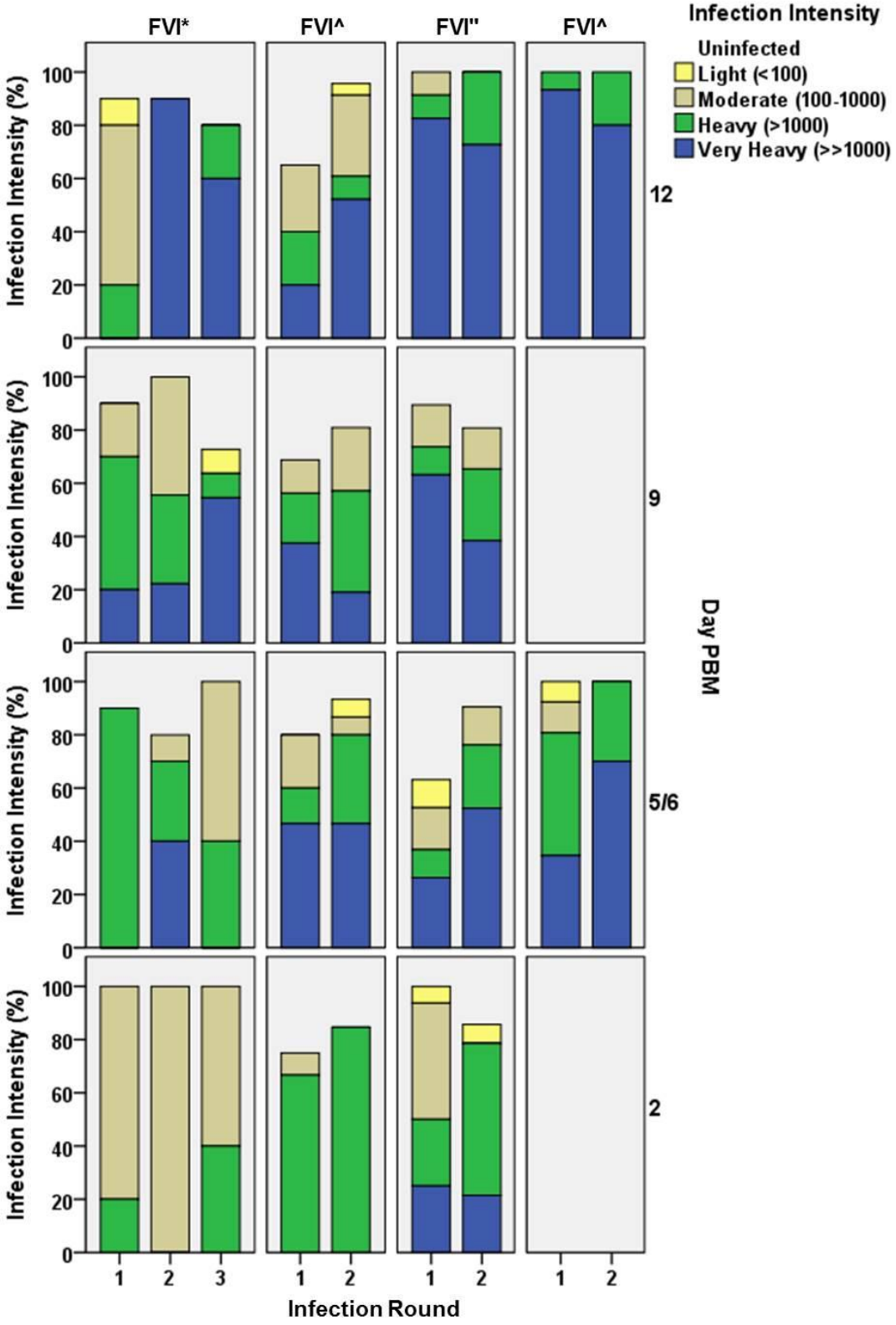


Fig.4.5 – Variability of observed infection intensity

The graphs show, exemplified on FVI, how infection intensity in midguts can vary from infection round to infection round impacting on the cumulative data set for each day PBM. For example, at day 12 PBM in FVI* the infection round 1 is poor compared to 2 & 3 and significantly different ($P<0.001$).

In general, the statistical comparison of the infection intensity data based on light microscopy and qPCR showed relatively good correlation of the two data sets, although in most cases the light microscopic approach underestimated the qPCR results (Fig.4.3). In some case there were significant difference between strains checked at day 12 PBM by light microscopy and qPCR (FVI [P=0.006]; *Lmj*cDNA16 dKO [P<0.001]; *Lmj*SHERP sKI [P<0.001]; *Lmj*cDNA16 dKO* [P<0.001]; *Lmj*HASPB sKI [P=0.002]; *Lmj*cDNA16 sKI [P<0.001], *Lmj*HASPA1/2 sKI [P=0.001] & *Lmj*HASPA2 sKI [P=0.022]), yet the observed trends by both methods were for most checked strains similar with only a few exceptions (*Lmj*cDNA16 dKO, *Lmj*SHERP sKI and *Lmj*cDNA16 dKO*), where significant differences were observed (P<0.001 for each) (raw data in Appendix 10).

With a few exceptions, there was a varying degree of significant difference between the positive controls, FVI and *Lmj*cDNA16 sKI, and the other mutant strains in the light microscopic data, which suggested that FVI and *Lmj*cDNA16 sKI proliferated and survived better in the sand fly midgut than all other replacement mutant lines. This trend, however, was not so clear cut in the qPCR data, where, for example, *Lmj*cDNA16 dKO produced more very heavy infections in the second infection experiment than FVI by day 12 PBM (Fig.4.3). Despite the huge variation in the infection intensity data, it can be said that parasite survival is generally not impaired in the mutant lines.

4.2.5. Analysing *Leishmania* parasite morphology and metacyclogenesis in the sand fly vector

AMGs and TMGs in drops of saline solution were smeared individually onto glass slides by pressing cover slips onto them. The samples were then allowed to dry before fixing the parasites together with the midgut material by applying 100% methanol at room temperature to the slides, which were then left to dry again. For analysis of parasite morphology, fixed slides were stained with Giemsa's stain for 20 min. and imaged under the light microscopy with a 100x oil-immersion lens. The flagellar length, cell body length and cell body width were measured on the parasite images and they were classified either as procyclic, nectomonad, leptomonad or metacyclic promastigotes according to established criteria (*157*, *365*), which had been used previously by Sádlová *et al.* (2010).

Fig.4.6 shows the complete, unaveraged data set for day 5/6, 9 and 12 PBM

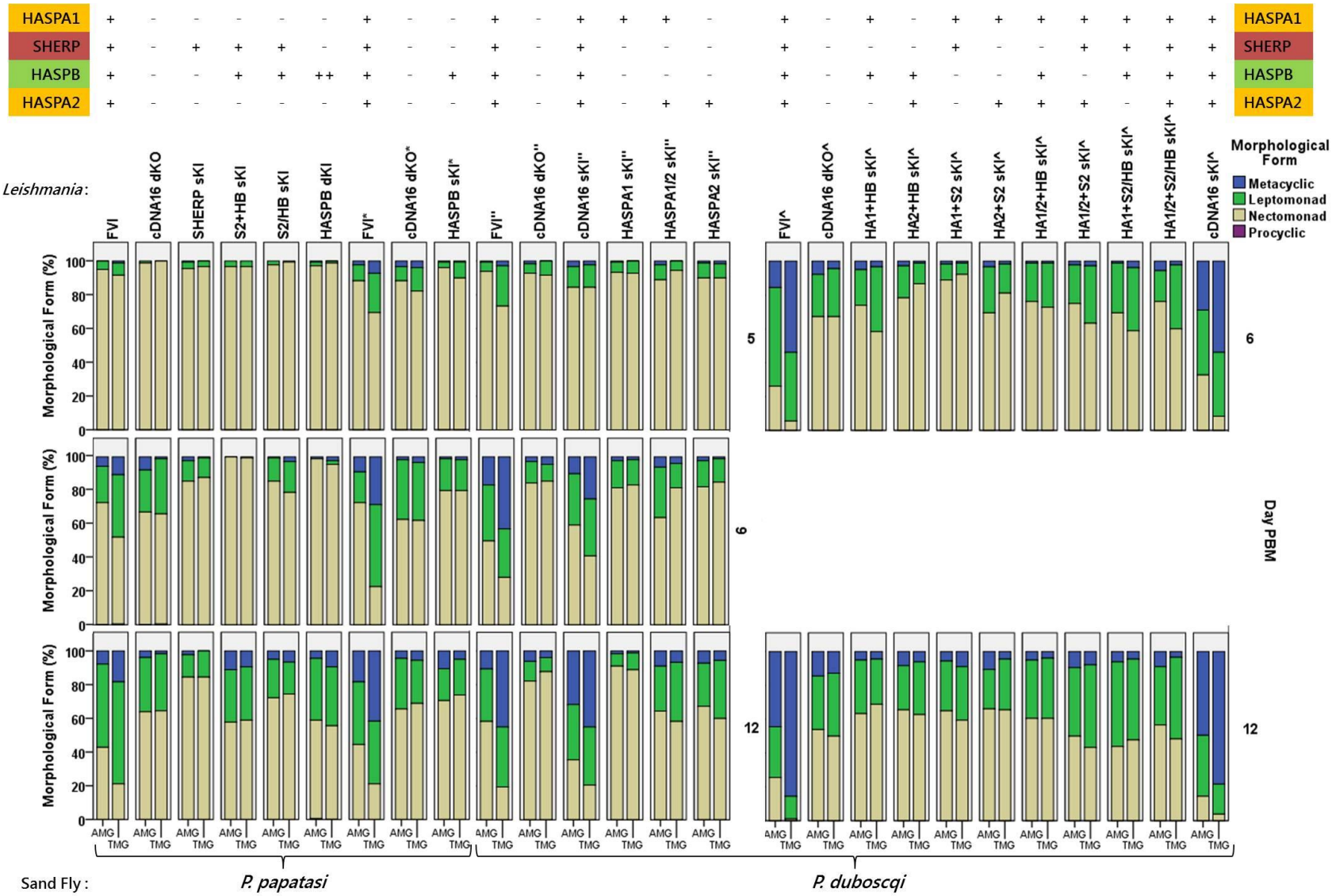

HASPA1
SHERP
HASPB
HASPA2
Leishmania:
FVI
cDNA16 dKO
SHERP sKI
S2+HB sKI
S2/HB sKI
HASPB dKI
FVI*
cDNA16 dKO*
HASPB sKI*
FVI''
cDNA16 dKO''
cDNA16 sKI''
HASPA1 sKI''
HASPA1/2 sKI''
HASPA2 sKI''
FVI^
cDNA16 dKO^
HA1+HB sKI^
HA2+HB sKI^
HA1+S2 sKI^
HA2+S2 sKI^
HA1/2+HB sKI^
HA1/2+S2 sKI^
HA1+S2/HB sKI^
HA1/2+S2/HB sKI^
cDNA16 sKI^
Morphological Form
Metacyclic
Leptomonad
Nectomonad
Procyclic
Morphological Form (%)
Day PBM
5
6
12
AMG
TMG
Sand Fly :
P. papatasi
P. duboscqi

Fig.4.6 – Complete morphology data set of sand fly midgut derived parasites

60 parasites per AMG and TMG, respectively, were measured per midgut smear slides per dissection day per strain. Three midgut smear slides were analysed per strain per dissection day, which amounts to 180 measured parasites per 100% bar in the graph. A total of 24120 individually measured parasites are represented in the data. Clear increases of metacyclics (blue) are observed in the positive controls, FVI and *Lmj*cDNA16 sKI, while this was not the case in all other mutant strains. Merely leptomonad generation was observed to varying degree among the mutants suggesting a lack of metacyclogenesis completion in the all mutant strains.

for the morphological analysis of all strains (raw data in Appendix 11). The data show that parasite differentiation took place in all tested strains from day 5/6 PBM towards day 12 PBM, but the numbers of generated leptomonads and metacyclics over time varied strongly between the different lines (Fig.4.7B). With the exception of *Lmj*HASPA1 sKI and *Lmj*HA1+HB sKI, all tested strains showed a significant increase in leptomonads and metacyclics over time. However, with the exception of FVI and *Lmj*cDNA16 sKI, none of the mutant strains had generated metacyclics efficiently by day 12 PBM (Fig.4.7). Not even the *Lmj*HA1/2+S2/HB sKI mutant, which technically had all the cDNA16 genes replaced back into the original cDNA16 locus with the exception of SHERP1 (98.8% gene and 100% ORF identity to SHERP2), generated metacyclics efficiently suggesting that it takes more than just replacing genes back into the former cDNA16 locus to rescue metacyclic generation. Looking at FVI and *Lmj*cDNA16 sKI, it was observed that a gradient of metacyclics was clearly established by day 12 PBM from AMG towards TMG with a significantly larger amount of metacyclics in the TMG compared to the AMG ($P<0.001$) (Fig.4.7A). Considering that metacyclics are the mammalian infective forms, which need to localize as far forward in the digestive system as possible, it made sense that such a metacyclic gradient towards the SV should exist. The fact that none of the mutant strains tested was able to establish this gradient may be a direct consequence of the lack of efficient metacyclogenesis. Leptomonad generation, however, seems to improve with the increase in the number of replaced genes (Fig.4.8), although the data do not clearly indicate which specific gene(s) were beneficial to improved leptomonad generation. It could be the HASPA1 and HASPA2 may be required, because the data from the individual infection rounds showed better leptomonad generation in all three infection compared to mutants containing only HASPA1 or HASPA2 (Fig.4.9). This may be due to the increase HASPA2 expression in mutant lines containing HASPA1 and HASPA2 in a single construct (Fig.3.9). However, improved leptomonad generation could be an artefact related to infected sand fly species too. Unlike metacyclics, leptomonads did not form a gradient towards the TMG, although one of the leptomonads' main functions is the secretion of the PSG plug in the TMG, potentially, helping the parasites to colonize the TMG strongly. In particular, *Lmj*HA1/2+S2 sKI, *Lmj*HA1+S2/HB sKI and *Lmj*HA1/2+S2/HB sKI showed high levels of leptomonads in both AMG and TMG. Perhaps the accumulation of leptomonads was due to the lack of conversion of leptomonads into metacyclics, as observed in FVI and *Lmj*cDNA16 sKI.

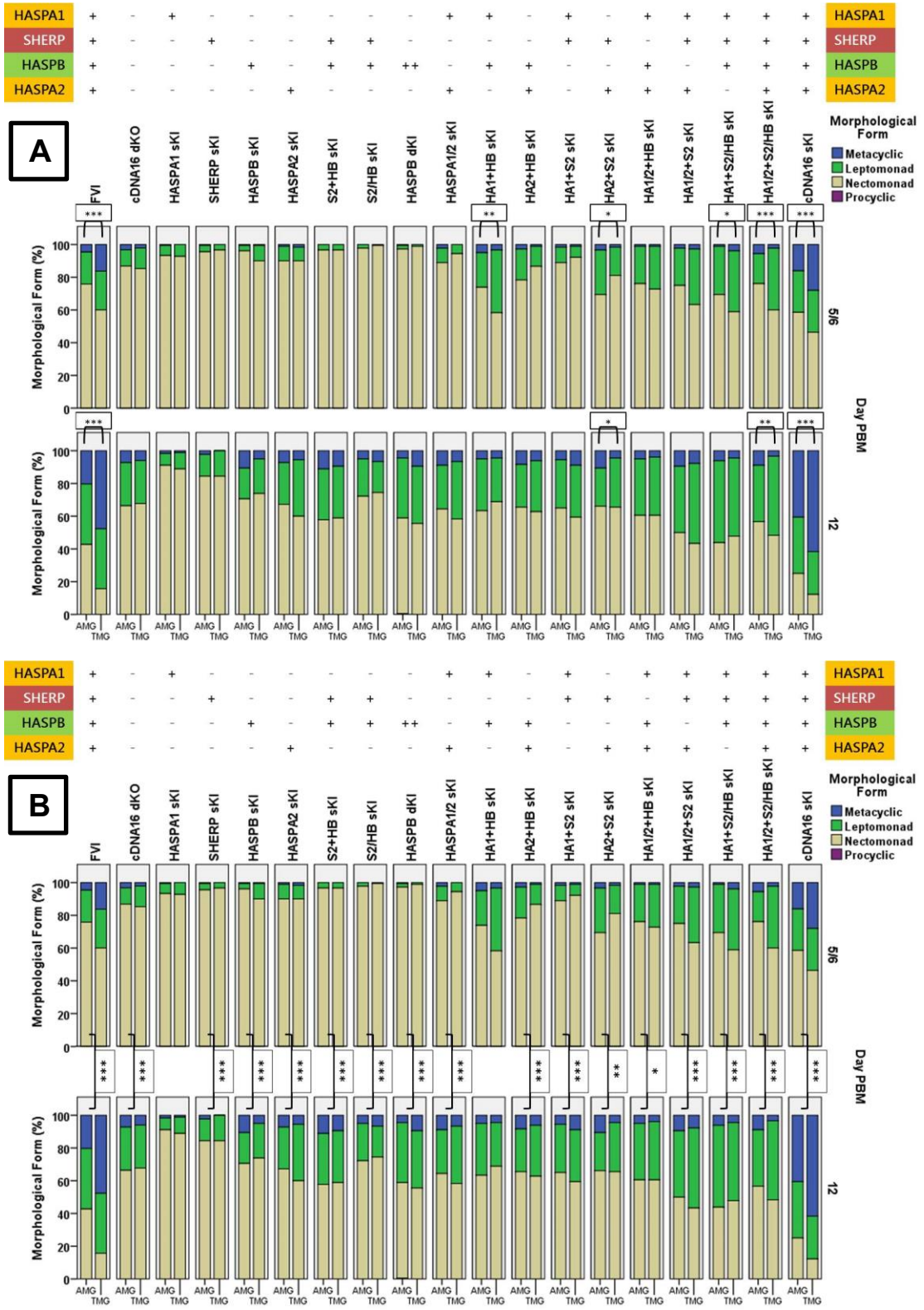


Fig.4.7 – Representation of statistically significant difference in parasite form distribution of summarized data

A) Differences in parasite form distribution between AMG and TMG. Only FVI and *Lmj*cDNA16 sKI showed a significantly different load of metacyclics between AMG and TMG (P<0.001). B) Parasite differentiation over time. Most strains showed a significant increase (P<0.001) in parasite differentiation from day 5/6 PBM towards day 12 PBM.

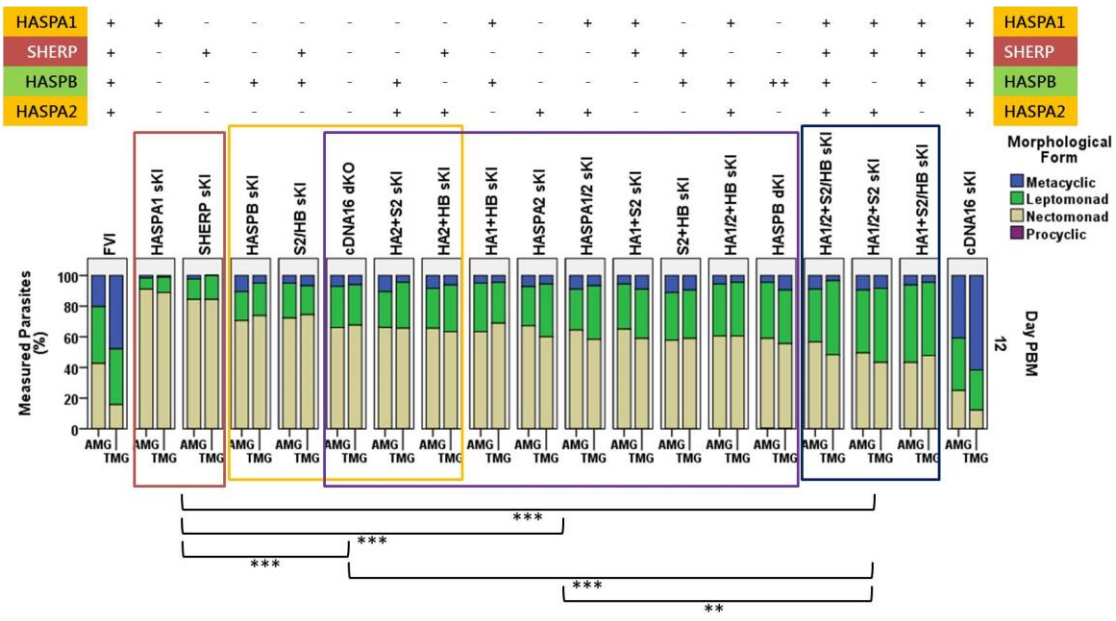


Fig.4.8 – Parasite morphology at day 12 PBM in the sand fly midgut

The boxes group strains, which showed no statistical significant difference in parasite differentiation capacity between on another. Parasites in the red and black boxes at the extremes of the spectrum always showed significant differences to those in the other boxes (P=0.006 to P<0.001), while the parasites in the orange and purple box did not show an statistically relevant difference between one another. Although individual replacement of HASP and/or SHERP genes did not recover efficient metacyclic generation a significant increase in leptomonad generation was observed, which correlated with the increased number of replaced HASP and SHERP genes. However, it remains unclear, which HASP and/or SHERP genes in particular contributed to the increased leptomonad generation.

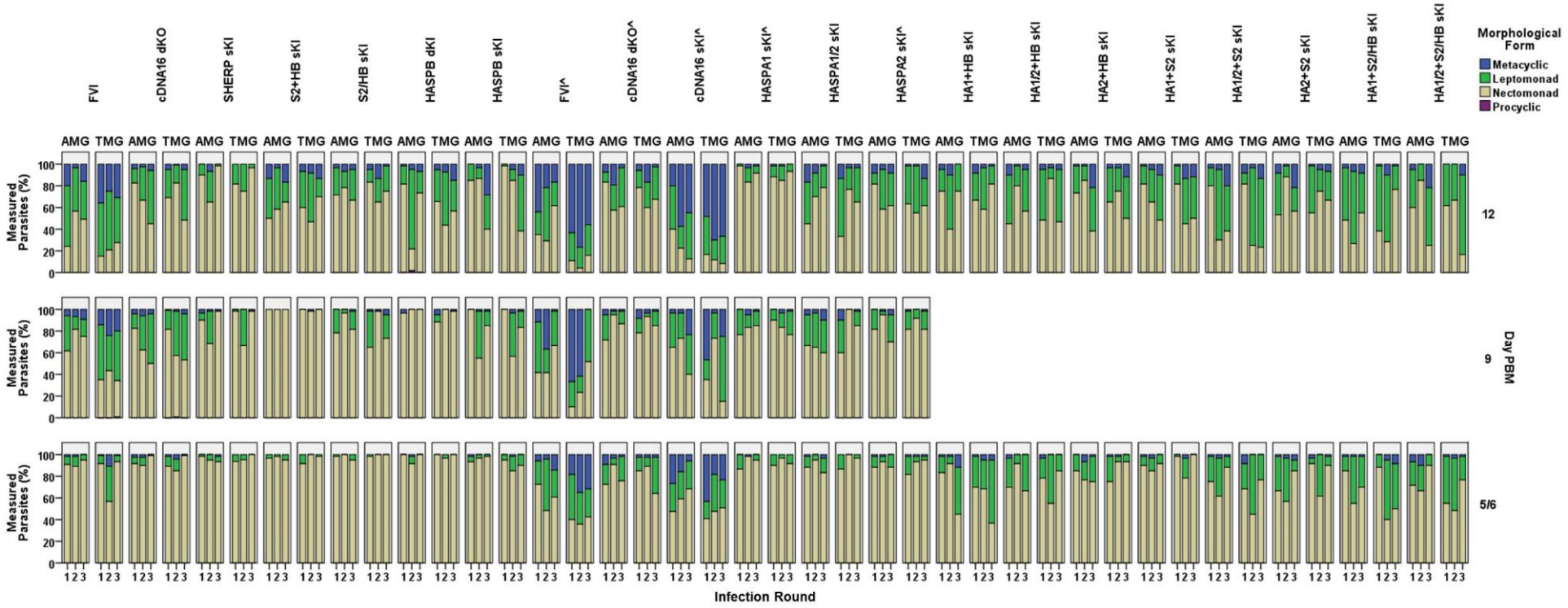


Fig.4.9 – Variability of observed parasite morphology between infection rounds

The graphs show the entire morphology data split into the individual infection rounds. In particular, at day 12 PBM strong variations in the ratio of observed morphological forms can be seen between the samples from individual infection rounds. Mutant lines with multiple HASP and/or SHERP genes replaced into the cDNA16 locus were able to produce leptomonads more efficiently.

4.2.6. Effects of parasite morphological plasticity on morphological data

There is a lack of stage-specific parasite protein markers that allow parasite stage distinction from one another, with the exception of HASPB and SHERP, which are considered specific for the metacyclic promastigote stage. HASPB and SHERP, however, could not be used as stage specific markers in this study, since they are the proteins of interest and are not expressed in all mutant stains tested. This left only cell body and flagellum measurement to distinguish the parasite stages. The criteria for parasite form distinction established by Walters *et al.* (1993) (*365*) and Cihakova & Volf (1997) (*157*), which had also been applied by Sádlová *et al.* (2010) (*146*), were used in this study (see 2.2.7.). Using only cell body and flagellum measurements for classification of parasite forms, however, bears an inherent problem. The change from one parasite form to another is gradual, which causes an unbroken gradient of measurements to appear with no clear distinction between parasite forms (Fig.4.10). For example, nectomonads are primarily distinguished from all other forms by having a cell body length ≥14 µm, however, cell body length measurements clustered strongly around the 14 µm mark putting into question if a parasite that measured 13.99 µm was indeed different from one that measures 14.00 µm (Fig.4.10A). The same applies for the distinction between leptomonads and metacyclics by measurement, where the main criterion for metacyclics is a coefficient >2, when flagellum length is divided by cell body length (Fig.4.10B). Kinetoplastids in general have a very flexible cell morphology and it is debatable whether measurements as the sole tool are sufficient for the distinction of parasite developmental stages (personal communication. J. Lukeš, Kinetoplastid Molecular Cell Biology Meeting V, April 2013).

To address this uncertainty in the morphology data, parasites falling close to break-off points were transiently excluded from consideration or thresholds were moved up or down the scale to see, whether these changes would affect the overall results. The data shown in Fig.4.11A-F indicated that overall there was little variation in the ratios of the different morphological parasite forms in the data by moving thresholds slightly up and down the scale or even excluding data points close to the artificial thresholds. Statistical re-analysis of the altered data set did not show any significant changes from the original data set. This showed robustness in the morphology data, increasing the confidence in the conclusions made.

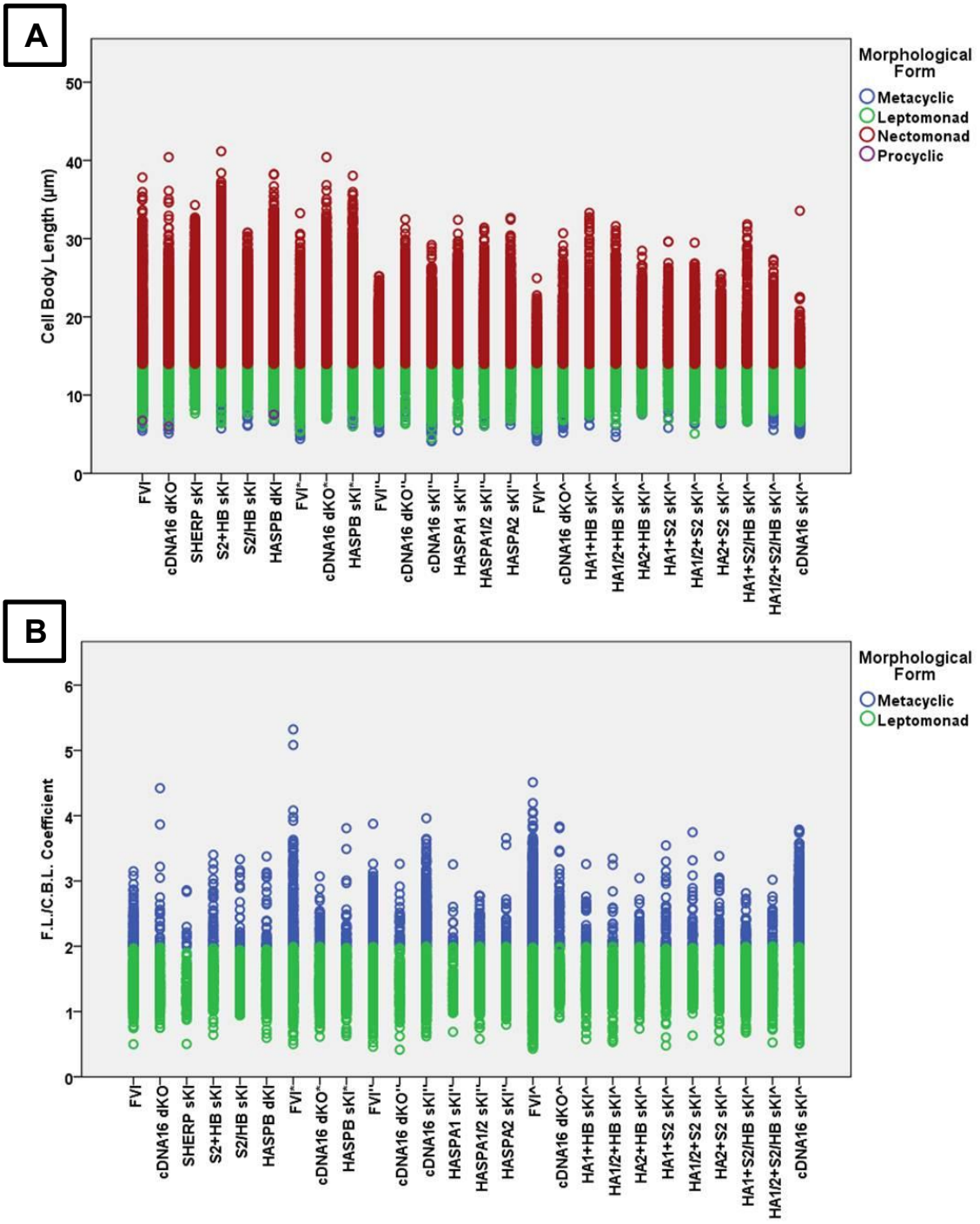
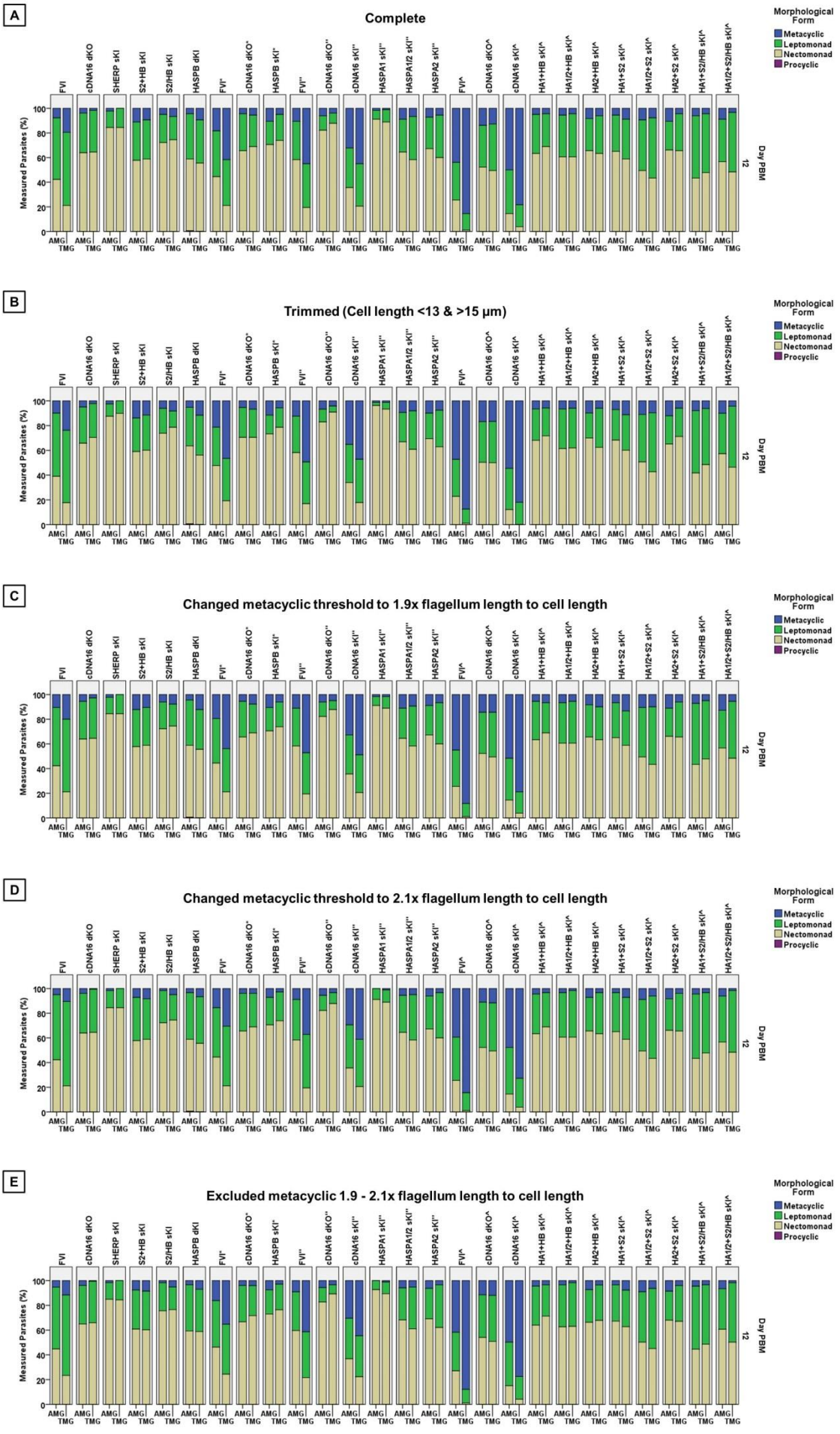


Fig.4.10 – Limitations of measurements to determine parasite form

A) shows gradient of cell body length in particular from leptomonads to nectomonads, where many parasite measurements cluster around the 14 µm threshold and not clear distinction is truly possible between the two. B) shows the gradient for the coefficient of flagellum length (F.L.) divided by the cell body length (C.B.L.), which marks the difference between leptomonads and metacyclics, if the threshold is ≥2. Also here is no clear distinction between the majority of parasites the cluster around this mark. Each circle is a single measurement.


A
Complete
B
Trimmed (Cell length <13 & >15 µm)
C
Changed metacyclic threshold to 1.9x flagellum length to cell length
D
Changed metacyclic threshold to 2.1x flagellum length to cell length
E
Excluded metacyclic 1.9 - 2.1x flagellum length to cell length
Morphological Form
Metacyclic
Leptomonad
Nectomonad
Procyclic
Measured Parasites (%)
Day PBM
12
FVI
cDNA16 dKO
SHERP sKI
S2+HB sKI
S2/HB sKI
HASPB dKI
FVI*
cDNA16 dKO*
HASPB sKI*
FVI**
cDNA16 dKO**
cDNA16 sKI**
HASPA1 sKI**
HASPA1/2 sKI**
HASPA2 sKI**
FVI^
cDNA16 dKO^
cDNA16 sKI^
HA1+HB sKI^
HA1/2+HB sKI^
HA2+HB sKI^
HA1+S2 sKI^
HA1/2+S2 sKI^
HA2+S2 sKI^
HA1+S2/HB sKI^
HA1/2+S2/HB sKI^
AMG
TMG

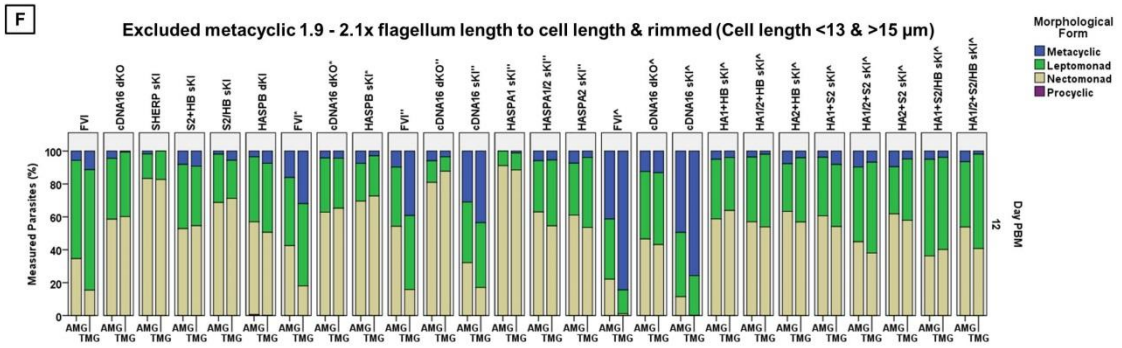


Fig.4.11 – Parasite morphology data with adjusted threshold

A) shows the unsummarized, unaltered morphology dataset. B) shows the data when parasites with a cell body length between 13 – 15 µm were excluded from consideration to achieve a clearer the threshold between leptomonads and nectomonads. C) – F) address the distinction between leptomonads and metacyclics by the ≥2 coefficient. In C) the data is shown for when the threshold was decreased to 1.9 and in D) when it was increased to 2.1. In E) parasites with a coefficient between 1.9 – 2.1 were excluded from consideration and F) combined the statistical aspects of B) and E). In F) this statistical exclusions deleted as little as 4.73% (FVI^) of all parasites up to 27% of parasites (HASPA2 sKI). However, all these modification did not change anything about the statistical relations as shown in the graphs between the different strains analysed, which showed robustness of the dataset with the chosen thresholds.

4.2.7. Discriminating between FVI metacyclics and mutant metacyclics

One point of concern that remained was the classification of metacyclics based solely on the coefficient between flagellum and cell body length. By day 12 PBM, parasites from sand fly midguts classified as metacyclics of FVI and *Lmj*cDNA16 sKI showed a very narrow cell body with a pointed posterior, while those of *Lmj*cDNA16 dKO, *Lmj*HASPB sKI, *Lmj*SHERP sKI, *Lmj*HASPA2 sKI and others showed a visibly broader cell body with a rounded posterior (Fig.4.12B), which gave the cell body a more leptomonad-like appearance. Looking at the average cell body width of classified metacyclics by day 12 PBM, it became apparent that classified metacyclics of FVI were significantly narrower ($P<0.001$) than those of *Lmj*cDNA16 dKO, *Lmj*HASPA2 sKI and *Lmj*S2/HB sKI, while difference of metacyclic cell body width compared to *Lmj*HASPB sKI ($P=0.057$) and *Lmj*HASPA1 sKI ($P=0.054$) were not significant, which could be due to the low number of measured metacyclics for these two strains (28 and 5 metacyclics, respectively, compared to 491 in FVI). Although *Lmj*SHERP sKI did not show a significant difference in the Kruskal-Wallis test, its position in the bar chart in Fig.4.12A showed that its metacyclics were visibly broader. This is underlined by the Giemsa stained parasites shown in Fig.4.12B. Although the majority of classified metacyclics in *Lmj*S2+HB sKI were significantly broader than FVI metacyclics, a few were equally narrow as in FVI. Due to the low efficiency of metacyclic generation in *Lmj*S2+HB sKI, however, it was not possible to establish whether HASPB and SHERP were sufficient to recover the narrow metacyclic cell body phenotype of FVI. The replacement of all HASP and SHERP genes back into the null background was sufficient in *Lmj*HA1/2+S2/HB sKI and *Lmj*cDNA16 sKI to re-establish the narrow cell body phenotype with the pointed posterior typical for FVI metacyclics at day 12 PBM. Interestingly, it was not sufficient to rescue the efficient metacyclic generation in *Lmj*HA1/2+S2/HB sKI, which was observed in *Lmj*cDNA16 sKI. In *Lmj*cDNA16 sKI, all HASP and SHERP genes had been replaced in a single construct which contained the native cDNA16 locus, while in *Lmj*HA1/2+/S2/HB sKI the genes had been replaced as two separate construct into the alleles. This suggested that gene organization within the locus may be important for parasite metacyclogenesis *in vivo*. It needs to be emphasised that, with the exception of *Lmj*cDNA16 sKI, none of the mutants recovered efficient metacyclic generation.

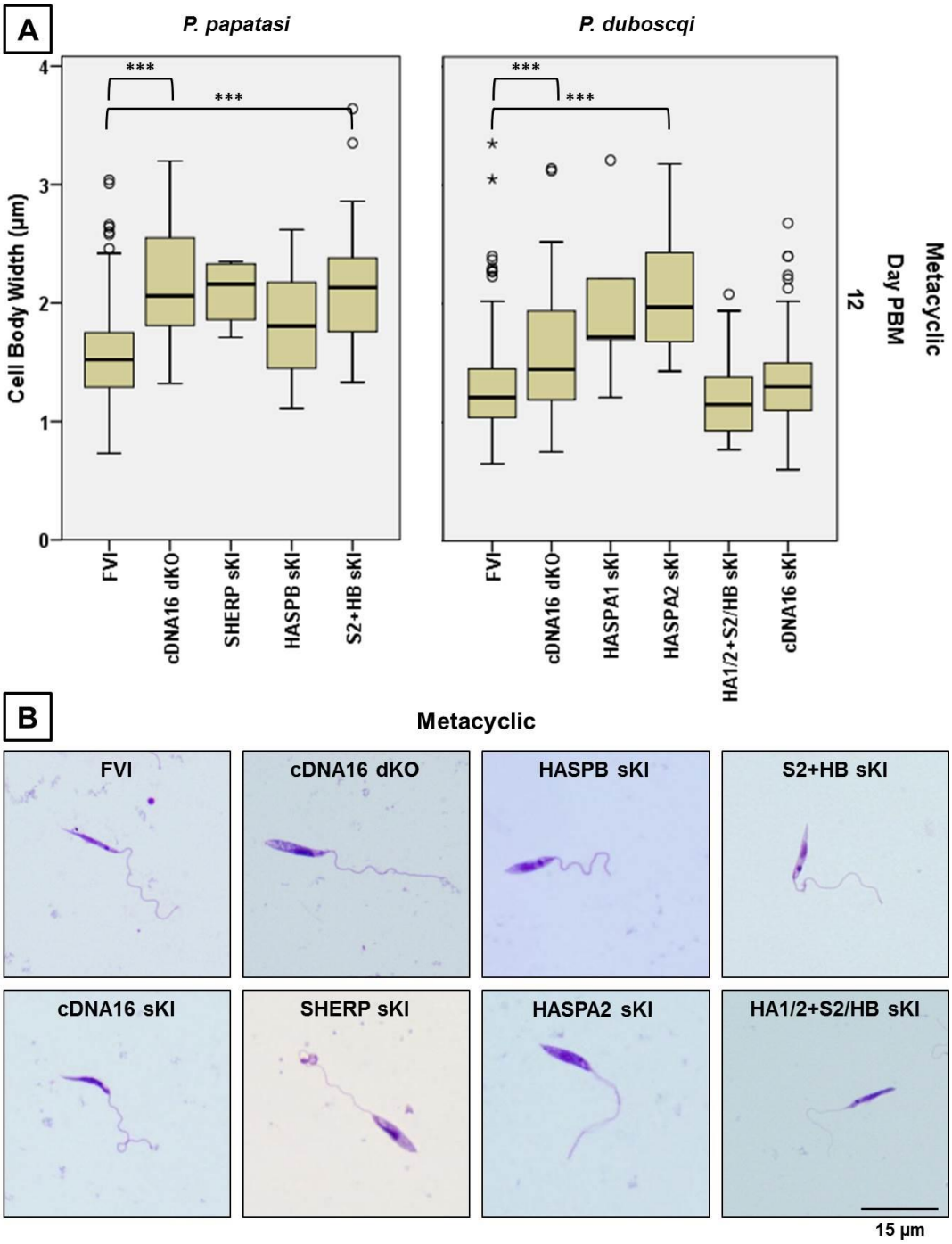


Fig.4.12 – Differences in cell shape among metacyclics from different strains

The stringency of the measurement threshold had a draw back when metacyclics were being identified. Cell bodies of metacyclics in many mutant lines were visibly different among strains. While FVI had metacyclics with narrow cell bodies and pointy posterior, metacyclics of some mutant strains (*Lmj*cDNA16 dKO, *Lmj*S2+HB sKI *&* *Lmj*HASPA2 sKI) were significantly broader (P<0.001), while *Lmj*HASPB sKI (P=0.057) and *Lmj*HASPA1 sKI (P=0.054) border on significance. Other mutant strains like *Lmj*cDNA16 sKI and *Lmj*HA1/2+S2/HB sKI showed the same narrow cell body as FVI. Also *Lmj*S2+HB sKI showed some of these narrow metacyclics, but the majority of classified metacyclics had the broad leptomonad like cell body.

4.3. Conclusions

The data described in this chapter show that simply replacing correctly-regulated genes back into the former cDNA16 locus did not rescue restoration of metacyclogenesis. This was unexpected, since all mutants had been meticulously tested in *in vitro* cultures and had been verified for correct expression and regulation of the HASPs and SHERP genes. Not even *Lmj*HA1/2+S2/HB sKI, which contained one gene of each of the four gene types found in the cDNA16 locus (HASPA1, SHERP, HASPB and HASPA2), rescued the parental line (FVI) phenotype; only *Lmj*cDNA16 sKI did. *Lmj*cDNA16 sKI contained a single copy of the whole cDNA16 locus in its native form in a single replacement construct, which distinguishes it from all other mutant lines. All HASP and SHERP gene replacement constructs contained one or two genes with their native 5' and 3' UTRs, but out of the native context of the cDNA16 locus. This could impact on the regulation and expression of the genes. For example, during mRNA maturation in *Leishmania*, the polycistronic transcripts are 5' *trans*-spliced and 3' polyadenylated in a coupled process (reviewed in (*308*, *309*)). 5' *trans*-splicing sites are generally 100 – 300 nt downstream of the 3' poly(A)site of upstream genes and it has been proposed that the 5' *trans*-splicing sites help to determine the poly(A)-sites of the upstream gene (*313*, *314*). Since the HASP and SHERP genes are out of their native context within the constructs, it could be that changes to their poly(A)-sites occur, which causes mRNA instability. This could be selective for the environmental conditions these mutant lines find themselves in, because within the rich M199 *in vitro* culture medium, gene regulation and expression worked as expected (Fig.3.9). Yet unidentified regulatory elements may exist up- and downstream of the HASP and SHERP genes that were excluded in the gene fragments, when they were amplified. These regulatory elements may be important for gene regulation within the sand fly vector rather than a rich medium like M199, where the parasite may be taking cues from sand fly-derived molecules to time their development. Further investigation is required to address these questions.

While the study of Sádlová *et al.* (2010) suggested that parasite metacyclogenesis was primarily stalled in the nectomonad stage, the data presented here suggest that the critical step stalled in parasite metacyclogenesis is the differentiation of parasites past the leptomonad stage into metacyclics and haptomonads. This conclusion is based on the morphology and localization study, respectively. The morphology data showed clearly that no mutant line, with the exception of *Lmj*cDNA16 sKI, was able to generate metacyclics in high numbers

by day 12 PBM, like FVI and *Lmj*cDNA16 sKI. The localization data suggested that haptomonads were not being produced, because the SV was never colonized in the mutant lines, with the exception of FVI and *Lmj*cDNA16 sKI. Gene replacement of combinations of HASP and SHERP genes does aid leptomonad generation in the sand fly midgut, however. The data show a trend of increased leptomonad generation with the number of HASP and SHERP genes replaced, although it is not clear which HASP and/or SHERP genes were most prominent in promoting leptomonad generation.

Sádlová *et al.* (2010) suggested a key role for HASPB during metacyclogenesis restoration. They tested an episomal HASPB replacement mutant in the null background. The data in this study, however, did not confirm this observation. In all replacement mutants, with the exception of *Lmj*cDNA16 sKI, HASPB replacement by homologous recombination did not restore metacyclogenesis. In fact, replacement mutants had similar phenotypes to the cDNA16 locus null mutant, *Lmj*cDNA16 dKO, confirming that replacing all single component genes back into the cDNA16 locus was insufficient to restore metacyclogenesis. This could mean that increased HASPB expression is the key for metacyclogenesis completion.

In summary, FVI and *Lmj*cDNA16 sKI parasites were able to establish stronger infections in both *Ph. (Ph.) papatasi* and *Ph. (Ph.) duboscqi* than all other mutant lines. FVI and *Lmj*cDNA16 sKI were also the only strains to colonize the SV efficiently, suggesting efficient haptomonad generation, and to produce high numbers of narrow bodied metacyclics. Further analysis of mutant parasites in the sand fly vector will be necessary to explain the failure of single component HASP and SHERP gene replacement to rescue the parental line phenotype.

5. Chapter V. – Further investigation into the cDNA16 locus *in vitro* and *in vivo*

5.1. Introduction

The results from the sand fly infection studies (Chapter IV) required further investigation into the mutants' behaviour within the sand fly midgut, to explain why none of the 17 replacement mutants had rescued full completion of metacyclogenesis. Most importantly, the expression and regulation patterns of the HASPs and SHERP from the replacement constructs had to be assessed within the sand fly midgut. The question as to whether *in vitro* growth in the rich M199 did alter parasite behaviour and/or HASP and SHERP regulation markedly compared to *in vivo* growth in the sand fly midgut also needed to be addressed, since it has been found that parasite differentiation has different kinetics in minimal conditions, such as 5% sucrose/PBS, compared to rich medium like M199 (personal communication, S. Kamhawi and D. Sacks, WorldLeish5, May 2013). The observed lack of strong TMG colonization in all mutant lines with the exception of *Lmj*cDNA16 sKI needed further investigation, too. Another important question was, whether the mutant lines could actually be transmitted from the sand fly to a new mammalian host, since neither the SV was colonized efficiently, which suggested no SV degradation, nor the PSG seemed to be secreted efficiently, which had been suggested by the lack of gel-immobilized parasites in dissected TMGs at days 9 and 12 PBM.

5.2. Confocal microscopic analysis of HASP and SHERP localization in *Leishmania* parasites derived from culture and sand fly midguts

Confocal microscopy had been used previously as a method for HASPB and SHERP detection and localization in the FVI, *Lmj*cDNA16 dKO and *Lmj*cDNA16 sKI lines (*146*, *336*). The same approach was used to investigate HASPB and SHERP expression and localization in the new mutant lines derived from M199 cultures and in midgut derived mutant lines. Parasites were fixed on glass slides and probed with specific antibodies (anti-336 for HASPB and anti-SHERP for SHERP) for the individual proteins. Alexa Fluor® secondary antibodies – usually 488 anti-rabbit – were used for the fluorescent signal required for visualization under a confocal microscope. DAPI was used to stain the DNA of the nucleus and kinetoplast of permeabilized parasite cells and was visualised with a 405 nm laser. *Lmj*cDNA16 dKO parasites served as a negative control, while FVI not probed with the primary antibody was used as an antibody control to check for non-specific binding of secondary 488 anti-rabbit antibodies.

HASPB was probed for in cultured and midgut derived FVI, *Lmj*cDNA16 dKO, *Lmj*HASPB sKI, *Lmj*HASPB dKI, *Lmj*S2+HB sKI and *Lmj*S2/HB sKI. The results for cultured parasites showed that HASPB was present on the cell body and flagellum of all tested mutant lines except for *Lmj*cDNA16 dKO as expected (Fig.5.1). All parasites positive for HASPB detection were measured and verified to be metacyclic (Table 5.1). These data demonstrated that HASPB expression and regulation in these mutants occurred as in the parental line (FVI). The strongly positive signal obtained from the *Lmj*HASPB sKI and *Lmj*HASPB dKI strains suggested that the other cDNA16 locus genes were not required for HASPB expression in the new mutants in culture. This was an interesting observation, because it had been debated whether SHERP, which locates to the cytosolic face of the ER and mitochondrial outer membrane, might interact with an ATPase-pump on vesicles and play a role in HASPB stable expression and/or trafficking (*348*). Also, SHERP was detected in all SHERP-containing mutant lines tested from culture (Fig.5.2), although the signal was not as compartmentalized in the cell body as previously observed (*341*). Thus, the confocal data confirmed the expression and regulation of HASPB and SHERP within mutant lines shown already on Western blots (see 3.3.1.4).

Since Western blots are inapplicable for sand fly derived parasites due to the great limitation of parasite material that can be generated from a sand fly midgut (about $5x10^3$ – $2x10^4$ parasites per midgut), HASP and SHERP expression and regulation were tested by confocal microscopy on midgut derived parasites in methanol-fixed midgut smears. The antibody probing, imaging and analysis was done using the same protocol as for the cultured parasites. A significant problem for HASPB and SHERP detection in midgut smears was the high level of background staining. Although a positive signal was detected for HASPB in the positive control (FVI), surprisingly, HASPB was not detected in *Lmj*HASPB sKI, *Lmj*HASPB dKI, *Lmj*S2+HB sKI and *Lmj*S2/HB sKI, while a clear positive signal had been observed in the same mutant lines in cells derived from M199 cultures (Fig.5.3). The same applied for SHERP, which was detected in midgut-derived FVI, but not in *Lmj*SHERP sKI, although a positive signal had been observed in the same strain derived from culture (Fig.5.4). Potentially, the rich culture condition in M199 could have a positive feedback on HASPB and SHERP expression. Alternatively, a sand fly derived signal or midgut metabolite could be involved, via a yet-to-be-identified signalling cascade, in the regulation of HASP and SHERP expression. Potentially, mRNA stability, translational or post-translational mechanisms might be active in the sand fly midgut only, thereby

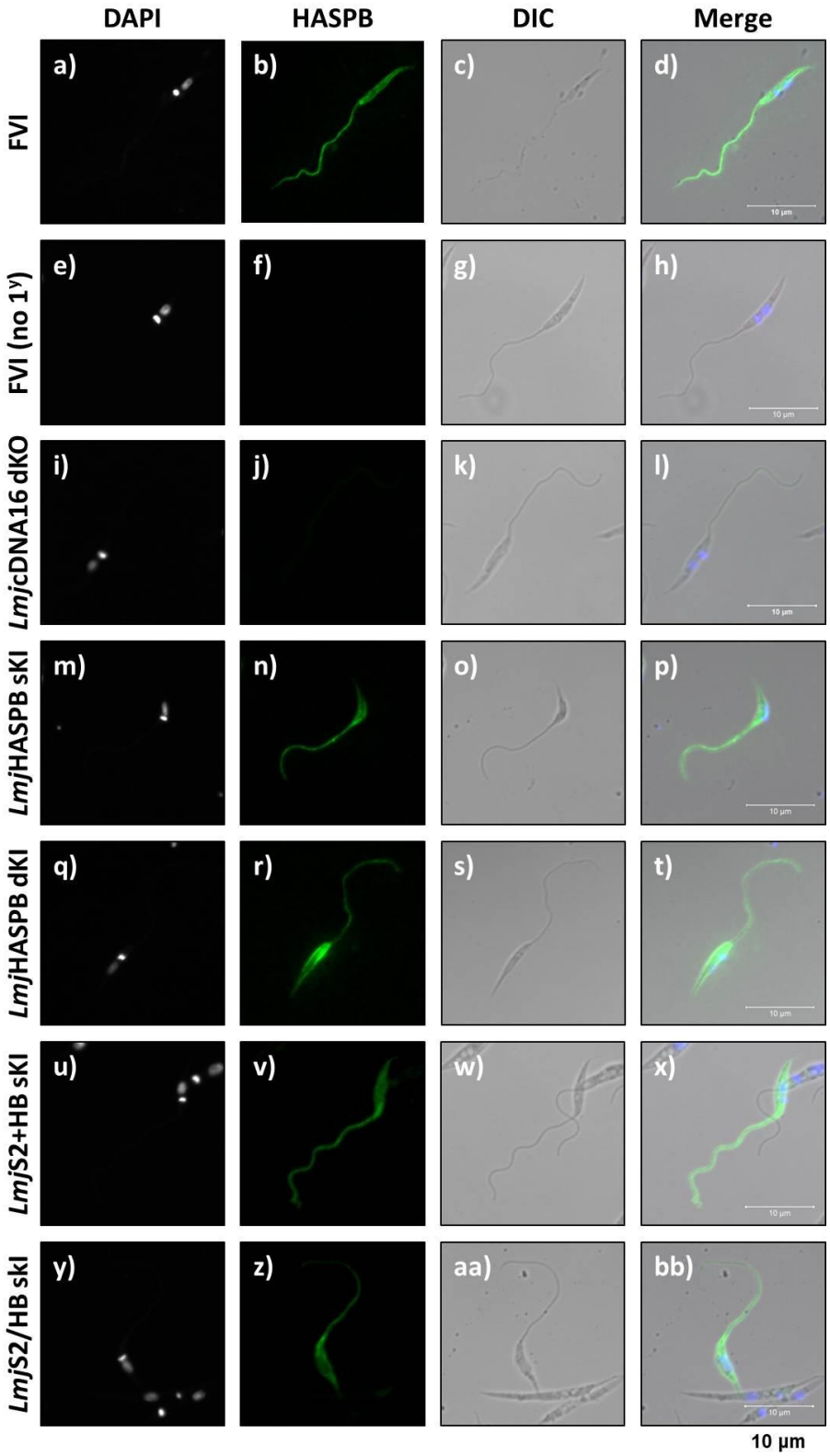


Fig.5.1 – Confocal images detecting HASPB in culture derived parasites

All parasites were probed with the primary ab336 and the secondary Alexa Fluor® ab488 with the exception of FVI (no 1^{y}) (e, f, g, h), which served as a non-specific-binding control for the secondary. Only *Lmj*cDNA16 dKO did not show a positive signal for HASPB probing (j), which was expected. HASPB was detected on the cell body and the flagellum in all positive samples as expected.

Table 5.1 – Measurements of Parasites on Confocal Images

Strains	Flagellum	Cell length	Cell width	Form
		Fig.5.1		
FVI	20.66 µm	10.13 µm	1.11 µm	Metacyclic
FVI (No 1^{o})	19.21 µm	9.57 µm	1.33 µm	Metacyclic
KO	18.69 µm	9.65 µm	1.55 µm	Leptomonad
HASPB sKI	16.06 µm	6.89 µm	1.05 µm	Metacyclic
HASPB dKI	19.74 µm	9.36 µm	1.38 µm	Metacyclic
S2+HB sKI	22.02 µm	8.89 µm	1.61 µm	Metacyclic
S2/HB sKI	21.37 µm	8.30 µm	1.51 µm	Metacyclic
		Fig.5.2		
FVI	18.92 µm	8.22 µm	1.08µm	Metacyclic
FVI (No 1^{o})	21.09 µm	6.34 µm	1.96 µm	Metacyclic
KO	17.49 µm	6.47 µm	1.08 µm	Metacyclic
SHERP sKI	19.71 µm	8.38 µm	1.81 µm	Metacyclic
S2+HB sKI	19.09 µm	7.95 µm	1.93 µm	Metacyclic
S2/HB sKI	16.50µm	8.98µm	1.42µm	Metacyclic
		Fig.5.3		
FVI	24.07 µm	8.47 µm	1.32 µm	Metacyclic
FVI (No 1^{o})	16. 56 µm	7.42 µm	1.76 µm	Metacyclic
KO	18.88 µm	9.18 µm	2.18 µm	Metacyclic
HASPB sKI	12.80 µm	8.54 µm	1.43 µm	Leptomonad
HASPB dKI	19.50 µm	9.76 µm	2.37 µm	Metacyclic
S2+HB sKI	23.07 µm	11.89 µm	1.78 µm	Leptomonad
S2/HB sKI	23.69 µm	9.78 µm	1.87 µm	Metacyclic
		Fig.5.4		
FVI	22.31 µm	7.84 µm	1.81 µm	Metacyclic
FVI (No 1^{o})	16.16 µm	7.95 µm	1.34 µm	Metacyclic
KO	19.19 µm	9.58 µm	1.48 µm	Metacyclic
SHERP sKI	22.31 µm	11.05 µm	2.23 µm	Metacyclic
S2+HB sKI	19.49 µm	11.63 µm	1.58 µm	Leptomonad
S2/HB sKI	21.96 µm	10.64 µm	1.68 µm	Metacyclic

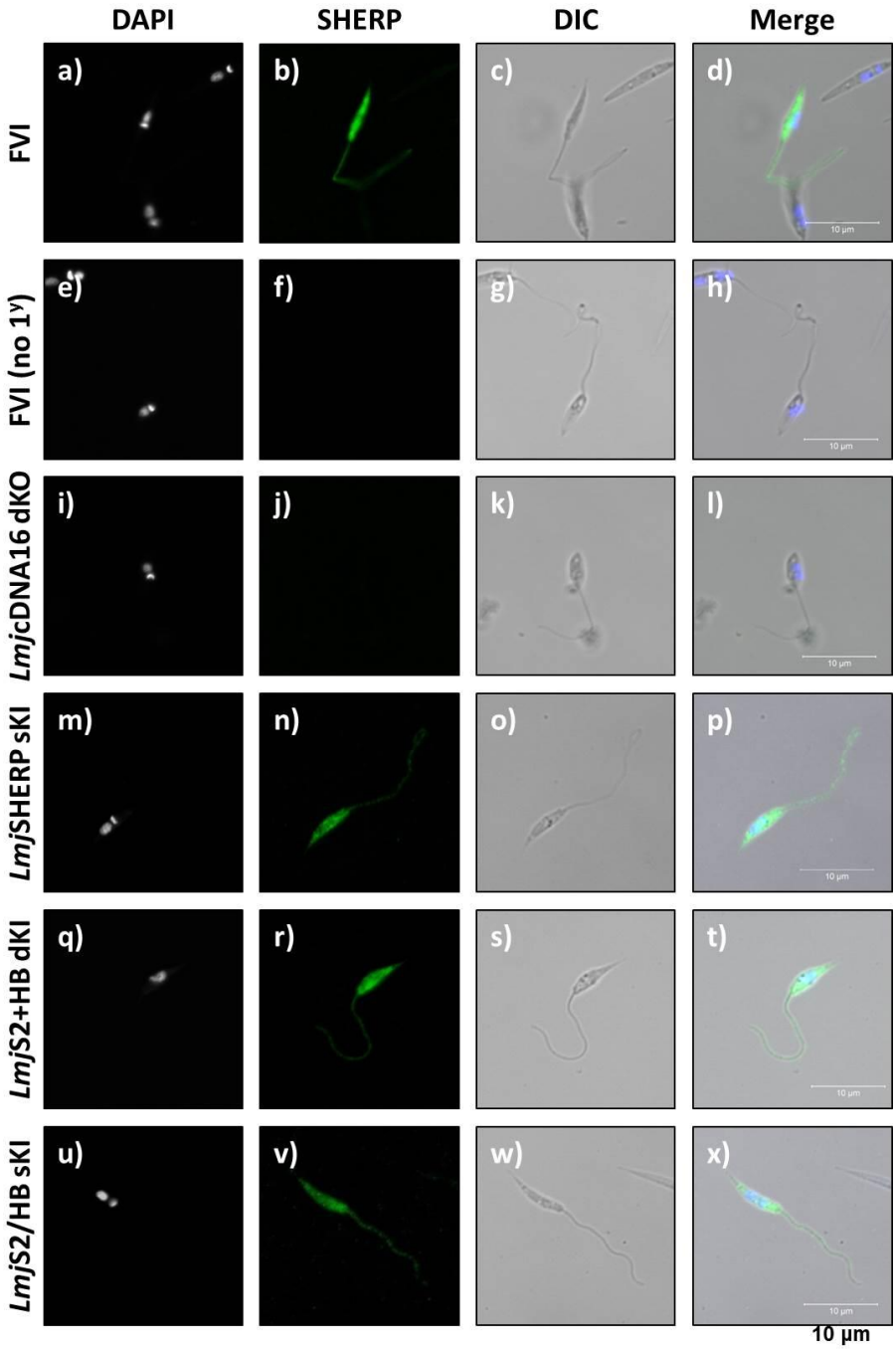


Fig.5.2 – Confocal images detecting SHERP in culture derived parasites

All parasites were probed with the primary abSHERP and the secondary Alexa Fluor® ab488 with the exception of FVI (no 1^{y}) (e, f, g, h), which served as a non-specific-binding control for the secondary. Only *Lmj*cDNA16 dKO did not give a positive signal as expected, while all other tested lines had a strong positive signal in the cell body.

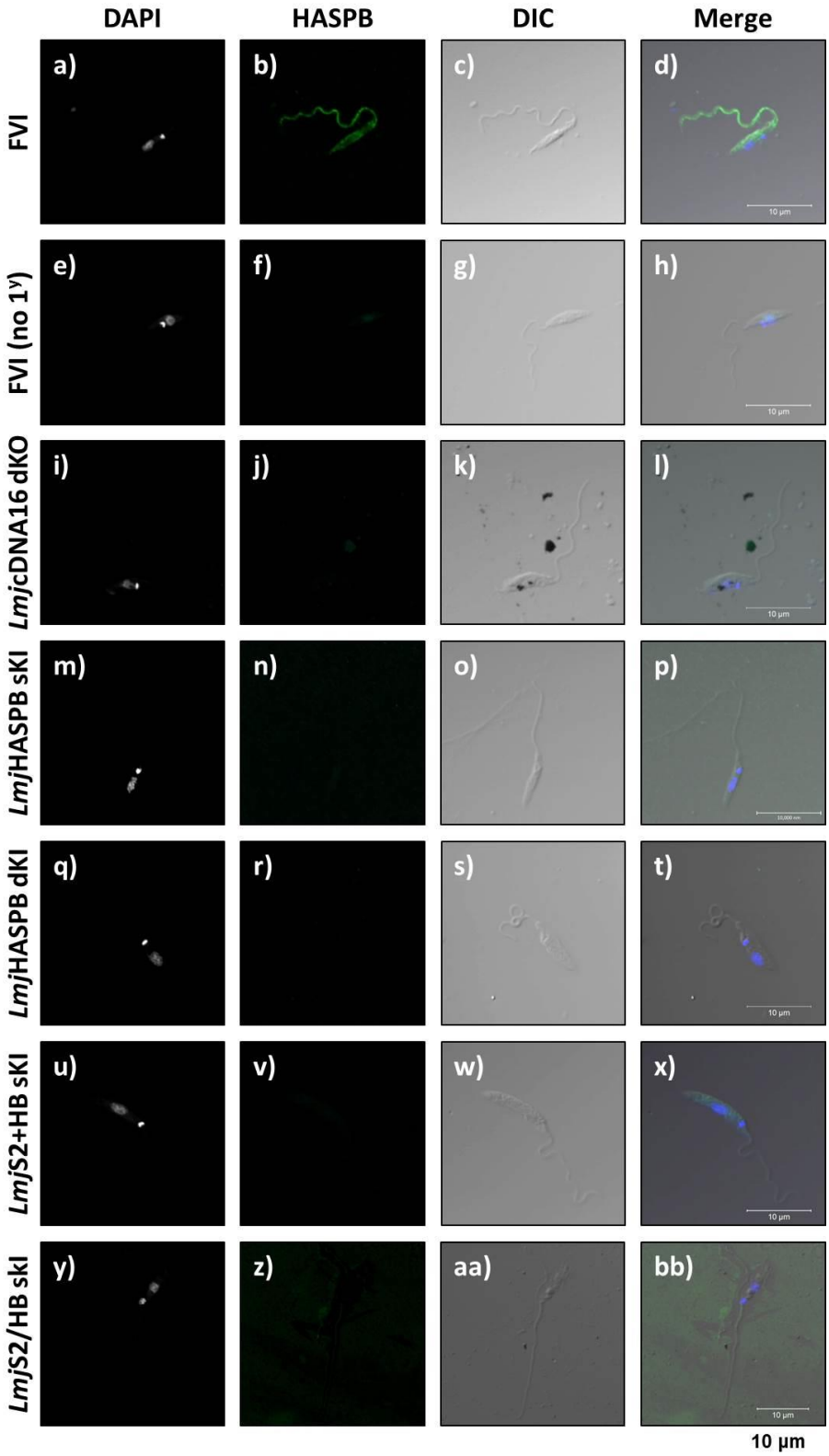


Fig.5.3 – Confocal images detecting HASPB in midgut derived parasites

For the confocal analysis, fixed parasites on midgut smear slides were treated as previously for the culture-derived parasites (Fig.5.1 & 2). Only FVI showed a positive signal for HASPB, while all tested mutant lines were negative together with the negative control *Lmj*cDNA16 dKO (in contrast to the data generated from cultured parasites).

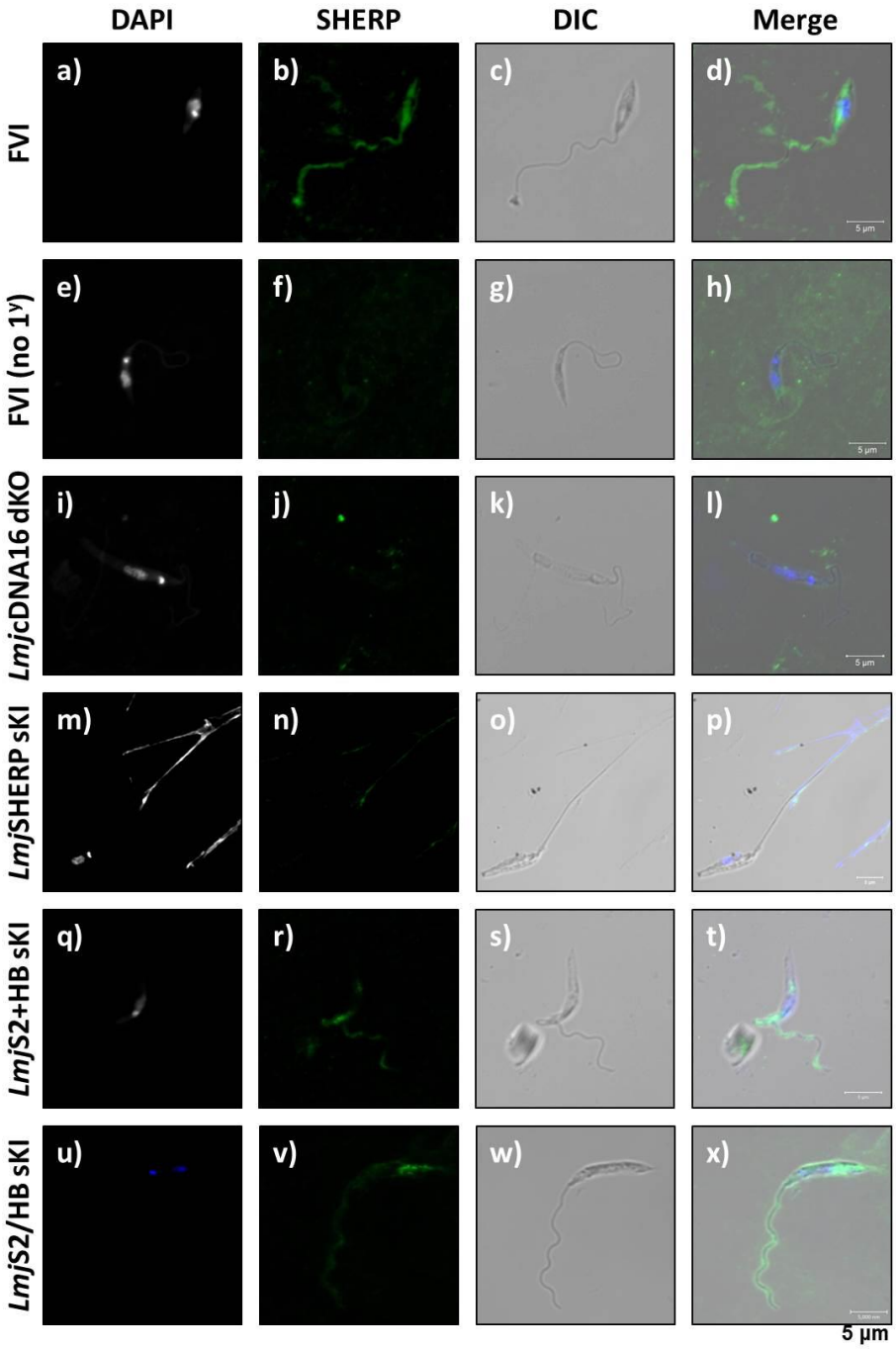


Fig.5.4 – Confocal images detecting SHERP in midgut derived parasites

Parasites derived from midguts were also probed for SHERP on midgut smears. While FVI had a positive signal for SHERP, the single SHERP replacement mutant did not show a positive signal. Interestingly, SHERP was detected in the SHERP and HASPB containing mutant line, *Lmj*S2+HB sKI.

preventing the stable expression of HASPB and SHERP.

Interestingly, SHERP expression seemed to be rescued in the *Lmj*S2+HB sKI and LmjS2/HB sKI mutant lines *in vivo*, which showed a compartmentalized signal in the cell body. This suggested that the presence of the HASPB gene promoted stable SHERP up-regulation, although HASPB itself was not up-regulated in these mutant lines *in vivo*. This was confusing, because SHERP has been shown to be up-regulated ahead of HASPB in culture (*341*). Due to a lack of a specific HASPA antibody, it was not possible to investigate HASPA1 and HASPA2 expression in these experiments. The non-affinity purified HASP antibody, used successfully for HASPA detection on the Western blots, was used to probe for HASPA in mutant lines not containing HASPB, but no detectable signals were observed in all tested lines from culture. The non-affinity purified HASP antibody produced a signal only in FVI, which, however, was likely to be due to HASPB which cannot be distinguished from HASPA with this antibody on a confocal slide. Attempts to raise a specific HASPA antibody in rabbits via two different protocols failed for unknown reasons and, therefore, HASPA expression could not be investigated by confocal microscopy.

Further investigations are required to explain the distinct differences in HASPB and SHERP detection in mutant lines in culture and in the sand fly midgut.

5.3. Parasite *in vitro* differentiation in 5% sucrose/PBS compared to M199 medium

One possibility for the differences in HASPB and SHERP expression *in vivo* compared to *in vitro* was that the relatively stable and nutrient rich conditions in M199 interfered with the proper regulation of the cDNA16 locus. In contrast, the conditions the parasites normally encounter in the sand fly midgut are highly dynamic, with nutrients depleted once the blood meal has been excreted. In nature, the sand fly's primary food sources are nectar and plant saps, which are essentially high concentration sugar solutions that are channelled into the midgut in regular intervals from the sand fly's crop. Observations indicated that there are distinct differences in the outcome of *Leishmania* metacyclogenesis between parasites grown in M199 and 5% sucrose/PBS, which mimicked the minimal conditions in the midgut after blood excretion (personal communication, S. Kamhawi and D. Sacks, 2013). In particular, metacyclics were distinct between parasites grown in M199 and 5% sucrose/PBS and in the latter case, two different types of metacyclics have been found, something never observed in M199. To

verify that M199 conditions did not influence HASP and SHERP gene expression from the constructs, HASPB and SHERP expression were tested in both M199 and 5% sucrose/PBS. Since starvation is a trigger for metacyclogenesis, parasites were grown until late log phase (day 2-3 p.i.) in M199 before transferring them into 5% sucrose/PBS after several washes. A time course of protein samples (day 3-6 p.i.) was taken from FVI, *Lmj*cDNA16 dKO, *Lmj*cDNA16 sKI, *Lmj*HASPB sKI and *Lmj*SHERP sKI for both culture conditions, and the lysates from these parasites were then analysed by Western blots. It was expected that if the minimal conditions of 5% sucrose/PBS interfered with the expression of HASPB and SHERP from the constructs in *Lmj*HASPB sKI and *Lmj*SHERP sKI, respectively, then no increase of HASPB and SHERP would be seen in 5% sucrose/PBS compared to M199. The results showed that no distinct difference in expression of HASPB and SHERP was detected in *Lmj*HASPB sKI and *Lmj*SHERP sKI, respectively, between the two conditions (Fig.5.5). The differences observed in FVI and LmjcDNA16 sKI HASP and SHERP expression were unexpected and may be artefacts. These results were interpreted as indicative of another factor (other than nutrient depletion) interfering with HASP and SHERP expression *in vivo*. Potentially, sand fly derived molecules or midgut metabolites were responsible for the differences in expression, acting through a yet unknown signalling pathway.

5.4. Assessing the potential effect of sand fly midgut molecules on parasite growth in liquid medium

To investigate the possibility that HASP and SHERP gene expression might be regulated by sand fly derived molecules or midgut metabolites, sand flies were fed on uninfected heat inactivated rabbit blood to stimulate changes in midgut enzyme content and simulate the conditions normally experienced by parasites during their development in the midgut. 50 midguts were dissected at days 6 and 12 PBM, respectively, into 100 µl M199. The midguts were homogenized, briefly spun down and the supernatant filtered through a 0.2 µm pore membrane to eliminate bacterial and fungal contaminants. This meant that only soluble midgut components were in the supernatant. The M199 medium contained antibiotics against bacteria and fungi, too, as a precaution. The filtered midgut extracts were diluted in 4 ml M199 (1 midgut / 80 µl) and four 1 ml culture were set up with FVI, *Lmj*cDNA16 dKO, *Lmj*cDNA16 sKI and *Lmj*HASPB sKI. Another four 1 ml cultures were set up without the midgut extracts of the same lines as negative controls. Since M199 is a very protein rich medium due to FCS supplementation, the change in protein content of the medium by the addition of the midgut extracts

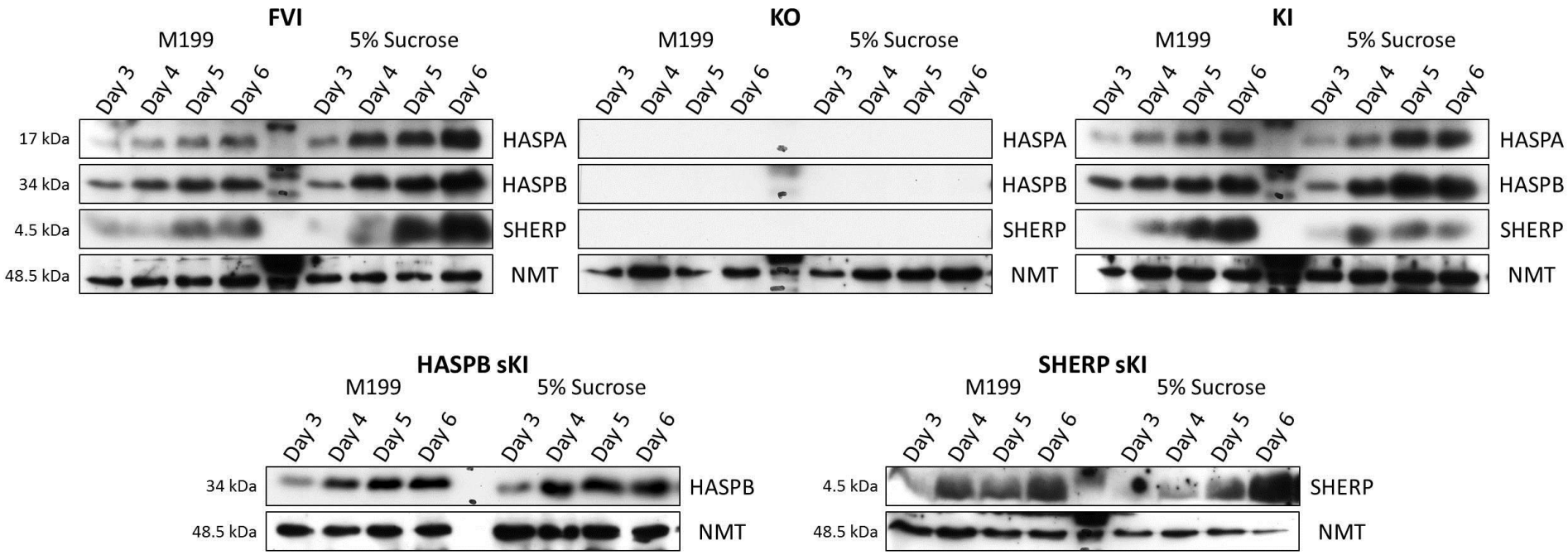


Fig.5.5 – HASP and SHERP expression in parasites grown in M199 versus 5% Sucrose/PBS

Whole lysate samples of parasites grown in rich (M199) and limited (5% sucrose / PBS) media were compared throughout metacyclogenesis in culture, over a time course from day 3-6 p.i. for HASP and SHERP expression. NMT served as a loading control.

was negligible. Cultures were incubated at 23 °C for 7 days before harvesting cells for lysate generation.

The hypothesis tested was that *Lmj*HASPB sKI would show suppressed HASPB expression in the presence of midgut extract if any soluble midgut molecules were present and at sufficient concentrations to affect HASPB expression from the constructs. Analysis of the parasite lysates is shown in Fig.5.6. The results showed that *Lmj*HASPB sKI did not show a marked difference in HASPB expression between growth in M199 with day 6 or 12 PBM midgut extract and the negative controls. This suggested that it was not necessarily the presence of midgut molecules that regulated HASPB expression in the sand fly midgut differently from culture. However, it could also be that the concentration of midgut extracts was too low to make an impact in these experiments. Potentially, relevant molecules may have decomposed over time or the key factors were in the insoluble fraction; perhaps a membrane protein of the midgut epithelia microvilli. Refinement of the method and testing of unfiltered homogenates containing the insoluble fraction would be required to confirm that HASPB construct regulation is not subject to regulation control via a midgut molecule-dependent mechanism. Interestingly, *Lmj*cDNA16 sKI showed comparably low expression of HASPB, which may be due to slower growth in the antibiotics, resulting in the entry into metacyclogenesis at a later day then the other strains tested. In addition, it appeared that *Lmj*cDNA16 sKI produced less HASPB grown in day 6 PBM midgut extracts than without, but this was not observed for growth in day 12 PBM midgut extracts. A repeat of the assay would be required to confirm the observed differences, which was not possible due to time constraints and limitations of sand fly material availability in this project. The difference in HASPB expression observed in FVI grown in day 12 PBM midgut extract could also be due to lower protein loading, looking at the NMT band compared to the negative control. However, a difference was observed in HASPA expression in FVI grown in day 6 PBM midgut extracts, which could not be explained by differences in protein loading. Repeated experiments and perhaps the use of earlier lysates would be required to confirm these findings. However, it is clear that the addition of the midgut extract had no impact on *Lmj*HASPB sKI.

5.5. Looking at HASP and SHERP mRNA levels within the sand fly vector

Since *in vitro* approaches did not offer an explanation for the difference in HASPB and SHERP signal in the confocal microscopy, gene expression was investigated at the mRNA level while the parasites were within the sand fly midgut. Parasite

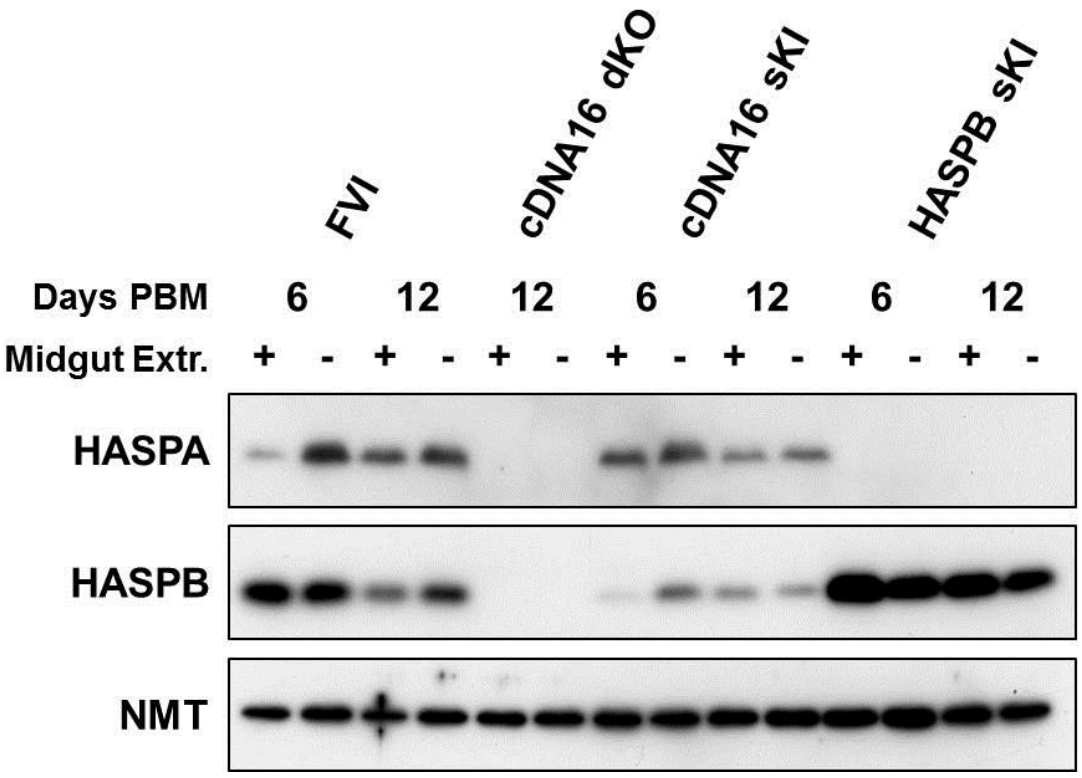


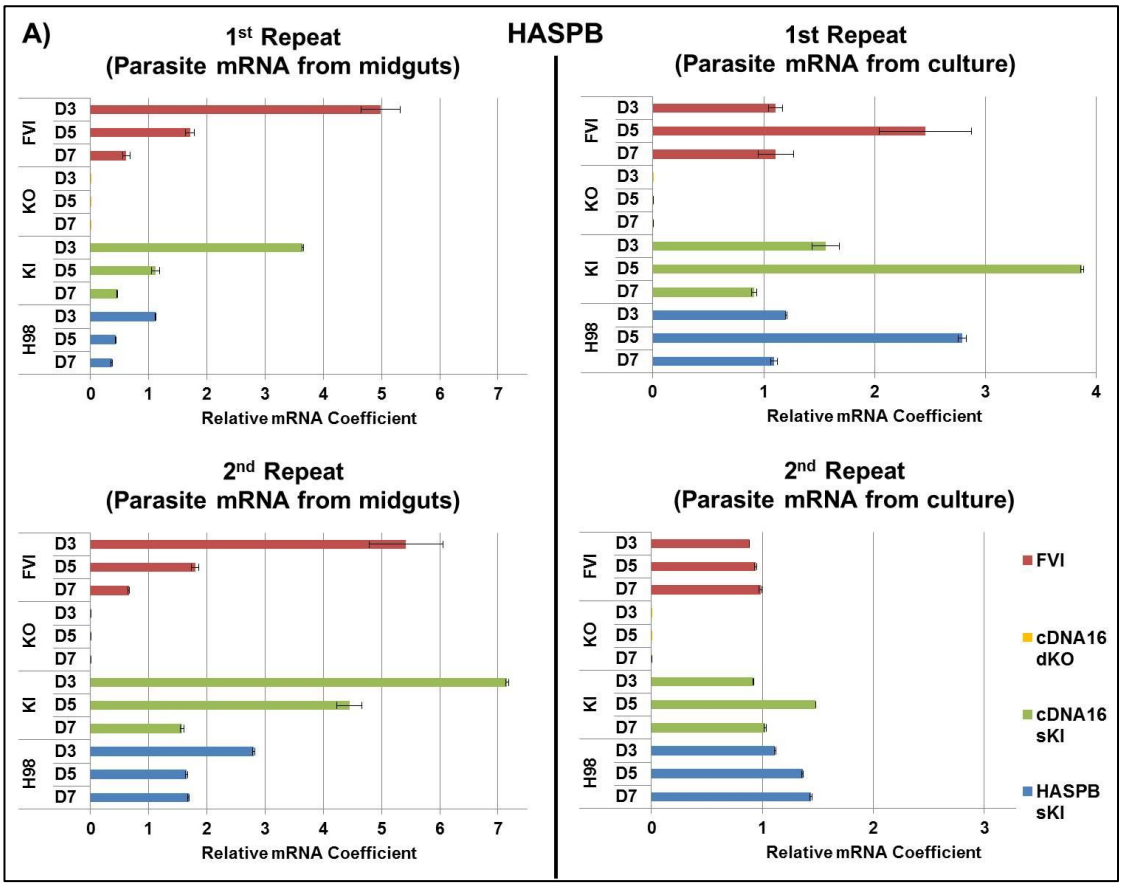
Fig.5.6 – Western blot to assess potential changes in HASPB expression profiles in the presence of midgut extracts

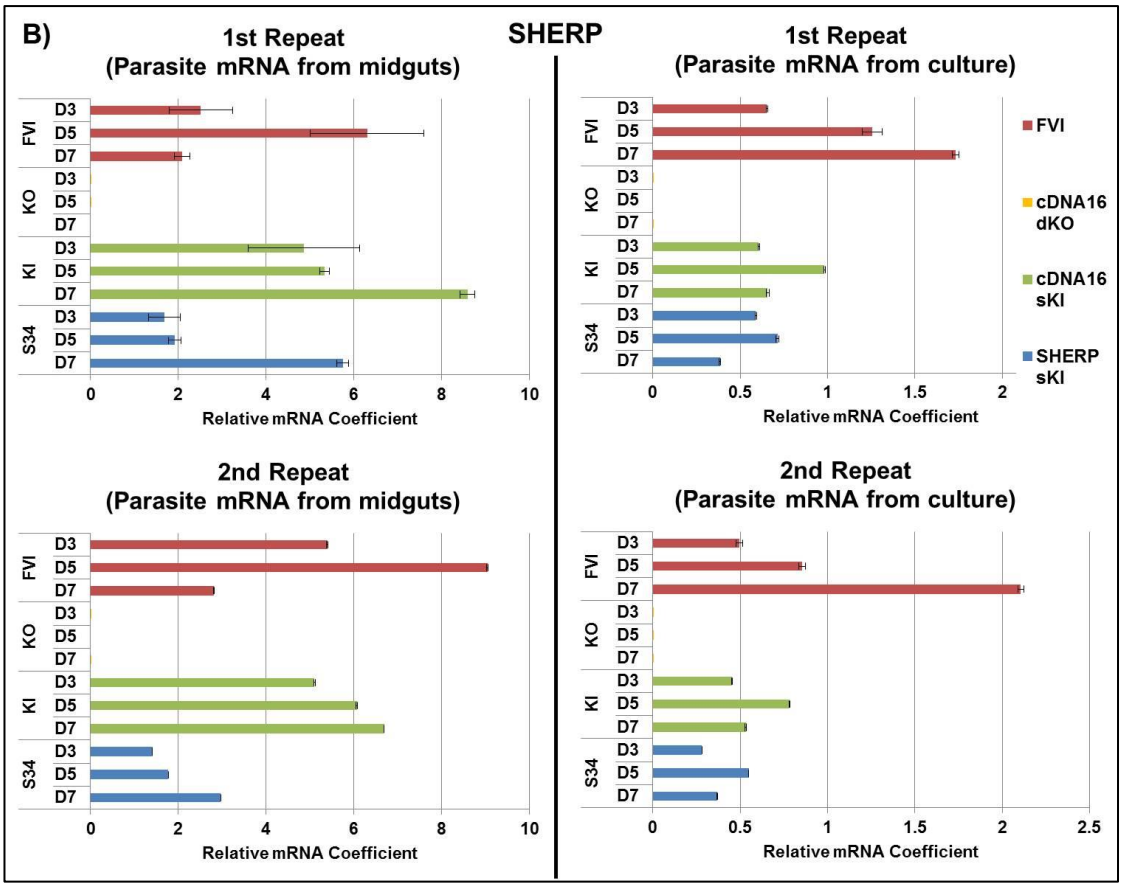
Four strains were grown in M199 for 7 days with (+) and without (-) the addition of extracts of blood fed sand fly midguts from days 6 and 12 PBM. The Western blot does not show a significant reduction in HASPB expression in the *Lmj*HASPB sKI line as anticipated, if midgut molecules would have impacted on the HASPB expression from the construct as observed by confocal microscopy. NMT served as a loading control to exclude that differences in HASPA and HASPB expression were due to uneven loading.

protein analysis by Western blot was difficult due to the limitations of recoverable parasite protein material due to the relatively low parasite numbers present in sand fly midguts (on average $5x10^3$ – $2x10^4$ parasites / midgut). HASP and SHERP mRNA from midgut derived parasites was extracted with magnetic oligo-dT beads from FVI, *Lmj*cDNA16 dKO, *Lmj*cDNA16 sKI, *Lmj*HASPB sKI, *Lmj*HASPA2 sKI and *Lmj*SHERP sKI at days 6, 9 and 12 PBM. 20 midguts were collected per strain per day PBM, homogenized and lysed by flash freezing in liquid nitrogen and lysis buffer for mRNA extraction. In addition, RNA was extracted from the same six lines grown in M199 culture on day 3, 5 and 7 p.i. by Trizol extraction to compare the profiles of HASP and SHERP mRNA level changes over time. The extracted mRNA samples were treated with reverse transcriptase for cDNA generation. These samples were analysed by qPCR for mRNA levels and PCR for mRNA length.

qPCR analysis of the mRNA samples was used to assess HASP and SHERP mRNA levels relative to NMT mRNA levels. NMT mRNA levels are stable throughout the promastigote stages and served as a sample control to normalize the HASP and SHERP mRNA level data. Since the HASPs and SHERP are regulated genes, changes in the mRNA levels over time relative to NMT mRNA levels were expected. Differences in these profiles between parasite lines and growth conditions might explain the difference in HASP and SHERP expression between *in vivo* and *in vitro* and the difference in metacyclogenesis rescue in *Lmj*cDNA16 sKI compared to all other replacement mutants.

The qPCR results are shown in Fig.5.7. The two repeats of midgut derived parasite samples showed that relative HASPB mRNA levels are increased 2.5 – 5-fold at day 6 PBM in FVI and *Lmj*cDNA16 sKI compared to the following time points (Fig.5.7A). They decrease as the midgut infection continues until day 12 PBM in both repeats of the assay. *Lmj*HASPB sKI also showed its highest levels at day 6 PBM, but in *Lmj*HASPB sKI relative HASPB mRNA levels were 2 – 4.5-fold lower than in FVI and *Lmj*cDNA16 sKI on day 6 PBM after normalizing the data. This reduction in HASPB mRNA levels in *Lmj*HASPB sKI could explain the lack of detectable HASPB levels *in vivo*. This hypothesis is supported by the observation that *Lmj*HASPB sKI expressed HASPB mRNA at similar levels and with a similar pattern compared to FVI and *Lmj*cDNA16 sKI *in vitro* in the first qPCR repeat of the culture derived parasite mRNA samples. Unfortunately, the second qPCR repeat did not confirm the results from the first run. A significant reduction of mRNA material in particular in the samples from day 5 p.i. may have


A)
HASPB
1st Repeat
(Parasite mRNA from midguts)
1st Repeat
(Parasite mRNA from culture)
2nd Repeat
(Parasite mRNA from midguts)
2nd Repeat
(Parasite mRNA from culture)
FVI
KO
KI
H98
D3
D5
D7
Relative mRNA Coefficient
FVI
cDNA16 dKO
cDNA16 sKI
HASPB sKI


B)
SHERP
1st Repeat
(Parasite mRNA from midguts)
1st Repeat
(Parasite mRNA from culture)
2nd Repeat
(Parasite mRNA from midguts)
2nd Repeat
(Parasite mRNA from culture)
FVI
KO
KI
S34
D3
D5
D7
Relative mRNA Coefficient
FVI
cDNA16 dKO
cDNA16 sKI
SHERP sKI

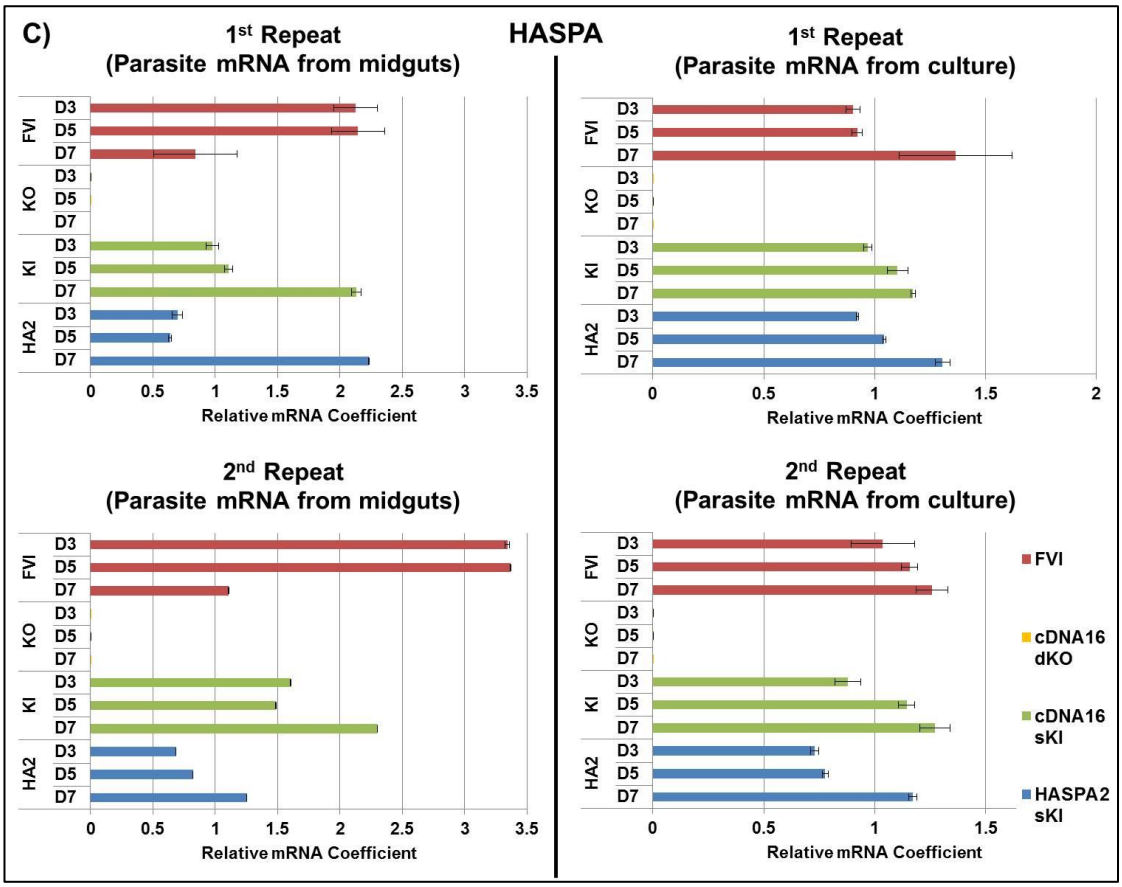


Fig.5.7 – HASP and SHERP mRNA levels relative to NMT mRNA levels

Relative HASP and SHERP mRNA levels of FVI, *Lmj*cDNA16 dKO, *Lmj*cDNA16 sKI, *Lmj*HASPB sKI, *Lmj*SHERP sKI and *Lmj*HASPA2 sKI were assessed by qPCR. Average HASP and SHERP mRNA quantities were normalized against NMT mRNA quantities in the respective samples. Two qPCR repeats of midgut and culture derived parasite HASP and SHERP mRNA are shown. Each sample was run in triplicate repeats per qPCR. Although variations in relative mRNA coefficient were observed between repeats, in general, the patterns of HASP and SHERP up- and down-regulation stayed the same for the different parasite lines tested. Error bars are based on the standard error calculated from the standard deviation from the triplicate repeats per sample. A) shows graphs for relative HASPB mRNA levels, B) for relative SHERP mRNA levels and C) for relative HASPA mRNA levels for midgut and culture derived parasite samples.

been the cause of the changed results in the second run.

Variations in the detected relative mRNA content were also observed in mRNA samples probed for relative SHERP mRNA abundance (Fig.5.7B). For SHERP these variations were comparatively small, however, and did not change the overall SHERP mRNA expression pattern both in midgut and culture derived parasite mRNA samples. Interestingly, the SHERP mRNA expression pattern of *Lmj*cDNA16 sKI matched that of *Lmj*SHERP sKI rather than FVI both *in vivo* and *in vitro*. FVI showed a 2 – 3-fold increase in relative SHERP mRNA levels from day 6 PBM to day 9 PBM followed by a 3 – 3.5-fold decrease at day 12 PBM. *Lmj*cDNA16 sKI and *Lmj*SHERP sKI lag behind in their SHERP mRNA upregulation and an increase is only observed by day 12 PBM. *In vitro* SHERP mRNA levels increase 3 – 4-fold from day 3 p.i. until day 7 p.i. in FVI. In contrast, *Lmj*cDNA16 sKI and *Lmj*SHERP sKI only showed a slight increase in SHERP mRNA levels from day 3 p.i. to day 5 p.i., but then a decrease again towards day 7 p.i.

A stark difference in HASPA expression was observed in the midgut derived parasite mRNA samples between tested strains (Fig.5.7C). FVI showed similar high HASPA mRNA level on day 3 and 5 p.i. and then a drop at day 7 p.i. in both assay repeats. In contrast, *Lmj*cDNA16 sKI and *Lmj*HASPA2 sKI did not increase their HASPA mRNA levels until day 7 p.i. Conversely, the HASPA expression profiles of culture-derived FVI, cDNA16 sKI and HASPA2 sKI parasites are very similar to one another.

The results suggest that mutant HASPA and SHERP mRNA expression patterns, including that of *Lmj*cDNA16 sKI, are distinct from FVI both *in vivo* and *in vitro*. The only mRNA expression patterns in *Lmj*cDNA16 sKI that matched the one of FVI both *in vivo* and *in vitro* were the ones for HASPB mRNA, suggesting that HASPB expression at the right time and the right level is key for *Lmj*cDNA16 sKI's capacity to complete metacyclogenesis. This is supported by the observation that *Lmj*HASPB sKI does not increase its HASPB mRNA levels to similar levels as FVI and *Lmj*cDNA16 sKI *in vivo* and does not complete metacyclogenesis. The importance of HASPB in metacyclogenesis completion was previously suggested by observations made in an episomal HASPB replacement mutant line by Sádlová *et al.* (2010) (*146*).

Since it has been proposed that poly(A)-sites could be influenced by downstream

splice acceptor sites due to the coupling of 5' splicing and 3' polyadenylation (*313*, *314*), it was possible that HASP and SHERP mRNA lengths expressed from some of the constructs might be different from the parental line (FVI), because of the foreign 5' splice acceptor sites downstream of the single HASP and SHERP genes in the constructs. These 5' splice acceptor sites were further downstream than the native ones in the cDNA16 locus (Fig.5.8A & B). If the 3' UTR of an mRNA was artificially extended in this fashion, it could destabilize the mRNA and target it for degradation. However, why this effect should be restricted to midgut conditions and does not reduce mRNA levels in cultured parasites is not known. To check whether HASPB and SHERP mRNA lengths were changed compare to the parental line (FVI), two fragments were generating, using a spliced leader forward primer and an internal reverse primer for one fragment and an internal forward primer and an oligo-dT reverse primer for the second fragment (Fig.5.8C).

Unfortunately, no conclusive results are available yet as the investigation is still underway after encountering a technical problem with the PCR amplification.

5.6. Assessing osmotaxis capacity in cDNA16 mutant strains

It was observed during this study that parasites spread relatively evenly through the AMG and TMG in all tested lines, except for FVI and *Lmj*cDNA16 sKI. To investigate the reasons for this, the osmotactic capacity of some mutant lines (*Lmj*cDNA16 dKO, *Lmj*cDNA16 sKI, *Lmj*HASPB sKI and *Lmj*SHERP sKI) were assessed and compared to FVI. To do this, an osmotaxis assay was adapted from Oliveira *et al.* (2000) (*366*) and Leslie *et al.* (2002) (*367*) using 100 mM sucrose as an attractant (see 2.2.9.). Initially, the assay was performed in three repeats on FVI, *Lmj*cDNA16 dKO and *Lmj*cDNA16 sKI only; during the re-run *Lmj*HASPB sKI and *Lmj*SHERP sKI were also analysed. The attraction coefficient (A.C.) was calculated by dividing the number of parasite counted in the sample collected from capillaries containing the attractant (100 mM sucrose) by the number of parasites counted in the samples taken from the capillaries without attractant. An A.C. of ~1 is indicative of zero attraction of the parasites for the sucrose. Since there were six attractant positive and six attractant negative capillaries per sample per run and six (FVI, *Lmj*cDNA16 dKO and *Lmj*cDNA16 sKI) or three (*Lmj*HASPB sKI and *Lmj*SHERP sKI) repeats of the assay per tested strain, the A.C. could have been calculated using two distinct methods, which on occasion gave significantly different results. Firstly, the largest parasite count of the positive capillary samples was divided by the largest parasite count of the negative control capillaries for all six samples. Then the individual A.C.s were

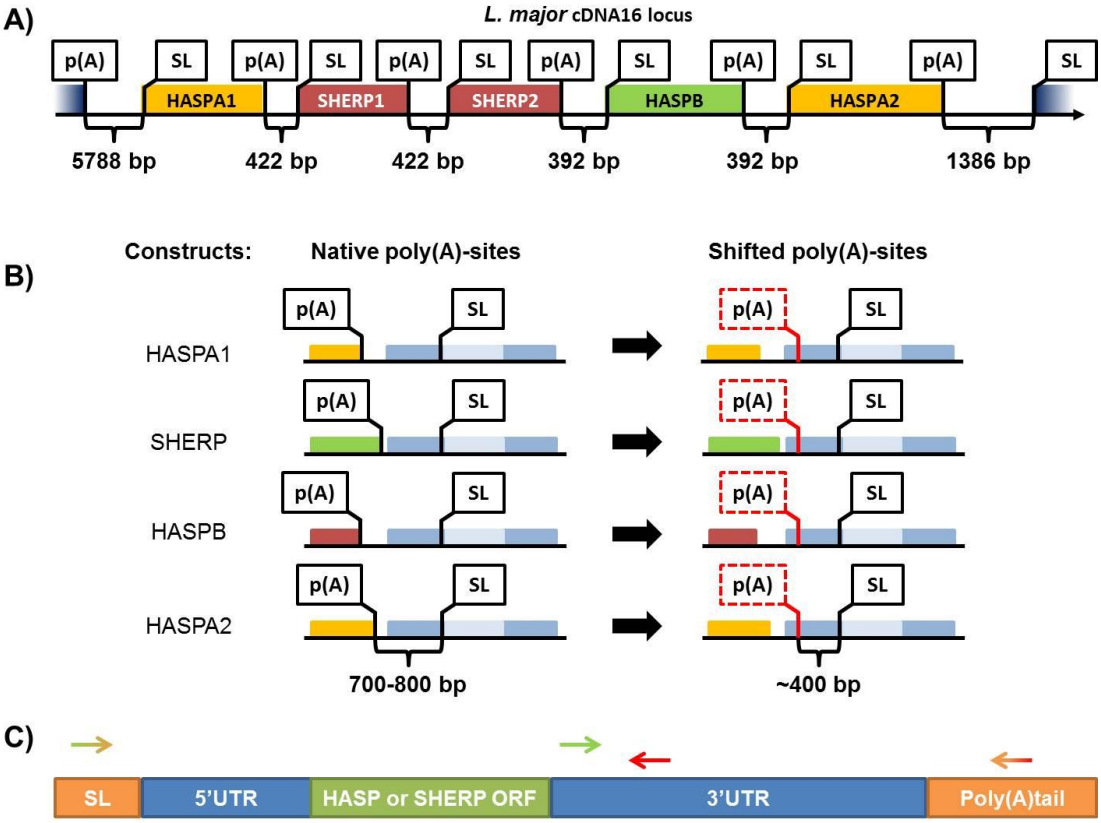


Fig.5.8 – PCR amplification to determine mRNA length

A) shows a schematic of the cDNA16 locus and the location and distance between 3' poly(A)sites and downstream 5' splice acceptor sites (not to scale). B) shows a schematic of the HASP and SHERP gene constructs (not to scale). The left side shows the native 3' poly(A)-sites and downstream 5' spliced leader acceptor site of the antibiotic resistance gene; the right side shows the 3' poly(A)-sites shifted downstream to be at the same distance from the downstream 5' spliced leader acceptor site as the genes are within the cDNA16 locus (see A). C) Schematic of the PCR approach chosen to investigate mRNA lengths. Two primer pairs (SL-primer (orange/green) / internal reverse primer (red) and internal forward primer (green) / oligo(dT) primer (red/orange)) were chosen to amplify the mRNA in all tested lines.

summed and divided by the number of paired samples to get the mean A.C.

Method 1: $$AC = \frac{\left(\frac{a_1}{b_1}+\frac{a_2}{b_2}+\frac{a_3}{b_3}+\frac{a_4}{b_4}+\frac{a_5}{b_5}+\frac{a_6}{b_6}\right)}{6}$$

Secondly, parasite counts were first summed in total and then, total parasites counts of the positive samples were divided by the total parasite counts of the negative samples to get the mean A.C.

Method 2: $$AC = \frac{a_1+a_2+a_3+a_4+a_5+a_6}{b_1+b_2+b_3+b_4+b_5+b_6}$$

The second method generally rendered a mean A.C. that was slightly smaller than the ones calculated using the first method; however, the trend pattern was always the same (Fig.5.10A). The assay is prone to strong variations between capillaries and test rounds, but any osmotaxis deficiency should have been readily detected according to literature (*366*, *367*). A Kruskal-Wallis test was used to establish any significant difference in the overall results, but none was found between the five tested strains whose mean A.C.s were all within each other's error range (Table 5.2). However, all five strains showed significant difference in their A.C.s compared to their negative controls ($P<0.001$ for all), which showed that all strains were osmotactically active (Fig.5.10B). Thus, all parasite mutant lines had the capacity to sense the posterior of the midgut and the reason why the mutant lines did not preferentially accumulate in TMG in late stage of sand fly infections, as in the case of FVI and *Lmj*cDNA16 sKI, was not due to parasite osmotactic deficiencies.

5.7. Looking for promastigote secretory gel secretion *in vivo*

The PSG is secreted by leptomonads in the TMG and both leptomonads and nectomonads attach to the gel with high affinity. It is possible that the PSG facilitates strong parasite colonization of the TMG by retaining parasites in the PSG matrix against peristalsis and the intake-flow of sucrose. A lack of PSG secretion had been suggested previously in *Lmj*cDNA16 dKO by Sádlová in her analysis of infected midgut wet mounts (unpublished) (Fig.4.2). This observation was supported in this study by the fact that mutant parasites swam immediately freely after releasing them from an infected TMG at day 12 PBM, while FVI and *Lmj*cDNA16 sKI parasites were immobilized initially after release. To confirm conclusively that a lack of PSG secretion occurred in the *Lmj*cDNA16 dKO, PSG extracts from day 12 PBM infected TMGs were prepared by collecting 10 infected

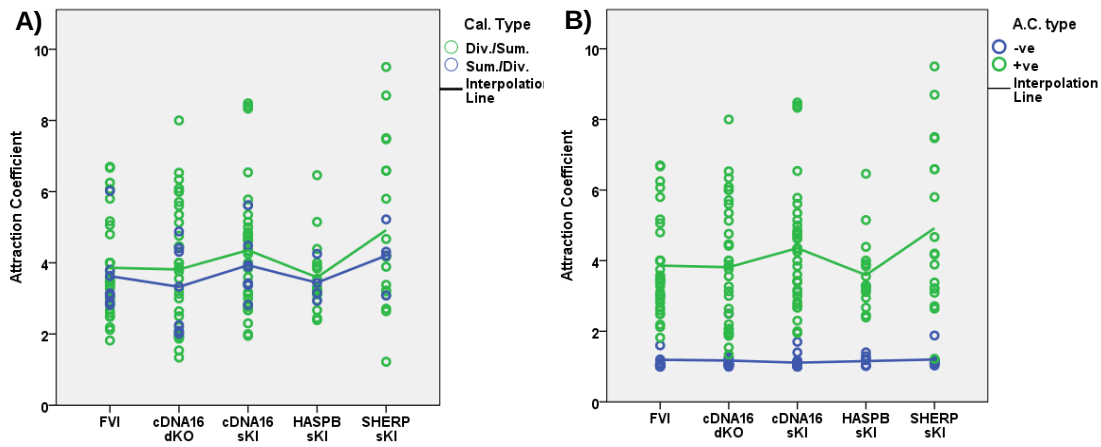


Fig.5.10 – Osmotaxis assay

A) The graph shows the A.C.s for the two calculation methods:

Method one (green): $$AC = \frac{\left(\frac{a_1}{b_1}+\frac{a_2}{b_2}+\frac{a_3}{b_3}+\frac{a_4}{b_4}+\frac{a_5}{b_5}+\frac{a_6}{b_6}\right)}{6}$$

Method two (blue): $$AC = \frac{a_1+a_2+a_3+a_4+a_5+a_6}{b_1+b_2+b_3+b_4+b_5+b_6}$$

The mean A.C.s are listed in Table 5.2. Neither approach showed any statistically significant differences between samples suggesting that osmotaxis is unimpaired in the mutants. B) The graph shows the A.C.s calculated by method 1 (green) and the A.C.s for the control experiment (blue), where no attractant was used. The results show a significant difference (P<0.001) in mean A.C.s in the presence of 100 mM sucrose compared to its absence, suggesting that parasites migrate ~4 times more often into capillaries with 100 mM sucrose than into sucrose negative capillaries.

Table 5.2 – Attraction Coefficient, Means and Standard Deviations

Strains	Method 1 Mean A.C. (Div./Sum)	S.D.	Method 2 Mean A.C. (Sum/Div.)	S.D.	Method 1 Mean A.C. (-ve Con.)
FVI	4.09	± 1.92	3.63	± 1.12	1.19
cDNA16 dKO	3.81	± 1.73	3.31	± 1.22	1.17
cDNA16 sKI	4.52	± 1.92	3.93	± 0.91	1.11
HASPB sKI	3.84	± 1.43	3.44	± 0.58	1.16
SHERP sKI	4.92	± 2.33	4.20	± 0.88	1.20

TMGs into 50 µl PBS (see 2.5.4). Samples were spun 6 times, with the supernatant transferred to a fresh tube after every spin to remove contaminating debris. 2-4 µl of supernatant were blotted onto an activated nitrocellulose membrane, which was blocked and probed with the LT15 antibody specific for the phosphoglycan disaccharide repeats [PO_4-6Gal(β1-4)Man(α1)] on LPG and selected PPGs, including fPPG (*54*).

The results showed that *Lmj*cDNA16 dKO had no positive signal for abLT15 binding suggesting that no soluble fPPG was present in the *Lmj*cDNA16 dKO supernatant, while FVI and *Lm*cDNA16 sKI had clear signals (Fig.5.11). The faint signal for *Lmj*cDNA16 dKO after only one round of spinning away debris suggested that parasite surface LPG and PPG contaminants were still present in the supernatant. Also, the *Lmj*HA1/2+S2/HB sKI supernatant lacked PSG, suggesting that all mutant lines did not produce PSG, except for *Lmj*cDNA16 sKI. In a repeat of the assay, debris pellets were also collected, lysed and blotted onto activated nitrocellulose membrane. FVI and *Lmj*cDNA16 sKI showed positive signals on abLT15 probing, while *Lmj*cDNA16 dKO did not (Fig.5.12A). This was surprising, because Fig.5.11 suggested the abLT15 would detect something in the *Lmj*cDNA16 dKO debris pellet, because the abLT15 had given a weak signal in the 1x spun *Lmj*cDNA16 dKO sample. *Lmj*HASPB sKI and *Lmj*HASPA2 sKI extracts were also analysed and did not show a positive signal either for pellet and supernatant. This confirmed the observations that *L. (L.) major* mutant lines without the complete cDNA16 locus do not produce PSG in the sand fly midgut. Interestingly, the abLT15 also did not detect anything in whole lysates of cultured parasites, including the FVI and *Lmj*cDNA16 sKI, although it was anticipated that parasite surface LPG and PPG would be detected by abLT15. This, however, does not suggest that mutant parasite lines grown in culture do not produce PSG, because the fPPG is secreted into the culture medium and removed with it, when isolating parasite cells for lysate production.

The PSG extracts and debris pellets were also probed for HASPB. It had been previously suggested that HASPB may be shedded within the midgut by metacyclics (*336*). Fig.5.12B shows that HASPB was readily detected in the parasite lysate samples of FVI, *Lmj*cDNA16 sKI and *Lmj*HASPB sKI as expected. Surprisingly, the *Lmj*cDNA16 dKO lysate sample also had a positive signal, although the mutant line does not contain a HASPB gene. This suggest unspecific antibody binding on the membrane. The only sample showing a positive signal for HASPB was the FVI PSG extract sample, while the debris pellet sample showed

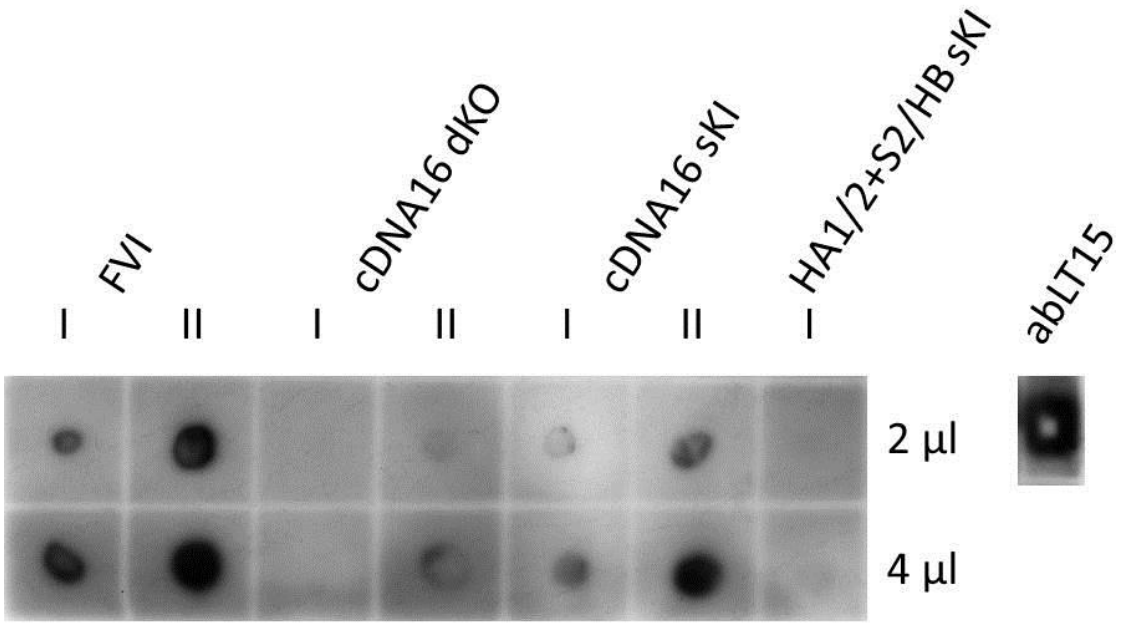


Fig.5.11 – Dot blot of PSG containing supernatants from midgut extracts

The dot blot was probed with the LT15 antibody for fPPG detection. Two sets of samples were blotted at two different volumes per parasite strain, with the exception of *Lmj*HA1/2+S2/HB sKI. The samples showed PSG extracts after 6 rounds of spinning them down and transferring them (I) and samples that had been spun down only once to clear away cell debris (II). After 6x spins (I) there is not more positive signal in *Lmj*cDNA16 dKO and *Lmj*HA1/2+S2/HB sKI samples, suggesting a severe lack of fPPG secretion at day 12 PBM resulting in the absence of PSG. The weak positive signal in the 1x spun down sample suggests the presence of contaminating parasite surface PPGs and LPG, which are also detected by the abLT15 according to M. E. Rogers, who kindly supplied the antibody. 1 µl of abLT15 had also been blotted as a secondary antibody control.

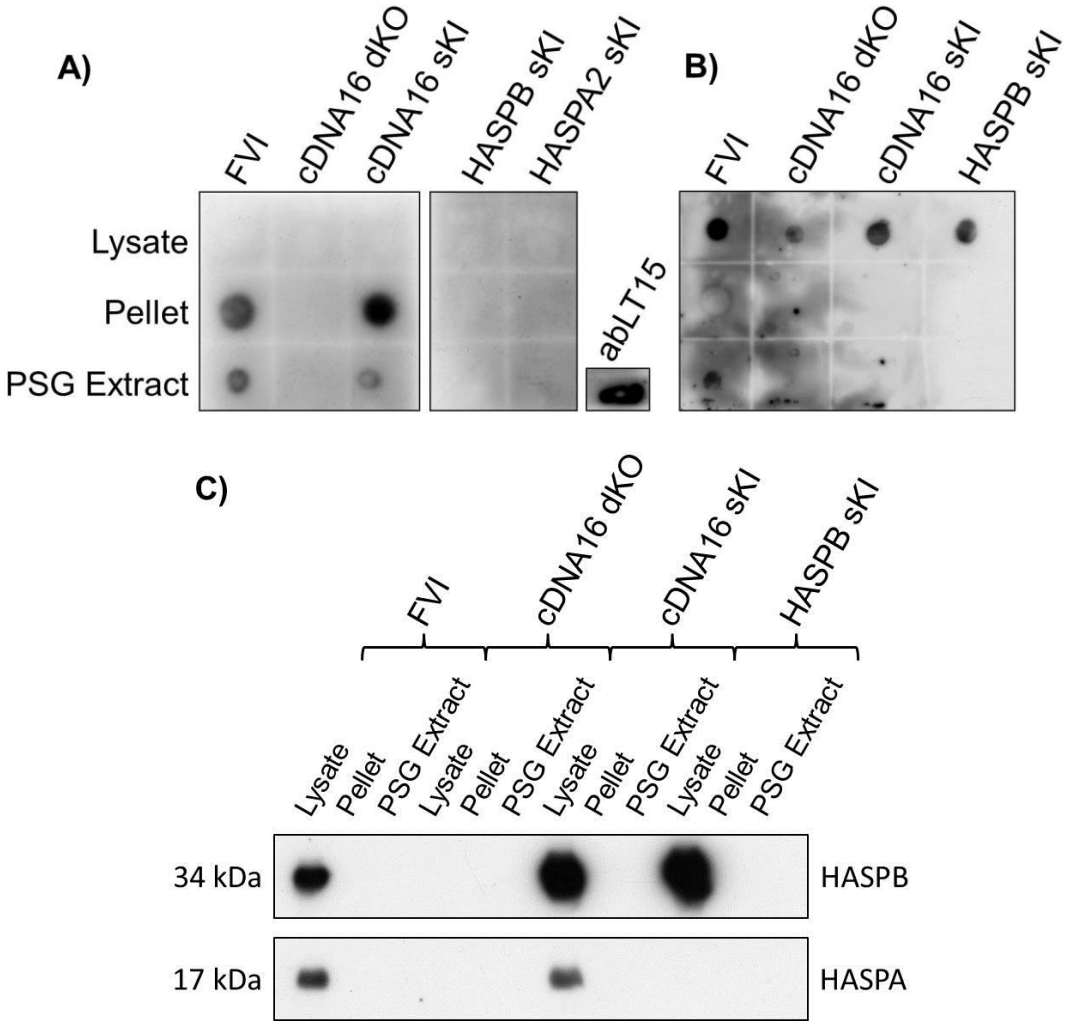


Fig.5.12 – Dot blot of PSG extracts from midgut

A repeat of the Dot blot with new extracts derived from thoracic sand fly midguts is shown. Whole lysates of cultured parasites were used as a control. A) Only FVI and *Lmj*cDNA16 sKI showed presence of PSG in pellet and supernatant, while *Lmj*cDNA16 dKO, *Lmj*HASPB sKI and *Lmj*HASPA2 sKI did not show any PSG signal when probed with the abLT15. The lysate samples were negative, although it had been expected that the abLT15 would detect parasite LPG and surface PPG too. B) The PSG extract were also probed with the non-affinity purified anti-HASP antibody to detect HASPB. The dot blot showed HASP staining in the lysates of FVI, *Lmj*cDNA16 sKI and *Lmj*HASPB sKI as expected, but unexpectedly also in *Lmj*cDNA16 dKO. Only the PSG extract from FVI was positive for HASPB, but not debris pellets. C) The dot blot results for HASPB staining were verified by Western blot. Here, no positive signal was observed in any the debris pellet lysates or PSG extracts, while HASPB was detected in the whole lysate samples from cultured parasites, as expected.

only a bleached out circle. *Lmj*cDNA16 sKI was also expected to show a positive signal for HASPB in the PSG extract and debris sample, but none was observed. However, since the *Lmj*cDNA16 dKO lysate did show a positive signal, it was put into question whether the positive HASPB signal in the FVI PSG extract was specific. A Western blot was produced of the same samples at the same quantities as had been used in the dot blot for verification. Fig.5.12C shows that only the parasite lysates of FVI, *Lmj*cDNA16 sKI and *Lmj*HASPB sKI were positive for HASPB. No PSG extract sample or debris pellet sample gave any positive signal even on longer exposure of the probed Western blot to film. Since HASPB expressing parasites had been present in the debris pellets of FVI and *Lmj*cDNA16 sKI before lysis, this could mean that the HASPB concentration was too low to be detected clearly on a Western blot.

5.8. Conclusions

The analysis of midgut derived parasites by confocal microscopy provided a method to look at HASP and SHERP expression without being limited by parasite numbers per midgut. Although midgut smears can be difficult to image by immunofluorescence due to high level background fluorescence, the analysis showed that there were marked differences for HASP and SHERP expression *in vitro* and *in vivo* for all mutant lines tested. Only FVI showed a positive signal for HASPB under both conditions, while SHERP could be detected in the *Lmj*S2+HB sKI mutant line, although HASPB was not detected in this line either. The lack of a positive signal in all mutant lines tested offered an explanation why no mutant line had managed to recover the parental line phenotype, but appeared in most aspects similar to the cDNA16 locus null phenotype.

Different possible explanations for the HASP and SHERP expression differences were explored in this part of the study. The rich M199 medium could be excluded from having an effect on construct regulation and the addition of blood fed sand fly midgut extracts from days 6 and 12 PBM to M199 cultures did not make a difference to HASPB expression in *Lmj*HASPB sKI either. This could mean that HASPB expression from the replacement constructs was not influenced by culture conditions or midgut derived molecules However, I cannot exclude that high dilution of midgut material or loss of molecules due to their size or insolubility. The methods would need to be refined for further analysis.

Further investigation into mRNA expression patterns showed that *Lmj*HASPB sKI had no mRNA peak at day 6 PBM, like FVI and *Lmj*cDNA16 sKI, which could

explain the lack of up-regulation of HASPB expression in that mutant line. Differences in the expression patterns of HASPA and SHERP between the parental line (FVI) and all mutant lines, including *Lmj*cDNA16 sKI, both *in vivo* and *in vitro*, suggest that regulation of HASPB at parental line levels and timing in *Lmj*cDNA16 sKI may be the key element for metacyclogenesis completion in *Lmj*cDNA16 sKI compared to all other mutant lines.

Investigations into the parasite's osmotaxis capacity to explain the even spread of all mutant lines in AMG and TMG, instead of accumulation in the TMG like FVI and *Lmj*cDNA16 sKI, showed that all mutant lines tested were attracted efficiently towards an independent sucrose source. This excluded osmotaxis as a potential mechanism requiring the HASP and SHERP function. However, the investigation into PSG secretion *in vivo* showed that all mutant lines tested had a clear lack of PSG secretion. This could be hypothesised to prevent parasite accumulation in the TMG, because nectomonads and leptomonads can bind to the PSG maintaining them in the TMG against peristalsis and intake-flow of sugar meals. Metacyclics can swim freely through the PSG and tend to accumulate at the poles of the PSG, which could mean that the fPPG matrix provides metacyclics with hold and orientation in the TMG. Further investigations would be required to determine if the HASPs and SHERP are directly involved in the pathways essential for fPPG synthesis and secretion or if this is only an indirect effect.

6. Chapter VI. – *L. (V.) braziliensis* orthologous HASP locus

6.1. Introduction

The broad aim of this part of the study was to investigate whether the locus equivalent to the *L. (L.) major* cDNA16 locus is also required for metacyclogenesis in the *L. (Viannia) spp.*, *L. (V.) braziliensis*. This required comparison of the phenotypes of *L. (V.) braziliensis* full orthologous HASP locus (OHL) deletion and replacement mutants *in vitro* and *in vivo* to the *L. (L.) major* full cDNA16 deletion (*Lmj*cDNA16 dKO) and replacement mutants (*Lmj*cDNA16 sKI). The OHL, which localizes in *L. (Viannia) spp.* to the same region on chromosome 23 as the cDNA16 locus in the *L. (Leishmania)* subgenus, was previously investigated by Depledge *et al.* (2010) (*294*). This study showed that the two annotated genes (LbrM.23.1110 and LbrM.23.1120), although distinct in sequence at the genomic level, expressed proteins with similar biochemical and structural properties to the HASPs. Due to these similarities and because frequently genes occurring in the same chromosomal context in different *Leishmania* spp. are functional homologues, it was hypothesised that the genes of the OHL in *L. (Viannia) spp.* may be functionally similar to those encoded by the cDNA16 locus in *L. (Leishmania) spp.* Based on these observations, it was hypothesised that the full deletion of the OHL may cause the same phenotype in *L. (V.) braziliensis* as the deletion of cDNA16 locus in *L. (L.) major*.

The available genomic sequence of OHL was recently assembled with reference to the cDNA16 locus in *L. (L.) major* which is known, however, to be misassembled due to the extensive repetitive sequence within the locus (see 1.7). Due to the similar high levels of sequence repetitiveness in the OHL, a repeat collapse had occurred in the assembly making the OHL appear smaller (~7 Kb) and with fewer genes than present in this segment of the genome (Fig.6.1). This was supported by the observation that the reanalysis of the automated genome assembly of the Lbr2904 genome (done by M. B. Rogers at the Sanger Institute) had a 4-5 fold increase in read depth compared to the surrounding sequence (Fig.6.2). The first step in this part of the study was to resolve the map of the OHL to be able to generate full OHL deletion and single OHL replacement mutants in *L. (V.) braziliensis* 2904 (reference strain).

6.2. Variation in the orthologous HASP locus between clinical *Leishmania (Viannia) braziliensis*

In order to address this repeat collapse, three genomic digests (*Pst*I – *Hind*III,

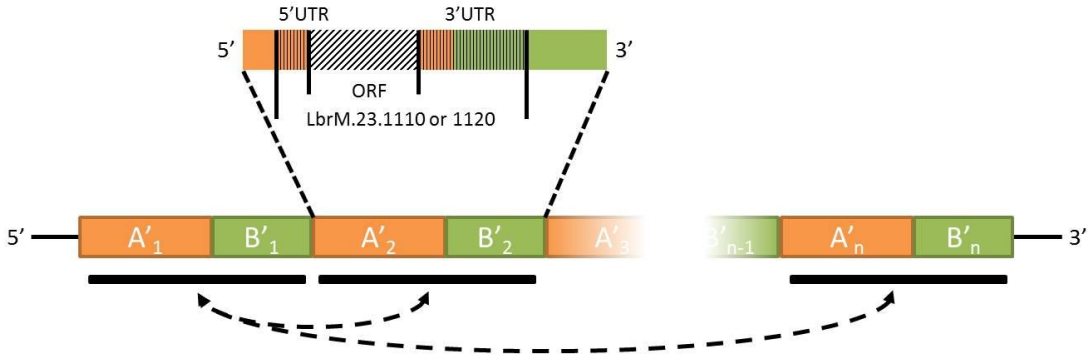


Fig.6.1 – Repeat collapse in OHL

The OHL consists of highly similar A'B' tandem repeats. A' contains the ORF and all single-nucleotide polymorphisms observed in the repeats. B' contains part of the 3' UTRs and the intergenic region, which are highly conserved. During automated sequence assembly, the A'B' motives were collapsed into a single A'B' unit, explaining the increased read depth for this region. Adapted from D. P. Depledge (PhD Thesis, 2009).

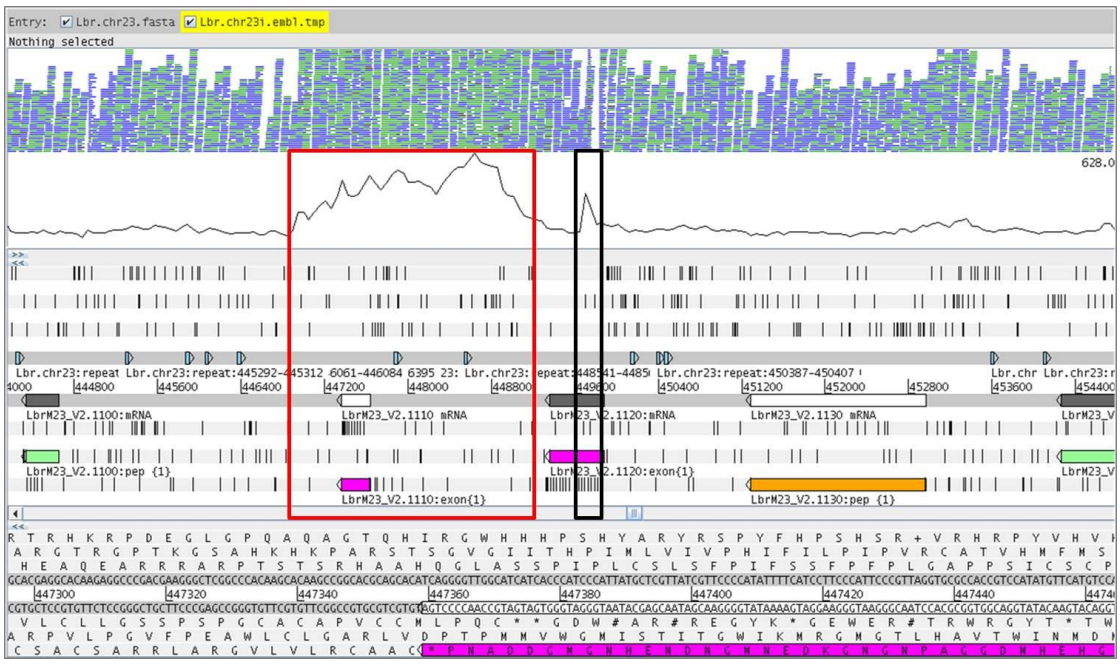


Fig.6.2 – Screen shot of Artemis software showing the reading depth for the OHL genes LbrM.23.1110 and LbrM.23.1120

The red and black boxes mark the areas of increased read depth compared to the neighbouring sequence, as represented in the elevated peaks of the graph in the second row (boxed). The ORFs of the two OHL genes are shown in pink. In particular, the area around the LbrM.23.1110 (1110) gene and the intergenic region between 1110 and LbrM.23.1120 (1120) show a 4-5 fold increase in read depth (red box), indicating the repeat collapse observed previously by Depledge *et al.* (2010). This analysis suggested that 1110 may occur as 4 – 5 copies in the OHL. Another smaller peak within the ORF of 1120 (black box) also suggested some sort of collapse, but one at a 2-fold level. As subcloning and sequencing showed later, there are two distinct ORFs for 1120. (The image was kindly provided by M. B. Rogers, Wellcome Trust Sanger Insitute)

*Xho*I or *Not*I – *Hind*III) were designed based on the available genomic sequence of *L. (V.) braziliensis* 2904 (*Lbr*2904) and performed individually. The digests were analysed on a Southern blot to determine the wild type OHL size. Together with *Lbr*2904, genomic digests of another available *L. (V.) braziliensis* strain (LTB 300) were run in parallel. The Southern blot (Fig.6.3) was probed with three different DIG-labelled probes (OHL I = binding to the intergenic region attached to 1110, OHL II = binding to the 5'flanking region, OHL III = binding to the 3'flanking region – Fig.6.4). Unfortunately, the first lane with LBT 300 gDNA was empty, because the gDNA digest was lost. Its size was later calculated based on the ratio of the other two pairs of *Lbr*2904 and LTB 300 measured bands observed. The resulting Southern blot (Fig.6.3) clarified the following points: (i) according to the available OHL sequence from GeneDB a genomic digest with either *Pst*I – *Hind*III, *Xho*I or *Not*I – *Hind*III should have a single fragment of the size 5.2 Kb, 5.6 Kb and 6.5 Kb, respectively (Fig.6.4A). Instead, bands were observed with 13.5 Kb (LTB300) and 14.9 Kb (*Lbr*2904) for *Pst*I – *Hind*III, 13.65 Kb (LTB 300) and 15 Kb (*Lbr*2904) for *Xho*I and 15 Kb (LTB 300) and 16.65 Kb (*Lbr*2904) for *Not*I – *Hind*III digests (Table 6.1). (ii) Comparing the digestion results for *Lbr*2904 and LTB 300, it was obvious that the locus size is not equivalent for these two *L. (V.) braziliensis* strains. The *Lbr*2904 OHL appeared larger for all three digests than the equivalent detected for LTB 300. These results confirmed that the data base sequence for the OHL was incomplete.

To investigate this variation of OHL size between *Lbr*2904 and LTB 300 further, a *Sac*I digest of gDNA was performed, which was designed to break the locus apart and generate individual fragments with LbrM.23.1110 and LbrM.23.1120, respectively (Fig.6.4). It was expected that each probe would only detect a single band on a Southern blot with the OHL I and OHL III probes detecting the same band. However, the OHL I probe was able to detect at least four distinct bands for *Lbr*2904 and three for LTB 300, of which two and one band was also detected by OHL III, respectively (Fig.6.5). Finding multiple bands with the OHL I probed suggested that the intergenic region was present several times rather than only once, underlining the observation made during the reanalysis of the automated re-assembly of the *Lbr*2904 genome, which had shown a 4-5 fold increase in reading depth for the region containing the 1110 gene (Fig.6.2).

To further analyse the *L. (V.) braziliensis* intraspecific OHL variability, gDNA samples from clinical *L. (V.) braziliensis* isolates taken from patients in Brazil (supplied by S. R. B. Uliana, Universidade de São Paulo, Brazil; investigated in

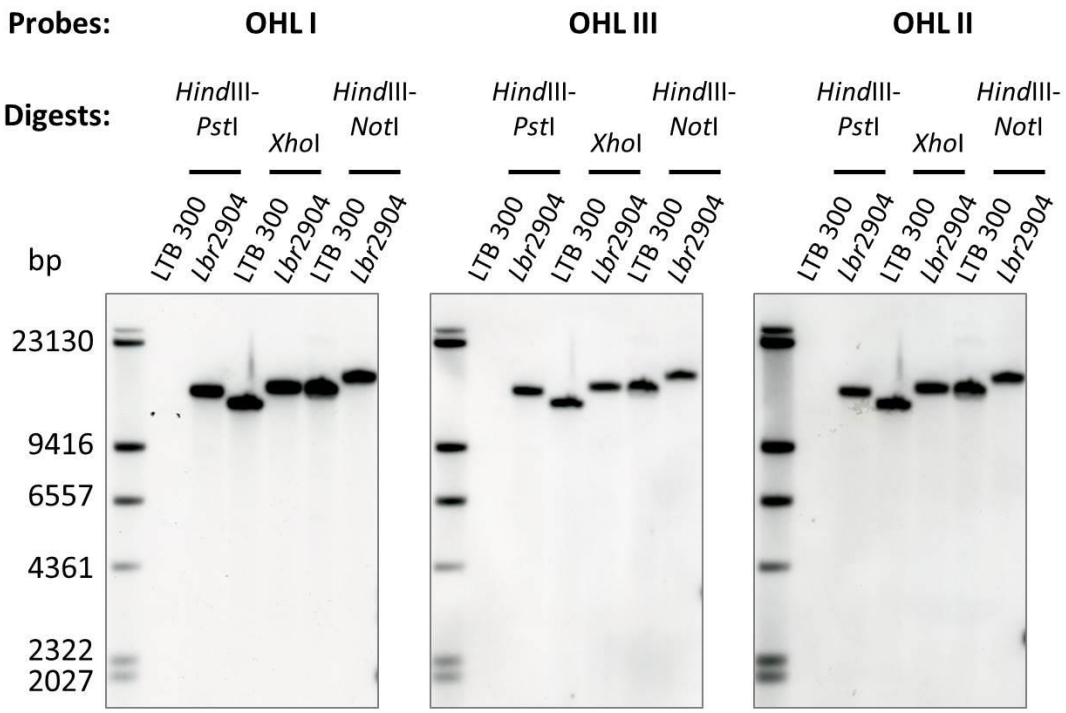


Fig.6.3 – Southern blot of three gDNA digests for two L. (V.) braziliensis strains
Genomic DNA from *L. (V.) braziliensis* 2904 and LTB 300 were restricted with three combinations of enzymes as shown, size separated and probed with three distinct DIG labelled probes (OHL I – III; Fig.6.4). All three DIG-labelled probes detected the same fragment for all three digests that were designed to cut out the entire OHL, confirming that the whole locus was isolated by the digest. Bands differ in size between LTB 300 and *Lbr*2904 for each digest, suggesting variation in the OHL content with the LTB 300 OHL smaller than the *Lbr*2904 OHL. The bands are also ~2.5x larger in size than the expected bands according to the available sequence for the OHL from GeneDB. (The gDNA *Hind*III - *Pst*I digest of the LTB 300 in the first lane was largely lost during precipitation, but a weak band, not visible on these blots, was visible against the light, which permitted determination of the migrated distance and band size).

Table 6.1 – Fragment size calculated by Fragment Size Calculator*

Restriction Enzymes	LTB300 (Kb)	Lbr2904 (Kb)
*Hind*III - *Pst*I	~13.5	~14.9
*Xho*I	~13.65	~15
*Hind*III - *Not*I	~15	~16.65

***Source:** http://www.basic.northwestern.edu/biotools/SizeCalc.html

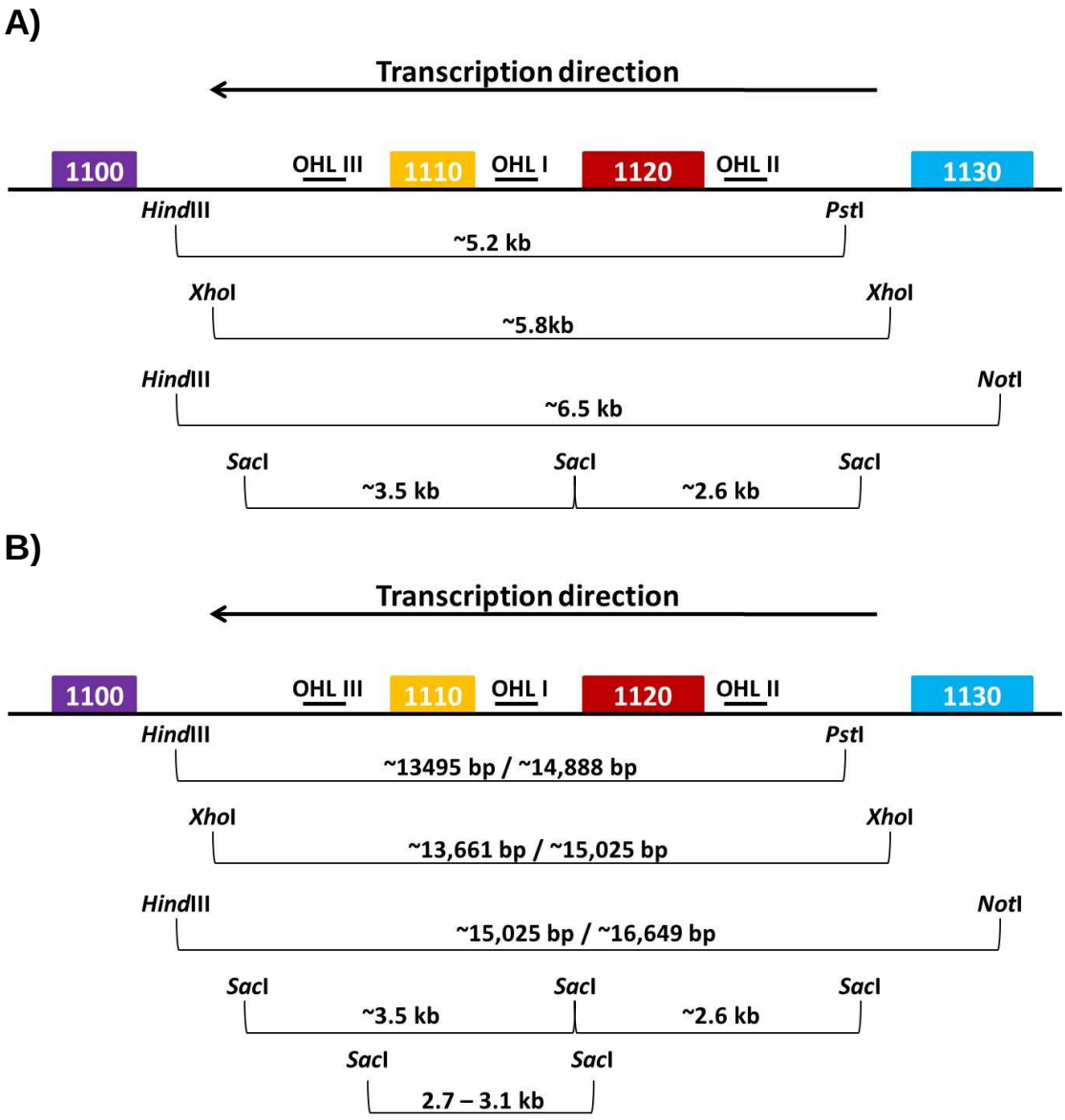


Fig.6.4 – Expected and measured sizes for the gDNA digests of LTB 300 and *Lbr*2904, respectively

A) The gDNA digests were planned using the available sequence of the OHL from GeneDB and the expected fragment sizes for each digest were calculated accordingly. B) The band sizes observed on the Southern blot (Fig.6.3) were estimated by measuring the distance travelled for each band against the distances travelled for the marker bands and applying those figures to the open source program "Fragment Size Calculator". (http://www.basic.northwestern.edu/biotools/SizeCalc.html). The calculated band sizes were on 2.3 – 2.6 times larger than the expected once. Notably, multiple bands were also detected after the *Sac*I digest containing the intergenic region and 1110 of varying sizes (Fig.6.5), suggesting that 1110 may occur as multiple gene copies in the OHL. This underlined the observed repeat collapse for this region, which was expect to have an at least 4-fold increase in reading depth during automated genome reassembly.

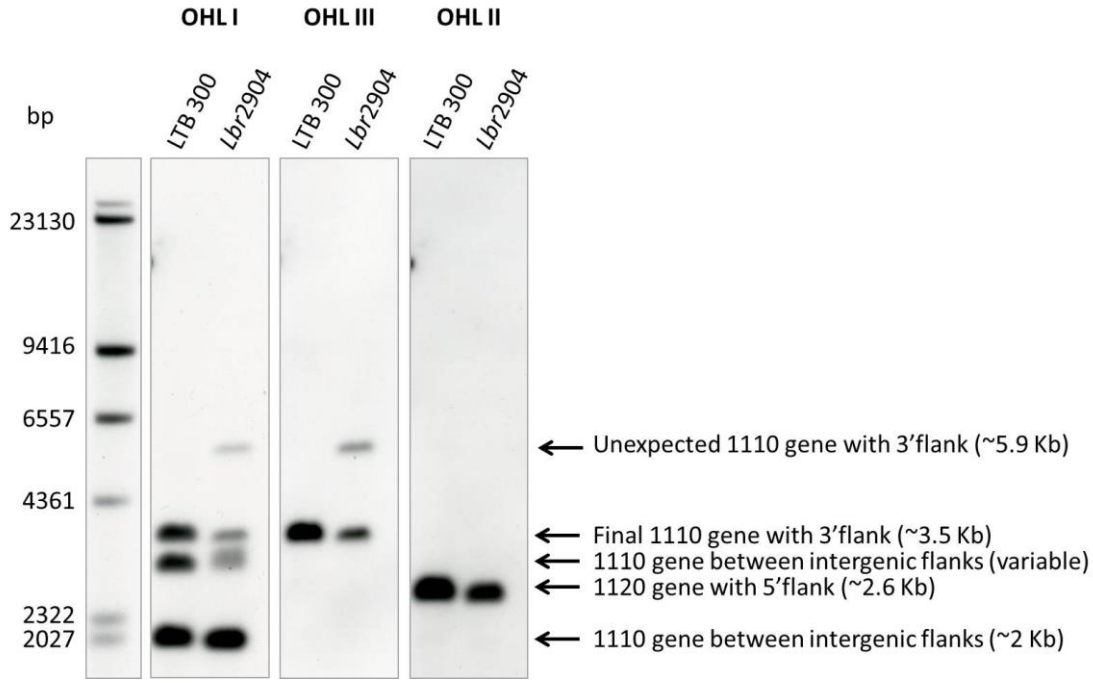


Fig.6.5 – Southern blot of *Sac*I digested gDNA of LTB 300 and *Lbr*2904

The *Sac*I digest was designed on the available sequence of the OHL to separate the 1110 and 1120 genes. For each DIG-labelled probe (OHL I – III; Fig.6.3B) a single band was expected, but for the OHL I probe binding to the intergenic region unique to the OHL 3 – 4 bands were detected, suggesting that the intergenic region occurs multiple times within the *L. (V.) braziliensis* genome. The 1120 fragment, which was not detected by the intergenic OHL I probe, was detected by probe OHL II and only a single band was detected as expected. Also the ~3.5 Kb fragment containing 1110, the 3' flanking region and the intergenic region up-stream of 1110 was detected. However, *Lbr*2904 also showed an unexpected band of ~5.9 Kb, which may be a fragment fused to one of the smaller unexpected fragments due to a mutation in the *Sac*I site.

Depledge *et al.* (2010)) were amplified using the GenomiPhi kit to generate enough material for a Southern blot. The efficacy of the kit was moderate, but enough gDNA was still available to perform a *Sac*I digest and run the gDNA digest on a gel for a Southern blot. Probing the blot with the OHL I probe, which hybridized to the intergenic region between the genes 1110 and 1120, a series of bands of variable size between the different clinical isolates were detected (Fig.6.6). This confirmed that the OHL was variable in size between *L. (V.) braziliensis* strains.

6.3. Addressing the *Leishmania (Viannia) braziliensis* orthologous HASP locus repeat collapse

In an attempt to resolve the full sequence and organisation of the OHL, different approaches were considered to overcome the problem of the high levels of sequence repetitiveness, which had caused the repeat collapse in the automated sequence assembly. Looking at the multiple bands detected by the OHL I probe, it was decided to focus on individual fragments of the OHL and determine their sequence. Several high fidelity PCR reactions with different sets of primers were set up including a reaction with a set of divergent primers in the intergenic region. Based on the available sequence, these divergent primers should not allow amplification of any fragments (Fig.6.7). However, all PCR reactions did generate fragments (Fig.6.8) and these were subcloned into the pCR-2.1 TOPO vector and transformed into chemically competent *E. coli* XL-1 cells. Clones were picked, screen and amplified. The plasmids were extracted and submitted for sequencing with appropriate primers.

The sequencing results for the different clones and fragments were assembled into contigs using the open source CAP3 Sequence Assembly Program and were analysed by sequence alignments using the open source ClustalW2 - Multiple Sequence Alignment software. The results showed two distinct ORF for 1120 (1120_v1 and 1120_v2), both with the same 5' and 3' UTRs and up- and downstream flanking regions (Fig.6.9). Nine distinct single-nucleotide polymorphisms (SNPs) were identified within these two sequences, which correlated with the expected number of 8 SNPs suggested by previous, yet unconfirmed, analysis (M. B. Rogers, unpublished). In addition, 1120_v2 contained one amino acid triplet and three repeat sequences more between its central section than 1120_v1. Interestingly, due to a premature stop-codon in 1120_v2, however, both ORF are 554 bp long. The upstream flanking region of both 1120 gene versions is also the upstream flanking region of the OHL and

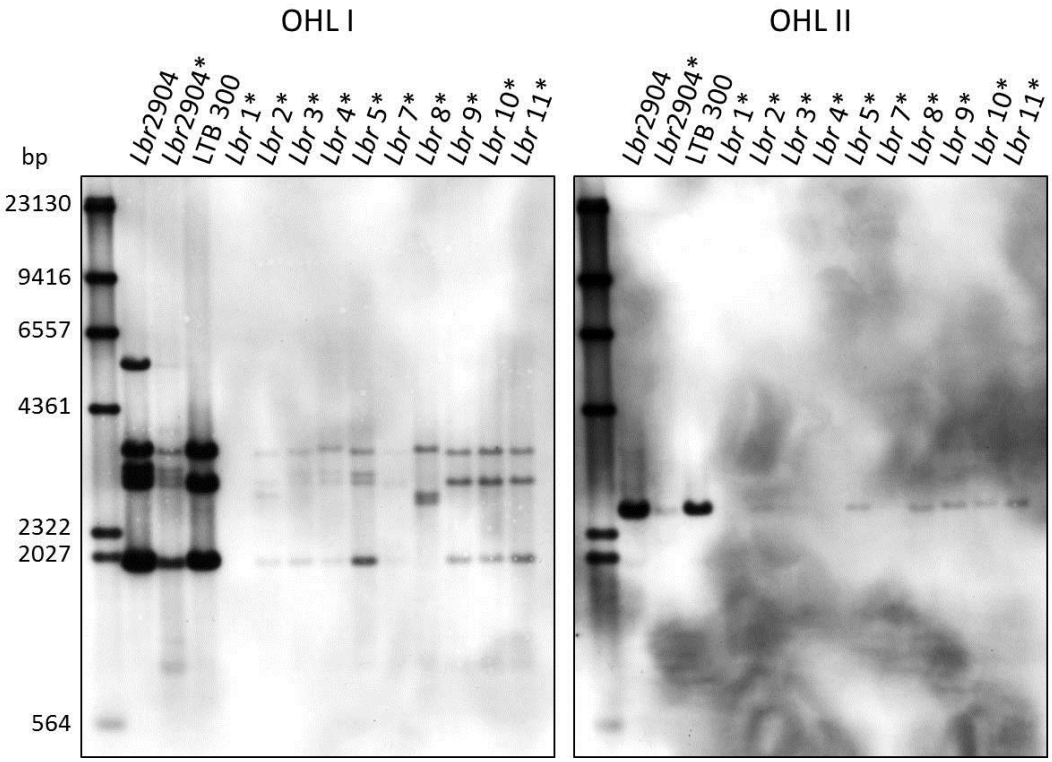


Fig.6.6 – Southern blot of *Sac*I digested GenomiPhi amplified gDNA from clinical *L. (V.) braziliensis* isolates to assess intra OHL variablilty

To assess OHL variability further between *L. (V.) braziliensis* strains, gDNA samples from clinical *L. (V.) braziliensis* isolates (*; supplied by Silvia R. B. Uliana, Universidade de São Paulo, Brazil; (*294*)) were amplified by the GenomiPhi kit (GE Healthcare) according to the supplier's protocol, *Sac*I digested and then size separated. Probing with the OHL I DIG-labelled probe in the intergenic region showed a variety of band patterns, supporting the hypothesis of OHL variability and multiple repeats of intergenic regions and LbrM.23.1110 genes. The OHL II probe detected only one band in all samples as expected (bands are not all visible in printed image, but are visible on the original blot), which suggested single copies of LbrM.23.1120 for all strains.

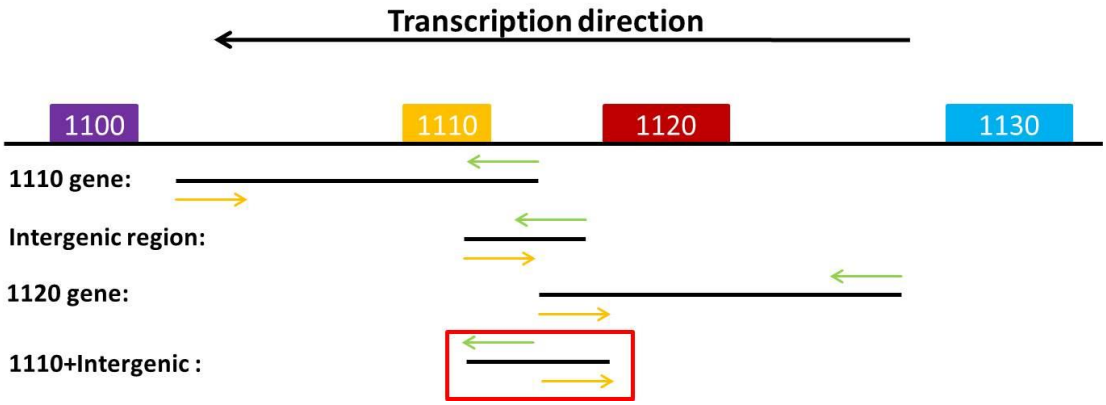


Fig.6.7 – Schematic for placement of primer in OHL used to amplify fragments for subcloning

Primer and amplification strategy was based on the available sequence for the OHL on GeneDB. Taking the findings from the *Sac*I digested gDNA Southern blots into account, which suggested multiple occurrences of the intergenic region, a primer pair (red box) was also picked facing away from one another in the intergenic region to test whether fragments could be amplified, confirming the presence of multiple intergenic region in the OHL.

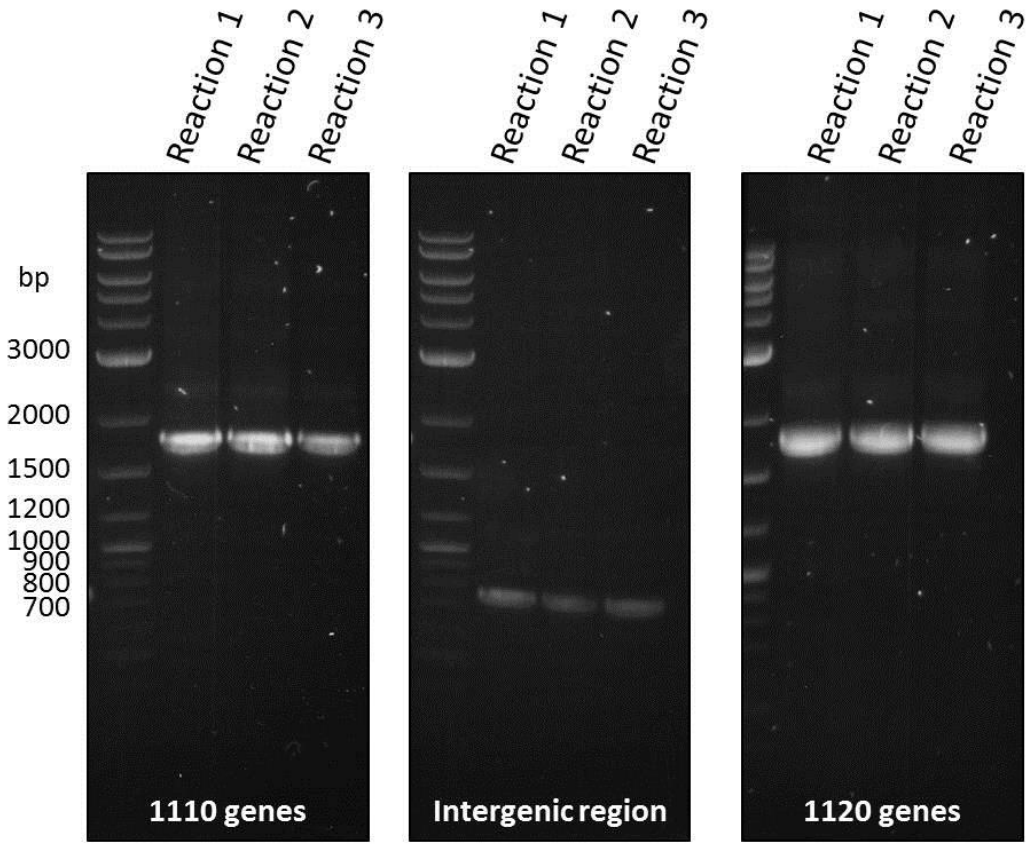


Fig.6.8 – PCR amplification of parts of the OHL for subcloning

To assess the OHL gene content four PCR reactions were set up with distinct primers to amplify fragments of the OHL for subcloning and sequencing (**Fig.6.7**). Three of these reaction were expected to generate single bands of expected sizes (1110 genes = ~1.8 Kb; Intergenic region = ~750 kb; 1120 genes = ~1.8 kb). These reactions always rendered the same band of the same size suggesting no size variation between alleles as suggested by Southern blots (Fig.6.5 & 6.6).

```
1120_v2_ CTCTCATACCAGCCCTCTGCGTTGTCGCTCGCTCACCGCCCCACCCCACTCCGTGTATCC 60
Lbr.1120 CTCTCATACCAGCCCTCTGCGTTGTCGCTCGCTCACCGCCCCACCCCACTCCGTGTATCC 60
1120_v1_ CTCTCATACCAGCCCTCTGCGTTGTCGCTCGCTCACCGCCCCACCCCACTCCGTGTATCC 60
         ************************************************************

1120_v2_ ACACCTTTCGCACCTCAACGCTTCTCTCCCTCAGCAGTTATCGCGCTATATTGGTGTGGC 120
Lbr.1120 ACACCTTTCGCACCTCAACGCTTCTCTCCCTCAGCAGTTATCGCGCTATATTGGTGTGGC 120
1120_v1_ ACACCTTTCGCACCTCAACGCTTCTCTCCCTCAGCAGTTACCGCGCTATATTGGTGTGGC 120
         ************************************ ******************

1120_v2_ TCTAAACCTACACTTACATCTGCTCCTCCTCTCTTCTTTCTCTCTCCCTTGACGCATACT 180
Lbr.1120 TCTAAACCTACACTTACATCTGCTCCTCCTCTCTTCTTTCTCTCTCCCTTGACGCATACT 180
1120_v1_ TCTAAACCTACACTTACATCTGCTCCTCCTCTCTTCTTTCTCTCTCCCTTGACGCATACT 180
         ************************************************************

1120_v2_ CTTCGCTGACGTGCGCCGTCGCAGGCTTTCCCCTTTACAGGCTCACCTACACGTCCCCCT 240
Lbr.1120 CTTCGCTGACGTGCGCCGTCGCAGGCTTTCCCCTTTACAGGCTCACCTACACGTCCCCCT 240
1120_v1_ CTTCGCTGACGTGCGCCGTCGCAGGCTTTCCCCTTTACAGGCTCACCTACACGTCCCCCT 240
         ************************************************************

1120_v2_ ACAGCGCAGTCTCCCAGTATTGTCACTTTCTAGTCTGGAATCCGCCAGCCTCAGCCCCTG 300
Lbr.1120 ACAGCGCAGTCTCCCAGTATTGTCACTTTCTAGTCTGGAATCCGCCAGCCTCAGCCCCTG 300
1120_v1_ ACAGCGCAGTCTCCCAGTATTGTCACTTTCTAGTCTGGAATCCGCCAGCCTCAGCCCCTG 300
         ************************************************************

1120_v2_ CGCACACTACACTCGATTCACTTATCCCCGTAGTAGCGCTTCACTCACCCTTAGCCGCTG 360
Lbr.1120 CGCACACTACACTCGATTCACTTATCCCCGTAGTAGCACTTCACTCACCCTTAGCCGCTG 360
1120_v1_ CGCACACTACACTCGATTCACTTATCCCCGTAGTAGCACTTCACTCACCCTTAGCCGCTG 360
         ********************************** ************************

1120_v2_ CCTTTCTTCCTCTACCCACTACTCTCTCACCCCACCTATAAGCAATGGGGACCATCTGTG 420
Lbr.1120 CCTTTCTTCCTCTACCCACTACTCTCTCACCCCACCTATAAGCAATGGGGACCATCTGTG 420
1120_v1_ CCTTTCTTCCTCTACCCACTACTCTCTCACCCCACCTATAAGCAATGGGGACCATCTGTG 420
         ************************************************************

1120_v2_ CGAAGCTGTCGCCGATGCCGCGCGGGACCAACCGTCCGACGAACCAGAAGGGCCGTGGCA 480
Lbr.1120 CGAAGCTGTCGCCGATGCCGCGCGGGACCAACCGTCCGACGAACCAGAAGGGCCGTGGCA 480
1120_v1_ CGAAGCTGTCGCCGATGCCGCGCGGGACCAACCGTCCGACGAACCAGAAGGGCCGTGGCA 480
         ************************************************************

1120_v2_ AAGGGAATAAGAAGAAGGGTGGTGGCCATCATAGACATGGGAAGAAGGATGGTGGCGACC 540
Lbr.1120 AAGGGAATAAGAAGAAGGGTGGTGGCCATCATAGACATGGGAAGAAGGATGGTGGCGACC 540
1120_v1_ AAGGGAATAAGA---AGGGTGGTGGCCATCATAGACATGGGAAGAAGGATGGCGGCGACC 537
         ************   ************************************ *******

1120_v2_ ATGGACATGAGAAGGTGAACGGCGGCGACCATGGACATGAACATATGGACGGCGGCCATC 600
Lbr.1120 ATGGACATGAGAAGGTGAACGGCGGCGACCATGGACATGAACATATGGACGGCGGCCATC 600
1120_v1_ ATGGACATGAGAAGGTGAACGGCGGCGACCATGGACATGAACATATGGACGGCGGCCATC 597
         ************************************************************

1120_v2_ ATGGACATGAGCATATGAACGGTGGCGACCATGGACATGAACATATGGACGGCGGCCAGC 660
Lbr.1120 ATGGACATGAGCATATGAACGGTGGCGACCATGGACATGAACATATGGACGGCGGCCAGC 660
1120_v1_ ATGGACATGAGCATATGAACGGTGGCGACCATGGACATGAACATATGGACGGCGGCCAGC 657
         ************************************************************

1120_v2_ ATGGACATGAACATATGGACGGTGGCGACCATGGACATGAACATATGGACGGTGGCGACC 720
Lbr.1120 ATGGACATGAACATATGGACGGTGGCGACCATGGACATGAACATATGGACGGTGGCGACC 720
1120_v1_ ATGGACATGAACATATGGACGGTGGCGACCATGGACATGAACATATGGACGGTGGCGACC 717
         ************************************************************

1120_v2_ ATGGACATGGGAATATGGACGGTGGCGACCATGGACATGAACATATGGACGGTGGCGACC 780
Lbr.1120 ATGGACATGGGAATATGGACGGTGGCGACCATGGACATGAACATATGGACGGTGGCGACC 780
1120_v1_ ATGGACATGGGAATATGGACGGTGGCGACCATGGACATGAACATATGGACGGTGGCGACC 777
         ************************************************************

1120_v2_ ATGGACATGAACATATGGACGGCGGCGACCATGGACATGAACATATGGACGGCGGCGACC 840
Lbr.1120 ATGGACATGAACATATGGACGGCGGCGACCATGGACATGAACATATGGACGGCGGCGACC 840
1120_v1_ ATGGACATGAACATATGGACGGTGGCGACCATGGACATGAACATATGAACGGTGGCGACC 837
         ********************** ********************** **** *******
```

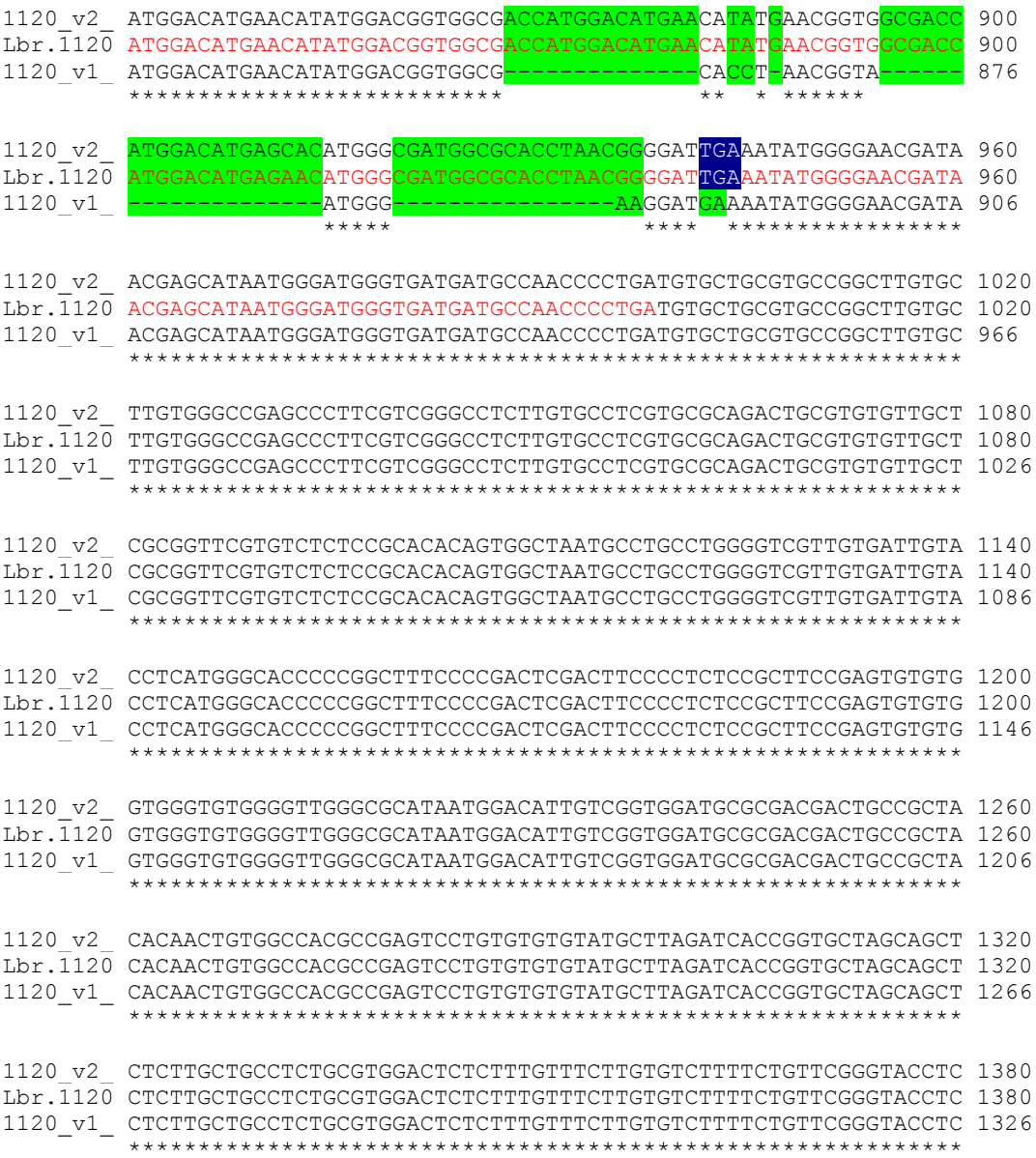


Fig.6.9 – Two heterozygous copies of LbrM.23.1120

Only two distinct sequences have been identified by subcloning and sequencing among several sequenced clones (shown here). Two distinct LbrM.23.1120 ORFs have been identified. The 5' flanking region of both LbrM.23.1120 copies is distinct from the intergenic region found between LbrM.23.1110 and LbrM.23.1120, which suggests that these two genes do not occur in a tandem repeat, but only once per allele and are heterozygous copies. The SNPs are marked in green and are restricted almost exclusively to the ORF (in red). Both ORFs have exactly the same length of 540 bp, although 1120_v2 copy contains three extra repeats and is also truncated due to a premature stop codon (in purple), reducing its size to 540 bp. The protein products are very similar at *N*-terminus and central repeat region, but have distinct *C*-termini (Fig.6.10).

distinct from the intergenic region between 1120 and 1110, which is flanking both 1120 gene versions downstream. Since the OHL upstream flank had not been detected anywhere else in the locus, this suggested that both 1120 gene version localized to the beginning of the OHL and were not tandemly repeated – that is, the two 1120 genes are heterozygous. At the protein level both products comprise of 179 aa and have a conserved *N*-terminus and internal repeat region, while the *C*-terminal regions are distinct (Fig.6.10A). A total of seven distinct 10 aa repeats were identified, which are all very similar (Fig.6.10B). Whether both protein products are functionally distinct remains to be determined.

The intergenic region showed the greatest level of sequence conservation, which was unexpected, because usually it is the non-functional intergenic regions which are prone to variations between species or even strains (Fig.6.11). This suggested that there is pressure for sequence conservation in this region, perhaps correlating with the presence of a SHERP homologue, which has a very short ORF (174 bp = 57 aa), within this region. Sequence analysis identified four potential ORFs within the ~1 Kb intergenic region (Appendix 12). The largest potential ORF contained two methionines, which could both serve as a translation initiation site resulting in a 363 and 405 bp ORF coding for a 120 and 134 amino acid long hypothetical protein, respectively. The other potential ORFs contained 219 bp, 192 bp and 177 bp and corresponded to 72, 63 and 58 amino acids long hypothetical proteins, respectively. However, BLAST searches against the amino acid sequences did not find any matches.

In the LbrM.23.1110 gene flanked downstream by the 3' OHL flanking region only two distinct sequences were observed among all the sequenced clones (Fig.6.12). These were distinct by a single synonymous SNP in position 147 (C→T) of the ORF. However, at least 5 distinct sequences were found for the LbrM.23.1110 ORFs flanked on both sides by the intergenic region with 7 SNPs in the ORF and three in the 3' UTR, ranging between 1 – 7 SNPs per sequence compared to the available sequence of Lbr.23.1110 on GeneDB (Fig.6.13). A total of 9 distinct SNPs were identified by comparison of all LbrM.23.1110 ORFs (Fig.6.14; Table 6.2); 6 matched SNP locations identified previously by comparing sequences of *L. (V.) braziliensis* 2904 and 2903 (available on TriTrypDB). 3 SNPs were synonymous changes, while 6 were non-synonymous. However, 5 of the six amino acid changes occurred within amino acid groups of strongly similar properties (Table 6.2); only one at amino acid position 7 was random (R→G). This suggested that all 1110 proteins are closely related and may have the same

A)

```
1120_v1   MGTICAKLSPMPRGTNRPTNQKGRGKGNKK-GGGHHRHGKKDGGDHGHEKVNGGDHGHE
1120_v2   MGTICAKLSPMPRGTNRPTNQKGRGKGNKKKGGGHHRHGKKDGGDHGHEKVNGGDHGHE
          ***************************** ***************************

1120_v1   HMDGGHHGHEHMNGGDHGHEHMDGGQHGHEHMDGGDHGHEHMDGGDHGHGNMDGGDHGH
1120_v2   HMDGGHHGHEHMNGGDHGHEHMDGGQHGHEHMDGGDHGHEHMDGGDHGHGNMDGGDHGH
          ************************************************************

1120_v1   EHMDGGDHGHEHMDGGDHGHEHMNGGDHGHEHMDGGAPNGNGKDENMGNDNEHNGMGDD
1120_v2   EHMDGGDHGHEHMDGGDHGHEHMDGGDHGHEHMDGGDHGHE---HMNGGDHGHEHMGDG
          ***********************:************    :   .  * *. *: ***

1120_v1   ANP--
1120_v2   APNGD
          *
```

B)

```
Rep-blue          GGDHGHGNMD
Rep-green         GGDHGHEKVN
Rep-red           GGHHGHEHMN
Rep-pink          GGDHGHEHMN
Rep-purple        GGDHGHEHMG
Rep-yellow        GGDHGHEHMD
Rep-grey          GGQHGHEHMD
                  **.*** .:
```

Fig.6.10 – Amino acid alignment of 1120_v1 and 1120_v2

A) Amino acid alignment of 1120_v1 and 1120_v2 proteins based on their ORFs. The alignment shows high levels of sequence identity in the *N*-terminal and central repeat region (coloured). 1120_v2 has an extra lysine at position 31 due to an inserted amino acid triplet (AAG). The change of T in the GTG codon to C (GCG) changes asparagine[141] to aspartic acid[141]. The three central repeat insertions in 1120_v2 and most polymorphisms were identified towards the *C*-terminus, which is distinct between 1120_v1 and 1120_v2. Both proteins consist of 179 aa, however. B) Alignment of distinct internal 10 aa repeats. Seven distinct repeat sequence of 10 aa were identified, with similarity to one another. The GGDHGHEHMD (yellow) is the most frequently occurring, either following itself or alternating with the other six sequences, which occur only once per protein version. The GGDHGHEHMG (purple) is distinct for 1120_v2, although it is only the *C*-terminal glycine, which distinguishes it.

```
Inter-reg-1 TCCCTTTGCTTCTGTGTGTGTGTGTGCGTGTGTGGTCTCGTTGGCCTCATGCCAGACTAG 60
Inter-reg-3 TCCCTTTGCTTCTGTGTGTGTGTGTGCGTGTGTGGTCTCGTTGGCCTCATGCCAGACTAG 60
Inter-reg-4 TCCCTTTGCTTCTGTGTGTGTGTGTGCGTGTGTGGTCTCGTTGGCCTCATGCCAGACTAG 60
Lbr2904     TCCCTTTGCTTCTGTGTGTGTGTGTGCGTGTGTGGTCTCGTTGGCCTCATGCCAGACTAG 60
Inter-reg-5 TCCCTTTGCTTCTGTGTGTGTGTGTGCGTGTGTGGTCTCGTTGGCCTCATGCCAGACTAG 60
            ************************************************************

Inter-reg-1 TACAGGATGCACATCCGCCCTTTCGCTGCCCCTGTCTTCTTTTCCTCGTCTGCTCTCTCT 120
Inter-reg-3 TACAGGATGCACATCCGCCCTTTCGCTGCCCCTGTCTTCTTTTCCTCGTCTGCTCTCTCT 120
Inter-reg-4 TACAGGATGCACATCCGCCCTTTCGCTGCCCCTGTCTTCTTTTCCTCGTCTGCTCTCTCT 120
Lbr2904     TACAGGATGCACATCCGCCCTTTCGCTGCCCCTGTCTTCTTTTCCTCGTCTGCTCTCTCT 120
Inter-reg-5 TACAGGATGCACATCCGCCCTTTCGCTGCCCCTGTCTTCTTTTCCTCGTCTGCTCTCTCT 120
            ************************************************************

Inter-reg-1 CTCTCTTCTCCTTGAGGGGCTTTTTCTTTCCTTCATCATTCCGTCTATCTCTTTGTGTAC 180
Inter-reg-3 CTCTCTTCTCCTTGAGGGGCTTTTTCTTTCCTTCATCATTCCGTCTATCTCTTTGTGTAC 180
Inter-reg-4 CTCTCTTCTCCTTGAGGGGCTTTTTCTTTCCTTCATCATTCCGTCTATCTCTTTGTGTAC 180
Lbr2904     CTCTCTTCTCCTTGAGGGGCTTTTTCTTTCCTTCATCATTCCGTCTATCTCTTTGTGTAC 180
Inter-reg-5 CTCTCTTCTCCTTGAGGGGCTTTTTCTTTCCTTCATCATTCCGTCTATCTCTTTGTGTAC 180
            ************************************************************

Inter-reg-1 GAGCTTGCGGTGCCTCTGTTTTCGAACATTTTCTCCTCTTTGGGGGGAGCCCTTCCCCCT 240
Inter-reg-3 GAGCTTGCGGTGCCTCTGTTTTCGAACATTTTCTCCTCTTTGGGGGGAGCCCTTCCCCCT 240
Inter-reg-4 GAGCTTGCGGTGCCTCTGTTTTCGAACATTTTCTCCTCTTTGGGGGGAGCCCTTCCCCCT 240
Lbr2904     GAGCTTGCGGTGCCTCTGTTTTCGAACATTTTCTCCTCTTTGGGGGGAGCCCTTCCCCCT 240
Inter-reg-5 GAGCTTGCGGTGCCTCTGTTTTCGAACATTTTCTCCTCTTTGGGGGGAGCCCTTCCCCCT 240
            ************************************************************

Inter-reg-1 CTTTCCCCGTCCGGTGCACGTGTTTGCCACTCTTTTTCGTTTCGTTCTTCTGATGGCAGC 300
Inter-reg-3 CTTTCCCCGTCCGGTGCACGTGTTTGCCACTCTTTTTCGTTTCGTTCTTCTGATGGCAGC 300
Inter-reg-4 CTTTCCCCGTCCGGTGCACGTGTTTGCCACTCTTTTTCGTTTCGTTCTTCTGATGGCAGC 300
Lbr2904     CTTTCCCCGTCCGGTGCACGTGTTTGCCACTCTTTTTCGTTTCGTTCTTCTGATGGCAGC 300
Inter-reg-5 CTTTCCCCGTCCGGTGCACGTGTTTGCCACTCTTTTTCGTTTCGTTCTTCTGATGGCAGC 300
            ************************************************************

Inter-reg-1 GAGCGGCGGCTGTGGCCTGCTGGGATGAGGTGTGGAGTGTGCCTGTCTGCGCACCGCTCT 360
Inter-reg-3 GAGCGGCGGCTGTGGCCTGCTGGGATGAGGTGTGGAGTGTGCCTGTCTGCGCACCGCTCT 360
Inter-reg-4 GAGCGGCGGCTGTGGCCTGCTGGGATGAGGTGTGGAGTGTGCCTGTCTGCGCACCGCTCT 360
Lbr2904     GAGCGGCGGCTGTGGCCTGCTGGGATGAGGTGTGGAGTGTGCCTGTCTGCGCACCGCTCT 360
Inter-reg-5 GAGCGGCGGCTGTGGCCTGCTGGGATGAGGTGTGGAGTGTGCCTGTCTGCGCACCGCTCT 360
            ************************************************************

Inter-reg-1 ATTTCCGTGTCTCCACCCTCCCAAGCTGCCCACGTCCCCGCACGCGAGTCTGCCGGGGCA 420
Inter-reg-3 ATTTCCGTGTCTCCACCCTCCCAAGCTGCCCACGTCCCCGCACGCGAGTCTGCCGGGGCA 420
Inter-reg-4 ATTTCCGTGTCTCCACCCTCCCAAGCTGCCCACGTCCCCGCACGCGAGTCTGCCGGGGCA 420
Lbr2904     ATTTCCGTGTCTCCACCCTCCCAAGCTGCCCACGTCCCCGCACGCGAGTCTGCCGGGGCA 420
Inter-reg-5 ATTTCCGTGTCTCCACCCTCCCAAGCTGCCCACGTCCCCGCACGCGAGTCTGCCGGGGCA 420
            ************************************************************

Inter-reg-1 GAAATGCTGTACTGCGCCGTAATAAAGGAAAACACGGAGACGAAGTGCGCCGGCGCCAGA 480
Inter-reg-3 GAAATGCTGTACTGCGCCGTAATAAAGGAAAACACGGAGACGAAGTGCGCCGGCGCCAGA 480
Inter-reg-4 GAAATGCTGTACTGCGCCGTAATAAAGGAAAACACGGAGACGAAGTGCGCCGGCGCCAGA 480
Lbr2904     GAAATGCTGTACTGCGCCGTAATAAAGGAAAACACGGAGACGAAGTGCGCCGGCGCCAGA 480
Inter-reg-5 GAAATGCTGTACTGCGCCGTAATAAAGGAAAACACGGAGACGAAGTGCGCCGGCGCCAGA 480
            ************************************************************

Inter-reg-1 GCACACGCTCACGCACACGTACACGGGCGCTGCGTGGTACGGTTTAGTGGATGGCACGCC 540
Inter-reg-3 GCACACGCTCACGCACACGTACACGGGCGCTGCGTGGTACGGTTTAGTGGATGGCACGCC 540
Inter-reg-4 GCACACGCTCACGCACACGTACACGGGCGCTGCGTGGTACGGTTTAGTGGATGGCACGCC 540
Lbr2904     GCACACGCTCACGCACACGTACACGGGCGCTGCGTGGTACGGTTTAGTGGATGGCACGCC 540
Inter-reg-5 GCACACGCTCACGCACACGTACACGGGCGCTGCGTGGTACGGTTTAGTGGATGGCACGCC 540
            ************************************************************

Inter-reg-1 TAAGCACTGTGAGAAAGCGGCGCGCTCTCTCTCCAAAAAGGCACACTCGCCATGGTGGTG 600
Inter-reg-3 TAAGCACTGTGAGAAAGCGGCGCGCTCTCTCTCCAAAAAGGCACACTCGCCATGGTGGTG 600
Inter-reg-4 TAAGCACTGTGAGAAAGCGGCGCGCTCTCTCTCCAAAAAGGCACACTCGCCATGGTGGTG 600
Lbr2904     TAAGCACTGTGAGAAAGCGGCGCGCTCTCTCTCCAAAAAGGCACACTCGCCATGGTGGTG 600
Inter-reg-5 TAAGCACTGTGAGAAAGCGGCGCGCTCTCTCTCCAAAAAGGCACACTCGCCATGGTGGTG 600
            ************************************************************
```

```
Inter-reg-1 AGCCCCTTTCGTCGCTTGCTTTTCTGTTTCAGCCCCCAGTGCGAATGCACATGTGTACTT 660
Inter-reg-3 AGCCCCTTTCGTCGCTTGCTTTTCTGTTTCAGCCCCCAGTGCGAATGCACATGTGTACTT 660
Inter-reg-4 AGCCCCTTTCGTCGCTTGCTTTTCTGTTTCAGCCCCCAGTGCGAATGCACATGTGTACTT 660
Lbr2904     AGCCCCTTTCGTCGCTTGCTTTTCTGTTTCAGCCCCCAGTGCGAATGCACATGTGTACTT 660
Inter-reg-5 AGCCCCTTTCGTCGCTTGCTTTTCTGTTTCAGCCCCCAGTGCGAATGCACATGTGTACTT 660
            ************************************************************

Inter-reg-1 ACGGGCTGCGGTTGCTGCGAAAGGAACAAGCTAACATGCCCGGGGCACCGCATTTCTGTT 720
Inter-reg-3 ACGGGCTGCGGTTGCTGCGAAAGGAACAAGCTAACATGCCCGGGGCACCGCATTTCTGTT 720
Inter-reg-4 ACGGGCTGCGGTTGCTGCGAAAGGAACAAGCTAACATGCCCGGGGCACCGCATTTCTGTT 720
Lbr2904     ACGGGCTGCGGTTGCTGCGAAAGGAACAAGCTAACATGCCCGGGGCACCGCATTTCTGTT 720
Inter-reg-5 ACGGGCTGCGGTTGCTGCGAAAGGAACAAGCTAACATGCCCGGGGCACCGCATTTCTGTT 720
            ************************************************************

Inter-reg-1 AAAGAACCTTCGATCCTGCTGCGTGTTTCTCT 752
Inter-reg-3 AAAGAACCTTCGATCCTGCTGCGTGTTTCTCT 752
Inter-reg-4 AAAGAACCTTCGATCCTGCTGCGTGTTTCTCT 752
Lbr2904     AAAGAACCTTCGATCCTGCTGCGTGTTTCTCT 752
Inter-reg-5 AAAGAACCTTCGATCCTGCTGCGTGTTTCTCT 752
            ********************************
```

Fig.6.11 – Conserved intergenic region between OHL genes

The intergenic region between LbrM.23.1110 and LbrM.23.1120 shown is highly conserved (100% identity) between all sequenced clones. No SNPs were found.

```
1110+3'f-2  CACTGCTCTCTCTCCCCCACCTATAAGCAATGGGGACCGCCTGTATGAGGGAGTTGACGA 717
1110+3'f-1  CACTGCTCTCTCTCCCCCACCTATAAGCAATGGGGACCGCCTGTATGAGGGAGTTGACGA 720
Lbr1110+3'f CACTGCTCTCTCTCCCCCACCTATAAGCAATGGGGACCGCCTGTATGAGGGAGTTGACGA 717
            ************************************************************

1110+3'f-2  GGCCGCGCACGTTCGACCTTAAGGCTCACGGGATGGGCGGTGGCAAAGGGGATAGGGCGA 777
1110+3'f-1  GGCCGCGCACGTTCGACCTTAAGGCTCACGGGATGGGCGGTGGCAAAGGGGATAGGGCGA 780
Lbr1110+3'f GGCCGCGCACGTTCGACCTTAAGGCTCACGGGATGGGCGGTGGCAAAGGGGATAGGGCGA 777
            ************************************************************

1110+3'f-2  ACGGCGGCGAGCATGGACATGAACATATGAACGGCGGCGACCATGGACATGAACACATGG 837
1110+3'f-1  ACGGCGGCGAGCATGGACATGAACATATGAACGGCGGCGACCATGGACATGAACATATGG 840
Lbr1110+3'f ACGGCGGCGAGCATGGACATGAACATATGAACGGCGGCGACCATGGACATGAACATATGG 837
            ************************************************** ****


1110+3'f-2  ACGGTGGCGACCATGGACATGAACATATGGACGGTGGCGCACCTAACGGGAATGGGAAGG 897
1110+3'f-1  ACGGTGGCGACCATGGACATGAACATATGGACGGTGGCGCACCTAACGGGAATGGGAAGG 900
Lbr1110+3'f ACGGTGGCGACCATGGACATGAACATATGGACGGTGGCGCACCTAACGGGAATGGGAAGG 897
            ************************************************************

1110+3'f-2  ATGAAAATATGGGGAACGATAACGAGCATAATGGGATGGGTGATGATGCCAACCCCTGAT 957
1110+3'f-1  ATGAAAATATGGGGAACGATAACGAGCATAATGGGATGGGTGATGATGCCAACCCCTGAT 960
Lbr1110+3'f ATGAAAATATGGGGAACGATAACGAGCATAATGGGATGGGTGATGATGCCAACCCCTGAT 957
            ************************************************************

1110+3'f-2  GTGCTGCGTGCCGGCTTGTGCTTGTGGGCCGAGCCCTTCGTCGGGCCTCTTGTGCCTCGT 1017
1110+3'f-1  GTGCTGCGTGCCGGCTTGTGCTTGTGGGCCGAGCCCTTCGTCGGGCCTCTTGTGCCTCGT 1020
Lbr1110+3'f GTGCTGCGTGCCGGCTTGTGCTTGTGGGCCGAGCCCTTCGTCGGGCCTCTTGTGCCTCGT 1017
            ************************************************************
------------------------------------------------------------------------

1110+3'f-2  AAAGAAAAGCATAGTAGCGCAGCGCTCTCGTTCGTGCTGGTCCCTCGCTCTCGCCCCCCT 2637
1110+3'f-1  AAAGAAAAGCATAGTAGCGCAGCGCTCTCGTTCGTGCTGGTCCCTCGCTCTCGCCCCCCT 2640
Lbr1110+3'f AGAGAAAAGCATAGTAGCGCAGCGCTCTCGTTCGTGCTGGTCCCTCGCTCTCGCCCCCCT 2637
            *.************************************************************

1110+3'f-2  TTTTATTGTTTACTGAGTCTGCATGCGCTTGTGGAGGTTGCTCTCGCACCACGCCATCGA 2697
1110+3'f-1  TTTTATTGTTTACTGAGTCTGCATGCGCTTGTGGAGGTTGCTCTCGCACCACGCCATCGA 2700
Lbr1110+3'f TTTTATTGTTTACTGAGTCTGCATGCGCTTGTGGAGGTTGCTCTCGCACCACGCCATCGA 2697
            ************************************************************

1110+3'f-2  TGAGGAAAGGGCGAGATAAAAACAACCCATCAGGTAACTTTCATGATATCAGCCTCTCTC 2757
1110+3'f-1  TGAGGAAAGGGCGAGATAAAAACAACCCATCAGGTAACTTTCATGATATCAGCCTCTCTC 2760
Lbr1110+3'f TGAGGAAAGGGCGAGATAAAAACAACCCATCAGGTAACTTTCATGATATCAGCCTCTCTC 2757
            ************************************************************

1110+3'f-2  TGCCTGTCTCTGTCTGTCTCTCCGCGTCTGTGTCTGTGTGTGTGTGTGTGTTGTGCTATG 2817
1110+3'f-1  TGCCTGTCTCTGTCTGTCTCTCCGCGTCTGTGTCTGTGTGTGTGTGTGTGTTGTGCTATG 2820
Lbr1110+3'f TGCCTGTCTCTGTCTGTCTCTCCGCGTCTGTGTCTGTGTGTGTGTGTGTGTTGTGCTATG 2817
            ************************************************************

1110+3'f-2  GGGGTCTGCATGCGAAGCTTTGTGGCACTAACACCCCTCGCTTTGAAGGTAGCTCGAGAA 2877
1110+3'f-1  GGGGTCTGCATGCGAAGCTTTGTGGCACTAACACCCCTCGCTTTGAAGGTAGCTCGAGAA 2880
Lbr1110+3'f GGGGTCTGCATGCGAAGCTTTGTGGCACTAACACCCCTCGCTTTGAAGGTAGCTCGAGAA 2877
            ************************************************************

1110+3'f-2  AAGAATTGCTATAGTGAGTCACGTTAAACAGCGAACTTAGGGAAGGCGGAAAGGGTTGAT 2937
1110+3'f-1  AAGAAATGCTATAGTGAGTCACGTTAAACAGCGAACTTAGGGAAGGCGGAAAGGGTTGAT 2940
Lbr1110+3'f AAGAATTGCTATAGTGAGTCACGTTAAACAGCGAACTTAGG-AAGGCGGAAAGGTTTGAT 2936
            *****:****************************************** ************

1110+3'f-2  ACATGTCGCAAGGTGAGCTCTAATGCCCGTACCCCCTGCGCTACCCAACCTTCTCCCTCT 2997
1110+3'f-1  ACATGTCGCAAGGTGAGCTCTAATGCCCGTACCCCCTGCGCTACCCAACCTTCTCCCTCT 3000
Lbr1110+3'f ACATGTCGCAAGGTGAGCTCTAATGCCCGTACCCCCTGCGCTACCCAACCTTCTCCCTCT 2996
            ************************************************************

1110+3'f-2  CCGCCGCCACGCAGAGAGAGAGAGAGAGAGAGCAGCTGTTTTACCCAAAAGAAAAAGAGC 3057
1110+3'f-1  CCGCCGCCACGCAGAGAGAGAGAGAGAGAGAGCAGCTGTTTTACCCAAAAGAAAAAGAGC 3060
Lbr1110+3'f CCGCCGCCACGCAGAGAGAGAGAGAGAGAGAGCAGCTGTTTTACCCAAAAGAAAAAGAGC 3056
            ************************************************************
```

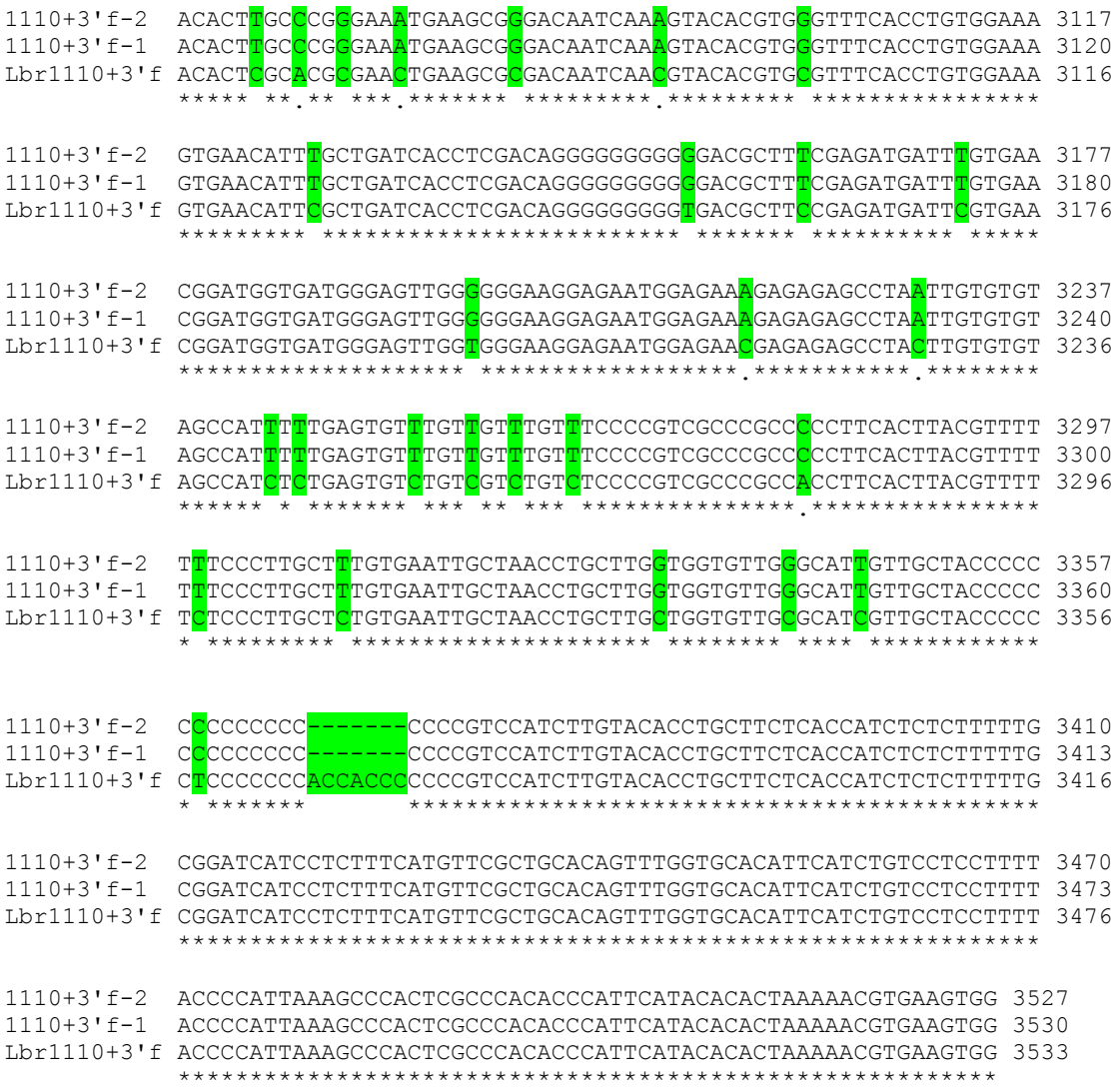


Fig.6.12 – Sequence alignment of down-stream LbrM.23.1110 gene copy with 3'flanking region of OHL

The most downstream copy of LbrM.23.1110 has a 3' flanking region distinct of the intergenic region, which separated the OHL from the downstream Lbr.23.1100 gene. Only two distinct fragment sequences were identified among several sequence clones. Here only a part of the sequenced fragment is shown (full sequence in Appendix 13). The first cut out shows the ORF of LbrM.23.1110 (red), where only one SNP (green) was identified. No SNPs were found in the UTRs. However, in second sequence cut out of the 3' flanking region of the OHL, many SNPs were identified compared to the available sequence on GeneDB (green).

```
1110+int-3 TGCTCTCTCTCCCCCACCTATAAGCAATGGGGACCGCCTGTATGAGGGAGTTGACGAGGC 720
1110+int-5 TGCTCTCTCTCCCCCACCTATAAGCAATGGGGACCGCCTGTATGAGGGAGTTGACGAGGC 720
1110+int-1 TGCTCTCTCTCCCCCACCTATAAGCAATGGGGACCGCCTGTATGAGGGAGTTGACGAGGC 720
1110+int-2 TGCTCTCTCTCCCCCACCTATAAGCAATGGGGTCCGCCTGTATGGGGGAGTTCACGAGGC 720
Lbr2904    TGCTCTCTCTCCCCCACCTATAAGCAATGGGGACCGCCTGTATGAGGGAGTTGACGAGGC 720
1110+int-4 TGCTCTCTCTCCCCCACCTATAAGCAATGGGGTCCGCCTGTATGGGGGAGTTGACGAGGC 720
           ******************************** *********** ******* *******

1110+int-3 CGCGCACGTTCGACCTTAAGGCTCACGGGATGGGCGGTGGCAAAGGGGATAGGGCGAACG 780
1110+int-5 CGCGCACGTTCGACCTTAAGGCTCACGGGATGGGCGGTGGCAAAGGGGATAGGGCGAACG 780
1110+int-1 CGCGCACGTTCGACCTTAAGGCTCACGGGATGGGCGGTGGCAAAGGGGATAGGGCGAACG 780
1110+int-2 CGCGCACGTTCGACCTTAAGGCTCACGGGATGGGCGGTGGCAAAGGGGATAGGGCGAACG 780
Lbr2904    CGCGCACGTTCGACCTTAAGGCTCACGGGATGGGCGGTGGCAAAGGGGATAGGGCGAACG 780
1110+int-4 CGCGCACGTTCGACCTTAAGGCTCACGGGATGGGCGGTGGCAAAGGGGATAGGGCGAACG 780
           ************************************************************

1110+int-3 GCGGCGAGCATGGACATGAACATATGAACGGCGGCGACCATGGACATGAACATATGGACG 840
1110+int-5 GCGGCGAGCATGGACATGAACATATGAACGGT---------------------------- 811
1110+int-1 GCGGCGAGCATGGACATGAACATATGGACGGTGGCGACCATGGACATGAACATATGAACG 840
1110+int-2 GCGGCGAGCATGGACATGAACATATGGACGGTGGCGACCATGGACATGAACATATGAACG 840
Lbr2904    GCGGCGAGCATGGACATGAACATATGAACGGCGGCGACCATGGACATGAACATATGGACG 840
1110+int-4 GCGGCGAGCATGGACATGAACATATGGACGGTGGCGACCATGGACATGAACATATGAACG 840
           ************************** ****

1110+int-3 GTGGCGACCATGGACATGAACATATGGACGGTGGCGCACCTAACGGGAATGGGAAGGATG 900
1110+int-5 --GGCGACCATGGACATGAACATATGGACGGTGGCGCACCTAACGGGAATGGGAAGGATG 870
1110+int-1 GTGGCGACCATGGACATGAACATATGGACGGTGGCGCACCTAACGGGAATGGGAAGGATG 900
1110+int-2 GTGGCGACCATGGACATGAACATATGGACGGTGGCGCACCTAACGGGAATGGGAAGGATG 900
Lbr2904    GTGGCGACCATGGACATGAACATATGGACGGTGGCGCACCTAACGGGAATGGGAAGGATG 900
1110+int-4 GTGGCGACCATGGACATGAACATATGGACGGTGGCGCACCTAACGGGAATGGGAAGGATG 900
             ************************************************************

1110+int-3 AAAATATGGGGAACGATAACGAGCATAATGGGATGGGTGATGATGCCAACCCCTGATGTG 960
1110+int-5 AAAATATGGGGAACGATAACGAGCATAATGGGATGGGTGATGATGCCAACCCCTGATGTG 930
1110+int-1 AAAATATGGGGAACGATAACGAGCATAATGGGATGGGTGATGATGCCAACCCCTGATGTG 960
1110+int-2 AAAATATGGGGAACGATAACGAACATAATGGGATAGGTGATGATGCCAACCCCTGATGTG 960
Lbr2904    AAAATATGGGGAACGATAACGAGCATAATGGGATGGGTGATGATGCCAACCCCTGATGTG 960
1110+int-4 AAAATATGGGGAACGATAACGAGCATAATGGGATGGGTGATGATGCCAACCCCTGATGTG 960
           ********************** *********** ************************

1110+int-3 CTGCGTGCCGGCTTGTGCTTGTGGGCCGAGCCCTTCGTCGGGCCTCTTGTGCCTCGTGCG 1020
1110+int-5 CTGCGTGCCGGCTTGTGCTTGTGGGCCGAGCCCTTCGTCGGGCCTCTTGTGCCTCGTGCG 990
1110+int-1 CTGCGTGCCGGCTTGTGCTTGTGGGCCGAGCCCTTCGTCGGGCCTCTTGTGCCTCGTGCG 1020
1110+int-2 CTGCGTGCCGGCTTGTGCTTGTGGGCCGAGCCCTTCGTCGGGCCTCTTGTGCCTCGTGCG 1020
Lbr2904    CTGCGTGCCGGCTTGTGCTTGTGGGCCGAGCCCTTCGTCGGGCCTCTTGTGCCTCGTGCG 1020
1110+int-4 CTGCATGCCGGCTTGTGCTTGTGGGCCGAGCCCTTCGTCGGGCCTCTTGTGCCTCGTGCG 1020
           **** ************************************************************

1110+int-3 CAGACTGCGTGTGTTGCTCGCGGTTCGTGTCTCTCCGCACACAGTGGCTAATGCCTGCCT 1080
1110+int-5 CAGACTGCGTGTGTTGCTCGCGGTTCGTGTCTCTCCGCACACAGTGGCTAATGCCTGCCT 1050
1110+int-1 CAGACTGCGTGTGTTGCTCGCGGTTCGTGTCTCTCCGCACACAGTGGCTAATGCCTGCCT 1080
1110+int-2 CAGACTGCGTGTGTTGCTCGCGGTTCGTGTCTCTCCGCACACAGTGGCTAATGCCTGCCT 1080
Lbr2904    CAGACTGCGTGTGTTGCTCGCGGTTCGTGTCTCTCCGCACACAGTGGCTAATGCCTGCCT 1080
1110+int-4 CAGACTGCGTGTGTTGCTCGCGGTTCGTGTCTCTCCGCACACAGTGGCTAATGCCTGCCT 1080
           ************************************************************
1110+int-3 GGGGTCGTTGTGATTGTACCTCATGGGCACCCCCGGCTTTCCCCGACTCGACTTCCCCTC 1140
1110+int-5 GGGGTCGTTGTGATTGTACCTCATGGGCACCCCCGGCTTTCCCCGACTCGACTTCCCCTC 1110
1110+int-1 GGGGTCGTTGTGATTGTACCTCATGGGCACCCCCGGCTTTCCCCGACTCGACTTCCCCTC 1140
1110+int-2 GGGGTCGTTGTGATTGTACCTCATGGGCACCCCCGGCTTTCCCCGACTCGACTTCCCCTC 1140
Lbr2904    GGGGTCGTTGTGATTGTACCTCATGGGCACCCCCGGCTTTCCCCGACTCGACTTCCCCTC 1140
1110+int-4 GGGGTCGTTGTGATTGTACCTCATGGGCACCCCCGGCTTTCCCCGACTCGACTTCCCCTC 1140
           ************************************************************

1110+int-3 TCCGCTTCCGAGTGTGTGGTGGGTGTGGGGTTGGGCGCATAATGGACATTGTCGGTGGAT 1200
1110+int-5 TCCGCTTCCGAGTGTGTGGTGGGTGTGGGGTTGGGCGCATAATGGACATTGTCGGTGGAT 1170
1110+int-1 TCCGCTTCCGAGTGTGTGGTGGGTGTGGGGTTGGGCGCATAATGGACATTGTCGGTGGAT 1200
1110+int-2 TCCGCTTCCGAGTGTGTGGTGGGTGTGGGGTTGGGCGCATAATGGACATTGTCGGTGGAT 1200
Lbr2904    TCCGCTTCCGAGTGTGTGGTGGGTGTGGGGTTGGGCGCATAATGGACATTGTCGGTGGAT 1200
1110+int-4 TCCGCTTCCGAGTGTGTGGTGGGTGTGGGGTTGGGCGCATAATGGACATTGTCGGTGGAT 1200
           ************************************************************
```

```
1110+int-3 GCGCGACGACTGCCGCTACACAACTGTGGCCACGCCGAGTCCTGTGTGTGTATGCTTAGA 1260
1110+int-5 GCGCGACGACTGCCGCTACACAACTGTGGCCACGCCGAGTCCTGTGTGTGTATGCTTAGA 1230
1110+int-1 GCGCGACGACTGCCGCTACACAACTGTGGCCACGCCGAGTCCTGTGTGTGTATGCTTAGA 1260
1110+int-2 GCGCGACGACTGCCGCTACACAACTGTGGCCACGCCGAGTCCTGTGTGTGTATGCTTAGA 1260
Lbr2904    GCGCGACGACTGCCGCTACACAACTGTGGCCACGCCGAGTCCTGTGTGTGTATGCTTAGA 1260
1110+int-4 GCGCGACGACTGCCGCTACACAACTGTGGCCACGCCGAGTCCTGTGTGTGTATGCTTAGA 1260
           ************************************************************

1110+int-3 TCACCGGTGCTAGCAGCTCTCTTGCTGCCTCTGCGTGGACTCTCTTTGTTTCTTGTGTCT 1320
1110+int-5 TCACCGGTGCTAGCAGCTCTCTTGCTGCCTCTGCGCGGACTCTCTTTGTTTCTTGTGTCT 1290
1110+int-1 TCACCGGTGCTAGCAGCTCTCTTGCTGCCTCTGCGCGGACTCTCTTTGTTTCTTGTGTCT 1320
1110+int-2 TCACCGGTGCTAGCAGCTCTCTTGCTGCCTCTGCGCGGACTCTCTTTGTTTCTTGTGTCT 1320
Lbr2904    TCACCGGTGCTAGCAGCTCTCTTGCTGCCTCTGCGCGGACTCTCTTTGTTTCTTGTGTCT 1320
1110+int-4 TCACCGGTGCCAGCAGCTCTCTTGCTGCCTCTGCGCGGACTCTCTTTGTTTCTTGTGTCT 1320
           ********** ********************** ************************

1110+int-3 TTTCTGTTCGGGTACCTCTACGCATGCGCATACCCCCTCCCTGCCTCTTCCCCTCTTTAC 1380
1110+int-5 TTTCTGTTCGGGTACCTCTACGCATGCGCATACCCCCTCCCTGCCTCTTCCCCTCTTTAC 1350
1110+int-1 TTTCTGTTCGGGTACCTCTACGCATGCGCATACCCCCTCCCTGCCTCTTCCCCTCTTTAC 1380
1110+int-2 TTTCTGTTCGGGTACCTCTACGCATGCGCATACCCCCTCCCTGCCTCTTCCCCTCTTTAC 1380
Lbr2904    TTTCTGTTCGGGTACCTCTACGCATGCGCATACCCCCTCCCTGCCTCTTCCCCTCTTTAC 1380
1110+int-4 TTTCTGTTCGGGTACCTCTACGCATGCGCATACCCCCTCCCTGCCTCTTCCCCTCTTTAC 1380
           ************************************************************

1110+int-3 CTCAGTGCGTCACACAGTGAGCTCCCTTCCCTCGACCTTATTTTGCTGCCTCTGGCGCCT 1440
1110+int-5 CTCAGTGCGTCACACAGTGAGCTCCCTTCCCTCGACCTTATTTTGCTGCCTCTGGCGCCT 1410
1110+int-1 CTCAGTGCGTCACACAGTGAGCTCCCTTCCCTCGACCTTATTTTGCTGCCTCTGGCGCCT 1440
1110+int-2 CTCAGTGCGTCACACAGTGAGCTCCCTTCCCTCGACCTTATTTTGCTGCCTCTGGCGCCT 1440
Lbr2904    CTCAGTGCGTCACACAGTGAGCTCCCTTCCCTCGACCTTATTTTGCTGCCTCTGGCGCCT 1440
1110+int-4 CTCAGTGCGTCACACAGTGAGCTCCCTTCCCTCGACCTTATTTTGCTGCCTCTGGCGCCT 1440
           ************************************************************
```

Fig.6.13 – Multiple copies and polymorphisms in LbrM.23.1110

A part of the sequences of the fragments amplified by divergent primers from the intergenic region, which contain SNPs, are shown (full sequence in Appendix 14). A set of fragments containing a single copy of LbrM.23.1110 flanked up- and downstream by the same intergenic region were generated with the divergent primers unexpectedly (Fig.6.7). Seven SNPs (green) were identified within the ORF (marked in red for the *Lbr*2904) and three in the 3' UTR of the gene. In one case (1110+int-5) a sequence repeat was deleted from the central region (blue).

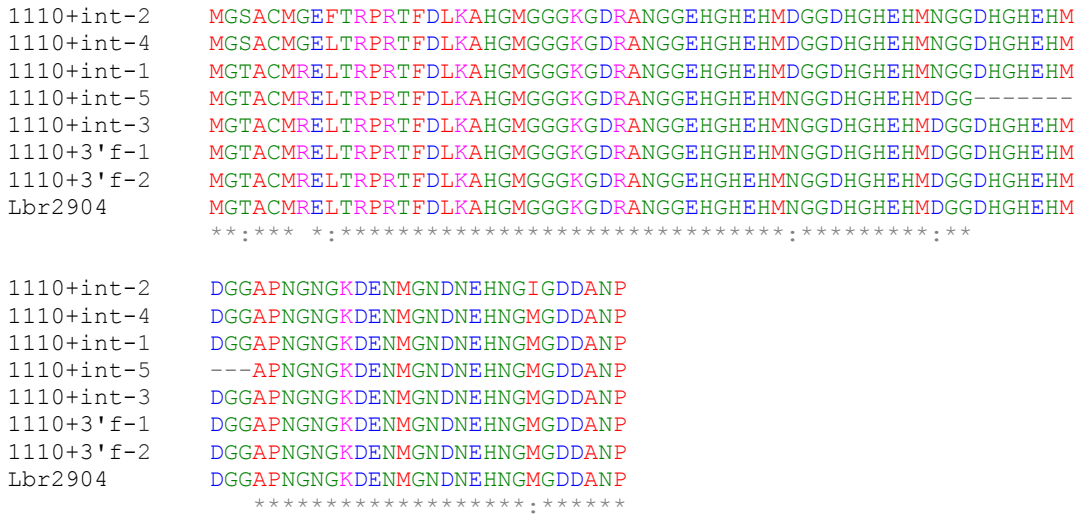


Fig.6.14 – Alignment of distinct 1110 amino acid sequences

The DNA sequence obtained by subcloning and sequencing of LbrM.23.1110 were translated and aligned. SNPs in the ORF location 7, 19, 27, 121, 126 and 151 had been previously identified by comparing sequences from *L. (V.) braziliensis* 2904 and 2903 for this ORF and are available on TriTryp. Five of six the non-synonymous amino acid changes occurred within groups of strongly similar properties and only one was random (position 7 in the amino acid sequence).

Table 6.2 – SNPs with the LbrM.23.1110 and Lbr.23.1120 ORFs

Gene in OHL	Position in ORF	Base Change	Non-/Syn.	Position in AA seq.	AA Change
LbrM.23.1110	7	A→T	Non-syn.	3	T→S
LbrM.23.1110	19	A→G	Non-syn.	7	R→G
LbrM.23.1110	27	G→C	Non-syn.	9	L→F
LbrM.23.1110	121	A→G	Non-syn.	41	N→D
LbrM.23.1110	126	C→T	Syn.	42	
LbrM.23.1110	147	T→C	Syn.	49	
LbrM.23.1110	151	G→A	Non-syn.	51	D→N
LbrM.23.1110	237	G→A	Syn.	79	
LbrM.23.1110	249	G→A	Non-syn.	83	M→I

function. In one case, the deletion of an entire sequence repeat (30 nt or 10 aa) within the ORF was observed, too, which changed the ORF size (Fig.6.13). However, all sequenced clones of LbrM.23.1110 flanked up- and downstream with the intergenic region were of very similar size and did not account for all bands detected by OHL I on the southern blot (Fig.6.15).

In summary, the OHL shows intra-species size variation. Two distinct heterozygous copies of LbrM.23.1120 and a single polymorphism in LbrM.23.1110 were found in the *L. (V.) braziliensis* 2904 gDNA. In addition, it was shown that multiple LbrM.23.1110 copies occur within the OHL flanked by the intergenic region as a tandem repeat array of 3-4 copies. It was also found that the intergenic region is unusually well preserved, which suggests the presence of a yet to be identified SHERP homologue within the intergenic region. Four hypothetical ORFs were found within the intergenic region, but no homologues were found for the hypothetical proteins. However, since SHERP is a unique protein to *L. (Leishmania) spp.*, it could be that this hypothetical gene within the OHL intergenic region is unique to *L. (Viannia) spp*.

6.4. Generating *L. (V.) braziliensis* OHL double deletion mutants

Rogers *et al.* (2011) had suggested that *L. (V.) braziliensis* 2409 was generally triploid. For phenotypic comparison between *L. (L.) major* cDNA16 locus and *L. (V.) braziliensis* OHL fill deletion mutants, it was, therefore, necessary to generate an *L. (V.) braziliensis* OHL triple deletion mutant by sequential homologous recombination. Three OHL deletion constructs were generated by the same method employed for the HASP and SHERP replacement constructs (see 3.2) using the OHL 5' and 3' flanking regions available on TriTrypDB and a cassette containing the antibiotic resistance gene (SAT, NEO or BSD) flanked by DHFR flanking regions (Fig.6.16). The minimal concentrations of SAT, NEO and BSD required for killing the *L. (V.) braziliensis* parental strain (*Lbr*2904) were determined by *Lbr*2904 growth in M199 with different antibiotic concentration.

The first OHL allele deleted was performed with a BSD-deletion construct. Initially positive/negative screening was done by PCR (Fig.6.15). Using a forward primer upstream of the integration within *Lbr*2904 genomic DNA and a reverse primer within the construct's 5' DHFR flanking region, a ~2.1 Kb band was expected to confirm construct integration. A Southern blot of *Sac*I digested gDNA samples of all PCR positive single OHL deletion mutants was used to determine correct integration of a single BSD-deletion construct copy (Fig.6.17). Out of 12 tested

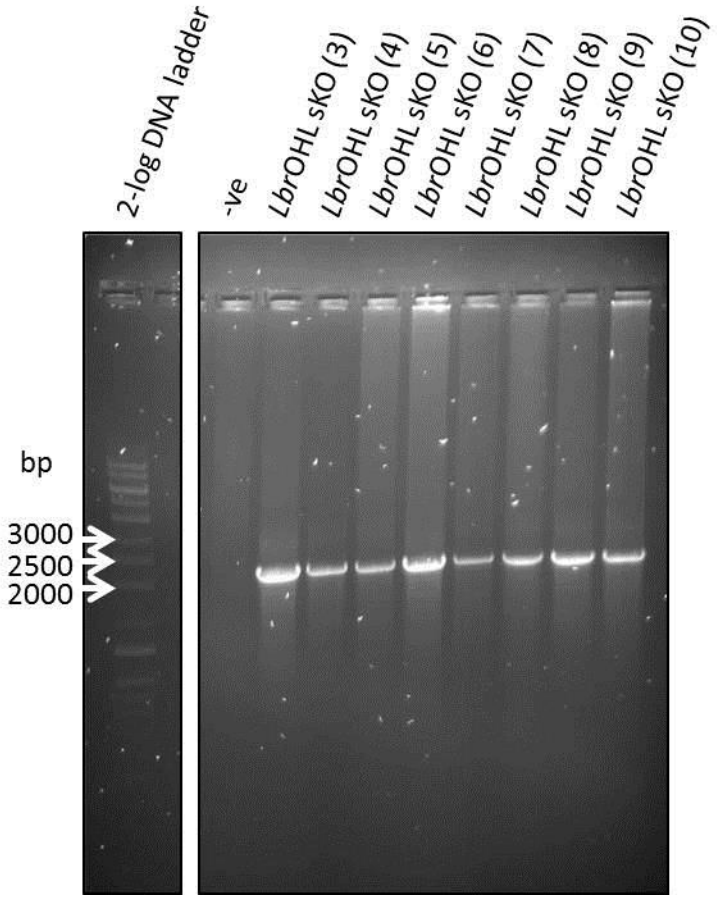


Fig.6.15 – PCR screen for verification of OHL single deletion

The gel images shows examples of *Lbr*OHL sKO clones, which checked positive. A ~2 Kb band was expected according to sequence map.

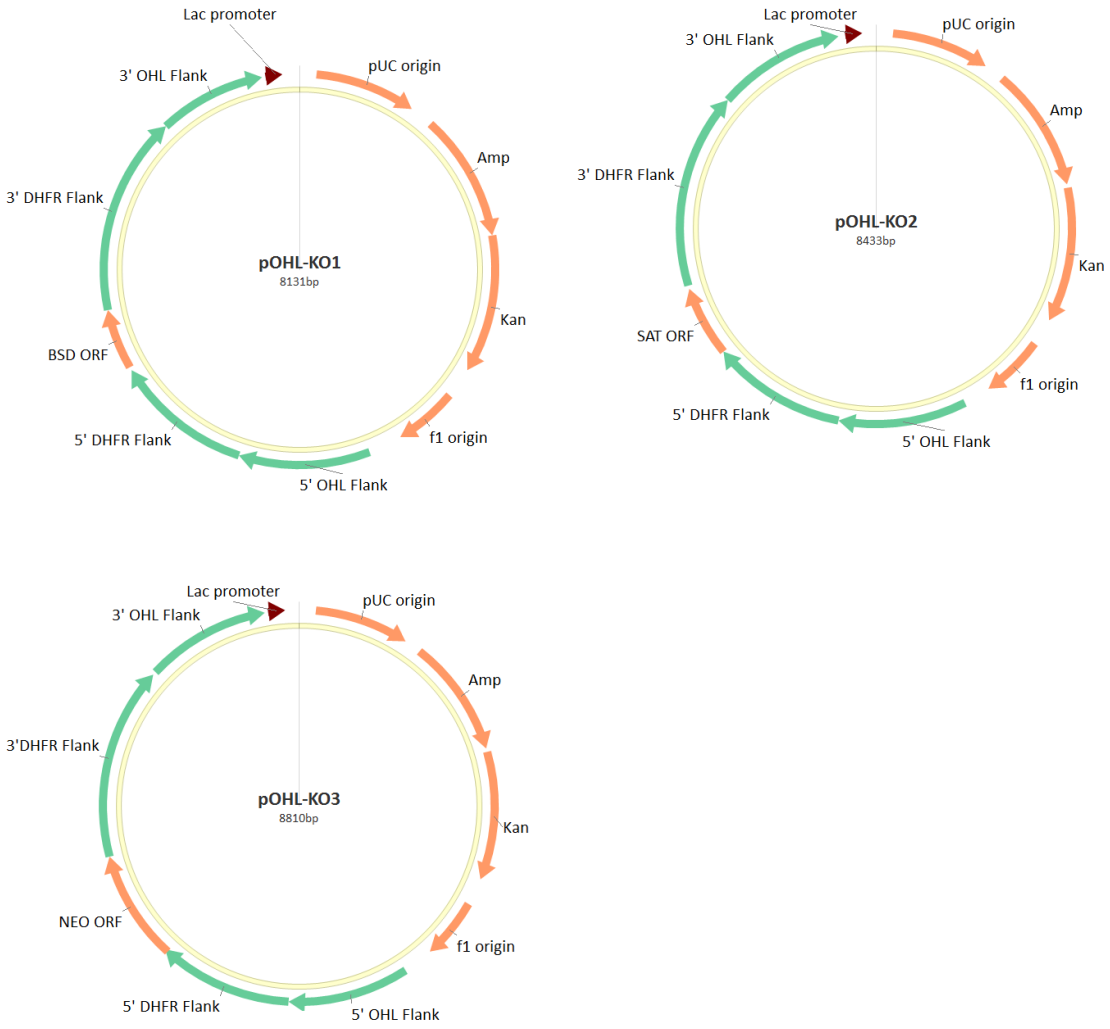


Fig.6.16 – Orthologous HASP locus deletion constructs in plasmid vectors

All vectors were generated within the pCR®2.1-TOPO® with identical 5' and 3' OHL flanks based on the available OHL sequence on GeneDB and TriTrypDB for homologous recombination. OHL-KO1 contains a blasticidin (BSD) resistance gene, OHL-KO2 a streptophricin (SAT) resistance gene and OHL-KO3 a neomycin (NEO) resistance gene.

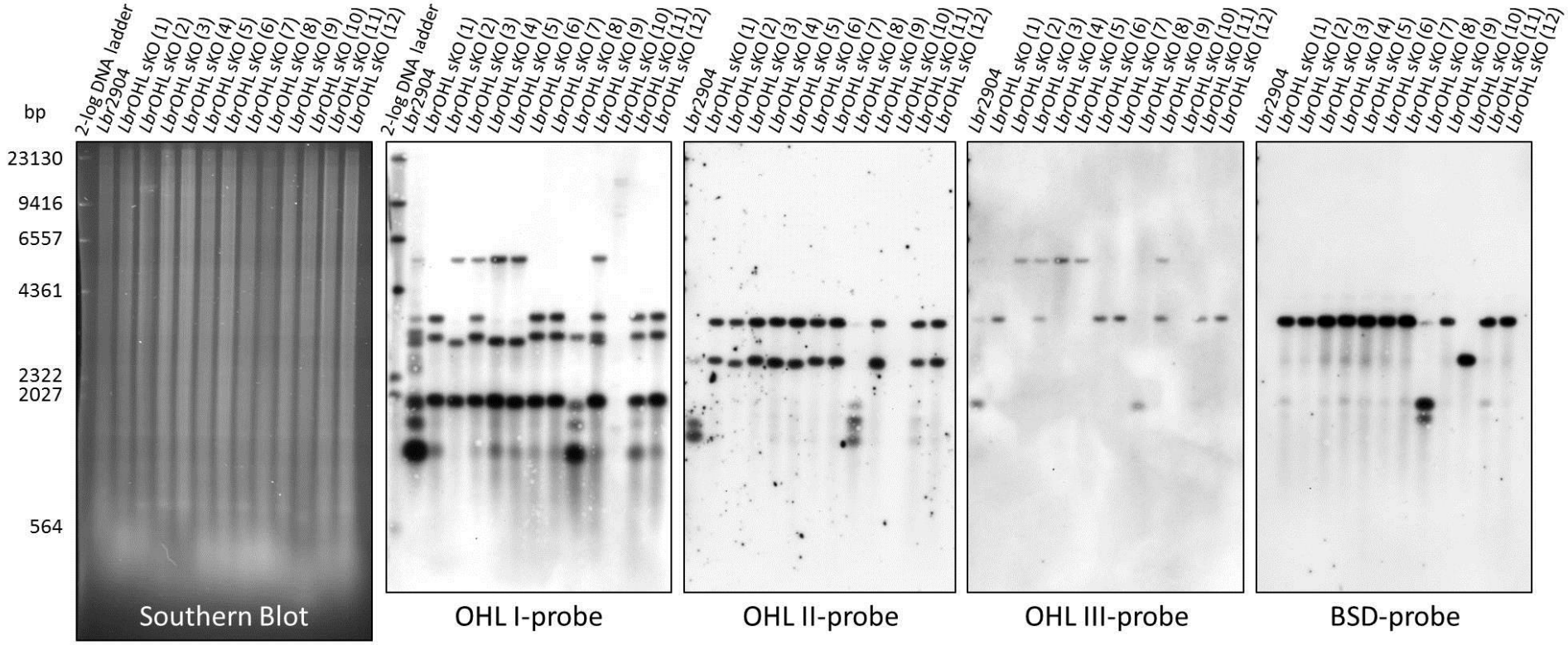


Fig.6.17 – Southern blot to verify successful single OHL deletion

The Southern blots showed that *Lbr*OHL sKO clones 8 and 10 did not have the expected ~3.95 Kb band when probed with the BSD DIG-labelled probe and were, therefore, excluded from further analysis. *Lbr*OHL sKO clones 3 and 9 were questionable, because the OHL III probe found two instead of only one expected band. Probing with the OHL I probe showed two distinct band patterns for *Lbr*OHL sKO clones 1, 6, 7, 11, & 12 and 2, 4 & 5, respectively. This suggested that both alleles are heterozygous and were both targeted with the deletion construct.

clones, two (*Lbr*OHL sKO 8 and 10) were shown to have incorrect BSD construct integration and in two others (*Lbr*OHL sKO 3 and 9), the OHL III probe detected two bands, although only one was expect (Fig.6.17). The remaining 8 clones showed correct deletion construct integration. 10 clones were tested in a qPCR for locus copy number (Fig.6.18). The single copy gene LbrM.23.1040 served as a control gene and the coefficient of average OHL abundance was calculated by dividing the mean quantities of the OHL with those from the control gene. Although the results showed some variation, 5 clones (*Lbr*OHL sKO 1, 4, 6, 7 & 11) had values close to 1 for one OHL copy compared to *Lbr*2904, which had a value close to 2. This indicated that one OHL copy had been successfully deleted in these 5 clones. *Lbr*OHL sKO 7 was picked for a second round of OHL deletion.

The generation of the OHL double deletion mutant was attempted 4 times targeting the second allele of this locus with two different deletion constructs (one containing SAT, the other NEO) 2 times each, but no colonies were obtained on the antibiotic agar plates, although colonies grew on antibiotic free control plates. This may suggest that the OHL in *L. (V.) braziliensis* is diploid against expectations and – unlike the cDNA16 locus in *L. (Leishmania) spp.* – is essential for parasite survival. Conversely, it could mean that the antibiotic concentration was still too high, although the previously determined minimal antibiotic concentration for *Lbr*2904 killing had been used. Potentially, the constructs for OHL deletion were unfit to delete the second locus copy due to yet unrecognized heterogeneity in the OHL flanking regions. This, however, seems unlikely considering the Southern blot results from the OHL single deletion mutant. When probing with the OHL I probe in the intergenic region between 1110 and 1120 three distinct band patterns were observed (one pattern for clones *Lbr*OHL sKO 1, 6, 7, 11, & 12; the second for clones *Lbr*OHL sKO 2, 4 & 5 and the third for *Lbr*OHL sKO 3 & 9), which suggested that all alleles had been targeted in the first deletion round with the flanking regions used. These data suggested that the OHL might be essential for viability in *L. (V.) braziliensis*.

To determine whether OHL is indeed essential for *L. (V.) braziliensis* viability, more work was required. For example, the full locus could be integrated into one of the ribosomal SSU loci prior to another deletion attempt of the second OHL copy. In this situation, if integrated colonies with both original OHL copies deleted grew, that would suggest that the OHL is essential for *L. (V.) braziliensis* survival and would mark an important difference to the cDNA16 locus. To test whether the SAT and NEO-containing constructs were fit for gene deletion, *Lbr*2904 could be

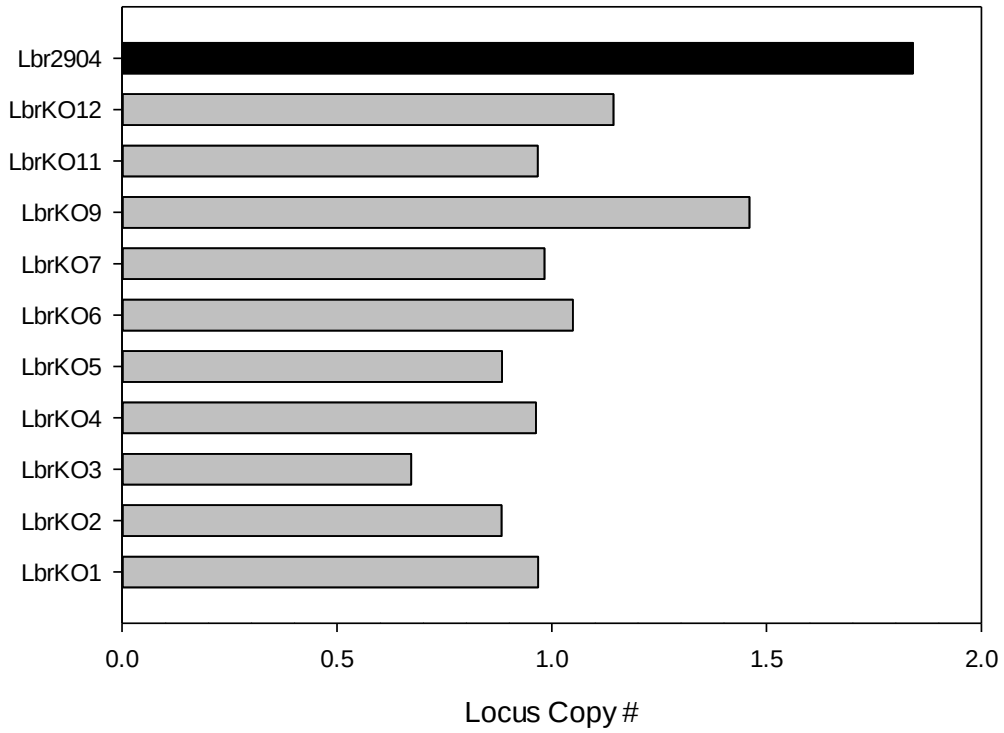


Fig.6.18 – qPCR verification of OHL copy number in single deletion mutants

A qPCR was used to determine OHL copy number within the *Lbr*OHL sKO clones. A coefficient of 1 was expected compared to two for the parental line *Lbr*2904. Clones 1, 4, 6, 7 and 11 had the closest coefficients to 1.

subjected to first round OHL deletion and any clones generated analysed as above. Further in-detail analysis of the OHL DNA sequences may also reveal further heterogeneity between the alleles, potentially explaining the failure to generate a full OHL deletion mutant with the available constructs.

Unfortunately, it was not possible to investigate the OHL any further during this project due to time constraints.

6.5. Conclusions

Regarding the map of the *L. (V.) braziliensis* orthologous HASP locus, the data generated in this study have confirmed that the currently available sequence of the OHL on GeneDB and TriTrypDB is incorrect and suffers from repeat collapses as previously suggested by Depledge *et al.* (2010) (*294*). Considering that the fragment containing the LbrM.23.1110 gene could be amplified with divergent primers from the intergenic region, this suggested that there are repeat copies of LbrM.23.1110 present within the OHL. These copies were distinct from one another by SNPs occurring in their ORFs, although most of the SNPs were either synonymous or only cause an amino acid change within groups of highly similar properties. This suggested that the 1110 proteins have the same function. Considering that seven of the nine identified SNPs in LbrM.23.1110 were previously identified comparing the sequences of two different *L. (V.) braziliensis* strains (2904 and 2903), this would suggest that these SNPs are not strain specific, but conserved in the multiply repeated LbrM.23.1110 gene between *L. (V.) braziliensis* strains. To prove this, higher resolution of the OHL map would be required. It was not possible to determine how many LbrM.23.1110 copies were present in the analysed OHL of *Lbr*2904, but considering a 4 – 5-fold increase in reading depth for this sequence would suggest 4 – 5 LbrM.23.1110 copies per OHL. One of these is the most downstream copy of LbrM.23.1110, which is flanked downstream by the OHL flanking sequence. Only two versions of this copy, distinct by a single SNP, have been identified in this study.

The identification of two distinct versions of LbrM.23.1120 (1120_v1 and 1120_v2), both flanked by the same upstream flanking region, which is distinct from the intergenic region, suggested that LbrM.23.1120 might be heterozygous. Considering that only a 2-fold increase in reading depth was observed for the 1120 ORF, this suggested that there was only one copy of LbrM.23.1120 per allele. Together with the data generated for LbrM.23.1110, I propose that a single heterozygous LbrM.23.1120 copy and 4 – 5 tandemly repeated copies of

LbrM.23.1110 are present within the OHL (Fig.6.19). The band patterns observed on the Southern blots of *Sac*I digested *Lbr*OHL sKO clones probed with the OHL I probe suggested that the OHL might be heterozygous (Fig.6.17). In this study, I attempted to produce a full length PCR product by long range high fidelity PCR with primers designed on the available map. The PCR cycle was calculated based on the fragment sizes for the OHL identified by restriction digest (Fig.6.3; Table 6.1), but so far it has not been possible to generate a fragment of this size. This could mean that the known sequences of 5' and 3' flanking regions of the OHL are also heterozygous. Further investigation is required to resolve the OHL map in *L. (V.) braziliensis*.

The high level of sequence conservation found within the intergenic region was surprising. Normally, non-functional intergenic regions are more prone to mutations than the flanking genes, but in the OHL, the inverse was observed. This could mean that a yet to be identified orthologue of SHERP may be present within that region. Based on the currently available sequence on GeneDB, verified by subcloning and sequencing, the intergenic region measures ~1 Kb and a potential 402 bp ORF is present. Considering that the SHERP ORF only measures 174 bp, it is possible that another gene is present in the intergenic region. Further analysis would be required to show if the ORF corresponded to a functional mRNA, translated into a functional protein and had any structural, biochemical and functional similarities to SHERP.

The attempts to generate an OHL double deletion mutant have failed so far. This could mean that the OHL is essential in *L. (V.) braziliensis*, which would be an important difference to the cDNA16 locus of *L. (Leishmania) spp.*, but the functionality of the three SAT-, BSD- and NEO-deletion constructs still needs to be proven in a single OHL deletion mutant generation. Further investigation is required to address the essentiality of the OHL in *L. (V.) braziliensis*.

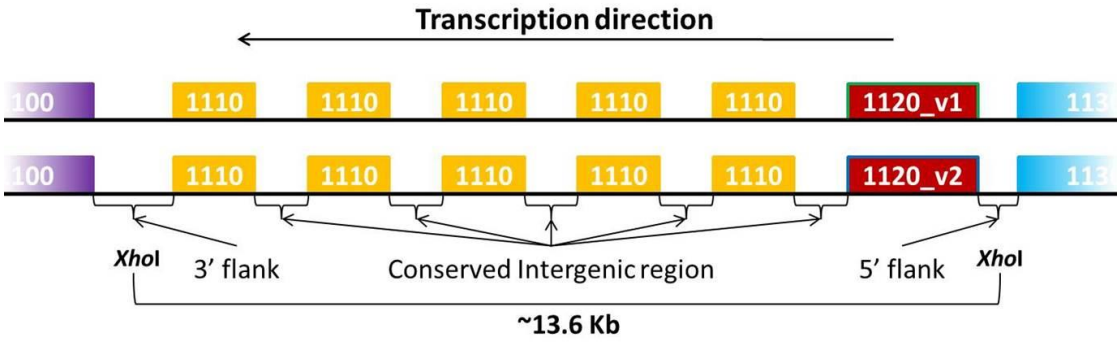


Fig.6.19 – Proposed *L. (V.) braziliensis* 2904 OHL map

The proposed map shows heterozygous versions of the OHL indicated by the two distinct LbrM.23.1120 copies (1120_v1 & 1120_v2; red). In this case five copies of LbrM.23.1110 are proposed (four flanked on both sides by the conserved intergenic region and one flanked downstream by the 3' flanking region of the OHL). The proposed map measures ~13.5 Kb from the *Xho*I restriction sites in the OHL flanking regions based on the accumulative lengths of the individual fragments, which is close to the estimated ~13.6 Kb for the fragment generated by *Xho*I digest (Fig.6.4). The individual LbrM.23.1110 genes are distinct from one another as indicated by specific SNP patterns in the ORFs. Two suggested sequences for the heterozygous OHL can be found in Appendices 15 & 16.

7. Chapter VII. – Discussion

7.1. Part One: Metacyclogenesis of *L. (L.) major* HASP and SHERP mutants in the sand fly vector

7.1.1. Data summary

Chapter III described the generation and analysis of newly generated HASP and SHERP replacement mutant lines by homologous recombination into the original cDNA16 locus based on previously established protocols (*146*, *332*). All mutant lines were grown in M199 culture and checked by PCR, Southern blot and qPCR for correct integration of a single HASP and/or SHERP construct into the cDNA16 locus in the null background of the *Lmj*cDNA16 dKO mutant line (Fig.3.5; Fig.3.7 & Fig.3.8, respectively), previously generated and characterized by McKean *et al.* (2001). Western blots confirmed gene construct expression and regulation at parental line (FVI) levels (Fig.3.9 & 3.10), while a biotinylation assay confirmed HASPB surface localization (Fig.3.11). Growth assays in M199 culture identified no fitness disadvantageous in the newly generated mutant lines (Fig.3.12). Based on this thorough mutant characterization two clones were picked per mutant line and passaged through BALB/c mice to recover parasite virulence based on previously established protocols (*293*). Mutant lines containing HASPA2, but not HASPA1, on a gene construct were observed to develop lesions at a much slower rate than all other passaged mutant lines. They also generated promastigotes on amastigote inoculation into M199 much slower than all other mutant lines. This observation is currently under investigation and a complete data set is not yet available.

Results in Chapter IV showed that the *in vitro* characterized mutant lines behaved differently in the sand fly midgut than expected from *in vitro* observations. The data showed that step-by-step replacement of HASP and SHERP genes was insufficient to rescue metacyclogenesis completion. Metacyclics were generated in very low numbers in all mutant lines except for *Lmj*cDNA16 sKI, which was the only mutant line to rescue metacyclogenesis to completion. In single gene replacement mutant lines, metacyclics derived from sand fly midguts had a cell body morphologically distinct from parental line (FVI) metacyclics. Reversion to the parental line metacyclic cell body morphology was achieved by the replacement of several HASP and SHERP genes. However, it is not clear which HASP and/or SHERP genes were

required to rescue parental line metacyclic cell body morphology. The discovery of few metacyclics with parental line morphology in *Lmj*S2+HB sKI and *Lmj*S2/HB sKI suggested that a combination of HASPB and SHERP genes may be sufficient to rescue metacyclic morphology (Fig.4.11B), even though not the metacyclic numbers, although HASPB was not upregulated in these mutant lines. The generation of metacyclics and presumably also haptomonads, considering the lack of SV colonization in all mutant lines except *Lmj*cDNA16 sKI (Fig.4.1), remained inefficient in all mutant lines. Leptomonad generation, however, could be improved significantly by replacement of combinations of HASP and SHERP genes. Unfortunately, the data did not clearly show which HASP and/or SHERP genes were the most essential for leptomonad generation, although replacement of HASPA1 and HASPA2 in a single construct appeared beneficial for generation of this stage. It is possible that improved leptomonad generation was related to the sand fly species infected since the vector species had to be changed from *Ph. (Ph.) papatasi* to *Ph. (Ph.) duboscqi* during the study for technical reasons. Since *Ph. (Ph.) duboscqi* is less specific for the infecting *L. (L.) major* strain than *Ph. (Ph.) papatasi* the improved leptomonad generation observed in the later experiments with *Ph. (Ph.) duboscqi* could be an artefact related to the sand fly species.

In Chapter V, it was revealed by confocal microscopy that HASPB was not up-regulated / expressed at detectable levels in all the tested replacement mutants (Fig5.1 & Fig.5.3). This was in contrast to the observations made by Western blot in culture derived parasites. Attempts to determine the reason for the difference in gene expression from the replacement construct within the cDNA16 locus were not able to explain this phenomenon conclusively. Neither difference in culture conditions (Fig.5.5), nor the addition of midgut extracts to M199 cultures was able to influence the expression of HASPB from the construct (Fig.5.6). Interestingly, profiles of HASPB mRNA levels from midgut and culture derived parasites revealed that, while the profile and relative levels of HASPB mRNA in *Lmj*HASPB sKI and *Lmj*cDNA16 sKI were comparable to the parental line (FVI) in culture, *in vivo* *Lmj*HASPB sKI showed considerably lower levels of HASPB mRNA than FVI and *Lmj*cDNA16 sKI at day 6 PBM, when HASPB mRNA was upregulated in FVI and *Lmj*cDNA16 sKI (Fig.5.7). Also the mRNA profiles for SHERP and HASPA were distinct for the *Lmj*SHERP sKI and *Lmj*HASPA2 sKI, respectively, compared to FVI. Interestingly, *Lmj*cDNA16 sKI expression patterns for

SHERP and HASPA both *in vitro* and *in vivo* were similar to the mutant lines rather than the parental line (FVI), suggesting that HASPB regulation may be the key difference between *Lmj*cDNA16 sKI and all the mutant lines and the reason for rescue of metacyclogenesis completion in *Lmj*cDNA16 sKI. The osmotaxis of tested mutant lines was not compromised (Fig.5.10), while the secretion of fPPG for the PSG gel plug generation was abolished in all mutant lines except *Lmj*cDNA16 sKI (Fig.5.11 & Fig.5.12). The fPPG synthesis pathway has not been described in detail in *Leishmania* yet. Synthesis and secretion of glycan-modified proteins and lipids usually involves secretion pathways through the ER and Golgi in other eukaryotic cells (*376*). Since *Leishmania* HASP and SHERP genes have no known orthologues in other eukaryotic systems and neither protein ever enters the ER or Golgi, it is difficult to predict how the HASP and/or SHERP proteins may influence fPPG synthesis and secretion. It has been hypothesised that HASPB may be shed in the midgut lumen from the *Leishmania* cell surface and that it may bind the fPPG supporting the PSG formation. However, it has not been possible to verify this hypothesis, because of the comparatively low parasite numbers in midguts, which causes HASPB levels to be too low for clear detection (Fig.7.1).

7.1.2. The HASPs and SHERPs in metacyclogenesis

The cDNA16 locus on chromosome 23 has been shown to be specific to members of the *L. (Leishmania)* subgenus (*294*). It contains genes of two unusual and unrelated gene families coding for the HASPs and SHERPs. These are stage specifically upregulated genes and with the exception of HASPA1, which is shown in this study to be amastigote-specific, they are preferentially expressed in promastigotes (*329*), although HASPB continues to be expressed in amastigotes weeks after mammalian host infection. *In vivo* HASPB and SHERP are expressed predominantly in metacyclics in *L. (L.) major*, although SHERP is detected at low levels in late leptomonads (*146*).

This thesis carries on the work on the HASP and SHERP genes of the *L. (L.) major* cDNA16 locus published by Sádlová *et al.* (2010), who demonstrated the essentiality of the cDNA16 locus in the sand fly midgut for metacyclogenesis completion (*146*). The aim of this thesis was to determine whether one or a subset of the HASP and /or SHERP genes were sufficient for metacyclogenesis completion, or if the whole locus was required. Sand fly infection studies by Sádlová *et al.* (2010), carried out with episomal HASPB

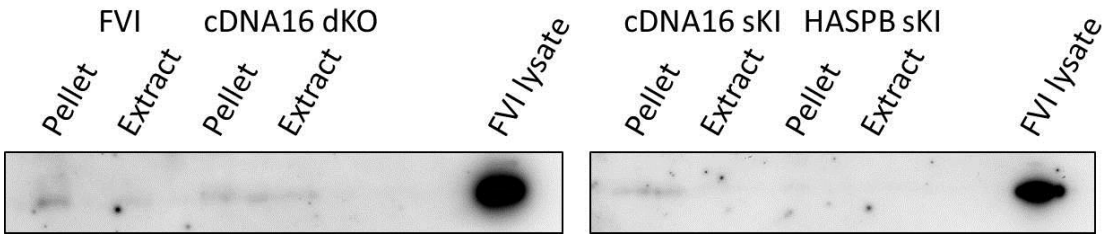


Fig.7.1 – HASPB detection in midgut extracts

All samples were prepared in 25 µl PBS + 25 µl Laemmli buffer. 15 µl of lysate from the parasite/midgut debris cell pellets from the extraction of midgut extracts from 20 sand fly midguts and 15 µl of midgut extracts from 20 midguts were loaded onto a 12% SDS-PAGE gel for a Western blot. Only 2 µl of a parasite cell lysate (8x 10^5 parasite cells) from culture were loaded as an antibody control. The blots show that HASPB concentrations are extremely low in the parasite cell pellet preparations from sand fly midguts. Only the FVI cell pellet gave a discernible band, which, however, is barely stronger than the background signal shown in the lanes of *Lmj*cDNA16 dKO. This is attributed to the low parasite numbers in the sand fly midgut ($5x10^3$ – $2x10^4$ per midgut) compared to a M199 culture. The 2 µl FVI cell lysate contains at least an estimated 6.5-25.5 times more parasite cells than the 15 µl of the parasite cell pellets. That would require a minimum of 130 heavily infected midguts to get a comparable signal to the 2 µl of FVI cell lysate.

and SHERP replacement mutant lines, suggested that HASPB may be essential for metacyclogenesis. However, gene expression from episomes is unregulated and often overexpressed. It was shown that both of these characteristics can cause overexpression phenotypes, which can vary significantly from the wild type phenotype (*332*). It had been shown that replacement by homologous recombination into the cDNA16 locus could recover wild type gene regulation (*146*, *332*). Therefore, the approach for mutant generation was the replacement of HASP and/or SHERP genes into the cDNA16 locus in the null background of *Lmj*cDNA16 dKO by homologous recombination, since individual gene deletion from the locus was not applicable due to the high level of sequence repetitiveness in the cDNA16 locus. All selected mutant lines had been rigorously tested in culture to verify genomic integration and re-establishment of wild-type gene expression regulation.

HASPB is a *N*-terminally dual-acylated protein targeted for the cell surface via membrane shedding (*335*, *336*, *345*). The previously demonstrated surface localization of HASPB was verified in a subset of replacement mutant lines too. The current hypothesis suggests that HASPB may be shed from the cell surface and is a target for B-1 B cell-derived natural antibodies (*377*). Whether HASPB is shed in the sand fly vector was addressed in this study, but it could not be proven that HASPB is present in midgut extracts (Fig.7.1). The suggested key function of HASPB in metacyclogenesis could also not be confirmed in this study, although it is suggested by the data generated. However, this is related to the lack of HASPB expression from the integrated construct, when the mutant lines were infected into the sand fly midgut.

Using mutant lines with individual HASPA1 and HASPA2 genes replaced into the cDNA16 locus, it was possible to show for the first time that HASPA2 is expressed in the promastigote stage, while HASPA1 expression is amastigote-specific, suggesting life-cycle specific functions for HASPA2 and HASPA1, while alternating their expression between promastigote and amastigote stage, respectively. Although HASPA1 and HASPA2 have the same ORF and are only distinct in their 3' UTRs, differences were observed in the behaviour of HASPA1 or HASPA2 containing mutant lines in the BALB/c mouse model. HASPA2 containing mutant lines without a HASPA1 gene copy developed lesions much more slowly and were much slower to produce promastigotes after amastigote inoculation into M199 medium than mutant

lines lacking HASPA2, including the cDNA16 null mutant line, *Lmj*cDNA16 dKO. Interestingly, introducing a HASPA1 copy into HASPA2 containing mutant lines abolished this phenotype. This may suggest a specific function for the HASPA1 in amastigotes that complements the HASPA2 function in promastigotes. Mutant lines containing only HASPA2 produced significantly more leptomonads than those containing only HASPA1 ($P<0.001$; Fig.4.8), while mutant lines containing both HASPA genes were among the most efficient leptomonad generators and formed lesions at a similar rate as FVI (detailed analysis still underway).

SHERP is a peripheral membrane protein that localizes to the cytosolic face of the ER and mitochondrion (*341*). It has previously been shown that interaction with anionic phospholipids is essential for SHERP to assume a globular form, which may be important for SHERP function (*348*). SHERP is proposed to form a complex with the subunit B of the vacuole H^+-ATP synthase (V-ATPase) potentially promoting V-ATPase assembly or preventing its disassembly (*348*). V-ATPases are important in the acidification of internal compartments like the endosomal/lysosomal system (*378*). It has been hypothesised that SHERP interaction with the V-ATPase may be important for parasite autophagy, based on SHERP's localization to the cytosolic face of the ER and mitochondrial membranes. These membranes are a source of phagophores that form the autophagosomes of the autophagic system (*350*). Autophagy was found to be important for parasite differentiation into metacyclics (*298*, *349*), which could make SHERP a key regulator for metacyclogenesis. However, data from this study did not confirm this hypothesis.

7.1.3. Differences in mutant parasite behaviour *in vitro* and *in vivo*

This study showed differences in mutant parasite behaviour between culture and sand fly midgut conditions. Unexpectedly, HASP and SHERP were not stage-specifically upregulated, when the mutant lines were infected into the sand fly vector, although gene expression and regulation had been confirmed in culture (Fig.3.9). This was interesting, since the phenotype of *Lmj*cDNA16 dKO only becomes apparent in the sand fly vector, but not in culture (*146*, *332*). It is not clear what causes these differences in mutant parasite behaviour between culture and midgut conditions, but it is possible that the parasites takes cues by midgut environment sensing for the upregulation of the HASP and SHERP genes. These cues would be absent in culture.

The in-/vertebrate hosts play important parts in the *Leishmania* life-cycle. Mammalian infective *Leishmania spp.* are digenetic kinetoplastid parasites and rely on an invertebrate vector to be spread from one long term mammalian host to another in nature. Parasites do not usually simply evade the in-/vertebrate host immune system until they have a chance for transmission. Instead, they make use of carefully regulated mechanisms to sense their environment and regulate their genetic programming for development according to detected changes in the host environment, to optimize their survival and chances for transmission. Better studied examples of parasites taking cues from the sensed host environment can be found in other parasite systems. For example, *Trypanosoma brucei* can detect the density of the *T. brucei* population in the blood stream of a mammalian host via a hypothesised stumpy induction factor (SIF) secreted by the parasite, a mechanism comparable to bacterial quorum sensing (*379*). *T. brucei* will only induce the production of stumpy forms, which are transmissible to tsetse flies, but are cell-cycle arrested, when the population of slender, proliferative forms is sufficiently dense, which is hypothesised to be indicated to the parasite by SIF concentrations in the blood stream. The generation of *Plasmodium* gametocytes transmissible to the invertebrate host is also hypothesised to be dependent on sensed changes in the parasite's environment (reviewed in (*380*)). More is known about parasite behaviour in vertebrate hosts than in invertebrate hosts, but it can be assumed that parasites monitor their environment just as carefully in the invertebrate host to time their development. Little is known, however, about the molecular mechanisms of parasite environmental sensing. One of the few well characterized coupled parasite sensing/development pathways in the invertebrate host is the CCA (citrate and/or *cis*-aconitate) signal-dependent differentiation initiation of *T. brucei* stumpy forms in the tsetse fly (reviewed in (*381*)). The temperature drop to ~20 °C on entry into the tsetse fly stimulates PAD (protein associated with differentiation) expression in *T. brucei*. PAD is a carboxylate transporter for CCA released from the blood meal and secreted by the tsetse fly (*322*, *382*, *383*). Only the combination of temperature drop and CCA presence induces stumpy form differentiation initiation. In *Leishmania*, the mechanism of purine scavenging in promastigotes has been described (*384*). *Leishmania* metacyclogenesis completion is dependent on scavenging purines from the sand fly midgut lumen; if levels are too low, *Leishmania* differentiation is arrested. The mechanism for detection of environmental purine levels is only now being unravelled. *Leishmania* detects environmental purine levels only

indirectly by detection of internal stores of scavenged purines (N. Carter, Kinetoplastid Molecular Cell Biology meeting, Woods Hole, 2013).

The regulation of HASP and SHERP genes could also be influenced by yet unknown environmental cues, which need to work synchronously to allow stable upregulation *in vivo*. The sand fly midgut is a very dynamic environment. Changes in pH, temperature, amino acid and enzyme content are being experienced by the parasites throughout their promastigote stages in the midgut. These changes do not occur in a culture unless artificially induced. Since parasites sense their environment to adapt quickly to changing conditions and promote their survival, changes in parasite behaviour have been observed in culture adapted stains versus natural strains. For example, culture adapted *T. brucei* were observed to have a 1,000x lower antigen switching rate for their variable surface glycoproteins (VSG) than in natural isolates (*385*, *386*). So far, it has not been possible to identify a potential sand fly midgut related trigger for HASPB regulation from the replacement constructs. The addition of midgut extracts from blood fed female sand flies to cultures was insufficient to markedly influence HASPB expression. This could suggest that another factor apart from sand fly midgut molecules/metabolites is required for HASPB upregulation. Multiple triggers that need to work synchronously for parasite gene regulation were shown in the example of *T. brucei* stumpy form differentiation initiation in the tsetse fly, where temperature drop and CCA detection need to occur together to induce stumpy form differentiation (*322*, *382*, *383*). Requiring multiple triggers tightens the parasites control over essential mechanisms. However, in case of the experimental set up of the midgut extract inoculation into M199 *Leishmania* cultures in this study, it is possible that the concentration of midgut extract per culture (1:80 dilution) was too low, that the relevant component was degraded or had been lost during the midgut extract filtering. Protein compounds can potentional bind to filter membranes and are so lost from the extract prior to dilution in M199. Refinement of the experimental approach is required to conclusively prove that construct expression is not suppressed by only midgut molecules and/or metabolites.

The differences in mutant line behaviour could also be of more technical in nature. The *Lmj*cDNA16 sKI mutant line has been the only one to rescue metacyclogenesis completion. *Lmj*cDNA16 sKI contains the whole cDNA16 locus in a single construct and was the only tested mutant line to contain all

HASP and SHERP genes in their natural order and context. The HASPB mRNA expression pattern of *Lmj*cDNA16 sKI was the only one similar to FVI *in vivo*, although not for SHERP and HASPA. However, if HASPB is the key element for metacyclogenesis completion, as suggested by Sádlová *et al.* (2010), it could be that the order and context in which the HASP and SHERP genes occur in the cDNA16 locus are relevant for gene regulation *in vivo*. It has already been shown *in vitro* that genomic location is important for gene regulation in *Leishmania* (*146*, *332*). Gene regulation in *Leishmania* occurs primarily post-transcriptionally and translationally (reviewed in (*315*, *316*)). It could be that mRNA stability is affected differently *in vitro* versus *in vivo* due to environmental conditions and/or signalling pathways. mRNA stability is determined through elements and secondary structures within 3' UTRs. It has been proposed that the 5' *trans*-splice site of a downstream gene can influence the determination of the 3' poly(A)-site of the upstream gene due to the coupled nature of these two processes in *Leishmania* (*313*, *314*). This could mean that the distance of 3' poly(A)-site and following 5' *trans*-splice site is fixed for a gene in *Leishmania*. If that is correct then moving the 5' *trans*-splice site following a 3' poly(A)-site farther downstream in a locus may also relocate the position of the 3' poly(A)-site of the gene upstream of the altered 5' *trans*-splice site. Within the replacement constructs, the HASP and SHERP genes are not within their natural locus context. The 5' *trans*-splice site following the HASPs and SHERP genes in the construct is the one of the antibiotic resistance genes, which is 2-3x further downstream of the known HASP and SHERP 3' poly(A)-sites than the 5' *trans*-splice site that normally follows them in the cDNA16 locus context (Fig.5.8A &B). This could mean that the HASP and SHERP mRNAs become longer, because the 3' poly(A)-sites are relocated farther downstream from their known locations. This could make the mRNA less stable due to 3' UTR extension. Potentially, this effect is amplified *in vivo* due to the sand fly midgut environmental conditions versus the stable culture conditions. It must also be considered that the phenotype of the cDNA16 null mutant is only observed in the sand fly midgut, but not in culture, which suggests that the midgut environment plays a vital role in the function of this locus. This possibility is currently still under investigation.

In many eukaryotic cells, gene regulation is supported at the post-transcriptional level by micro-RNAs (miRNAs), which are a class of endogenous non-coding short RNAs (reviewed in (*387*)). Since post-transcriptional gene regulation is the primary means of gene regulation in

Leishmania, miRNA seems like a reasonable mechanism for gene regulation in these parasites. However, post-transcriptional regulation via miRNA requires a functional RNA interference pathway, to which proteins like Dicer and argonoute belong, that process pre-miRNA and load it onto the RNA-induced silencing complex, respectively. It has been shown that only members of the *L. (Viannia)* subgenus possess a functional RNAi pathway, while proteins like Dicer and argonoute have no identifiable homologues in *L. (Leishmania) spp.* (reviewed in (*388*)). It was concluded that members of the *L. (Leishmania)* subgenus do not use miRNA for gene regulation (*389*). This seems to exclude the possibility of miRNA genes within the cDNA16 locus intergenic regions that allow HASP and SHERP regulation *in vivo*. Although RNAi has been demonstrated to be functional in *L. (Viannia) braziliensis* (*390*), its involvement in gene regulation and its evolution in these species are only now beginning to be revealed (*391*).

Another interesting observation regarding the regulation from the cDNA16 locus was made in this study, when a set of HASPB-GFP fusion gene mutant lines were generated. The HASPB-GFP fusion construct was transfected into several mutant backgrounds (*Lmj*cDNA16 dKO, *Lmj*SHERP sKI, *Lmj*HASPA1 sKI, *Lmj*HASPA1/2 sKI, *Lmj*HASPA2 sKI and *Lmj*cDNA16 sKI) into the second cDNA16 locus allele. The mutant lines were designed to be passaged through sand flies to detect the moment of HASPB upregulation. All mutant lines were rigorously tested by the same approaches as all other mutants in this study. However, whenever the construct was integrated into the cDNA16 locus, the HASPB-GFP fusion gene was never expressed efficiently in culture and GFP fluorescence was undetectable. When an episomal integration had occurred in any mutant background, however, the HASPB-GFP fusion gene was expressed at detectable levels by FACS and confocal microscopy, suggesting that the construct was intact and functionally sound (Fig.7.2). It is unclear, why the construct was not upregulated when integrated into the cDNA16 locus, but this observation points out that facets of mechanisms governing gene regulation in *Leishmania* remain unexplored. A demonstration of the *in vivo* specific suppression of expression of HASP and SHERP would be possible by inoculating mutant lines from the sand fly midgut into M199 culture to see if HASP and/or SHERP expression can be rescued after culture inoculation. Attempts for this have been made in this study, but so far the overgrowth of the cultures by some antibiotic resistant bacterium or fungus has prevented successful parasite growth after inoculation. It may be interesting to integrate

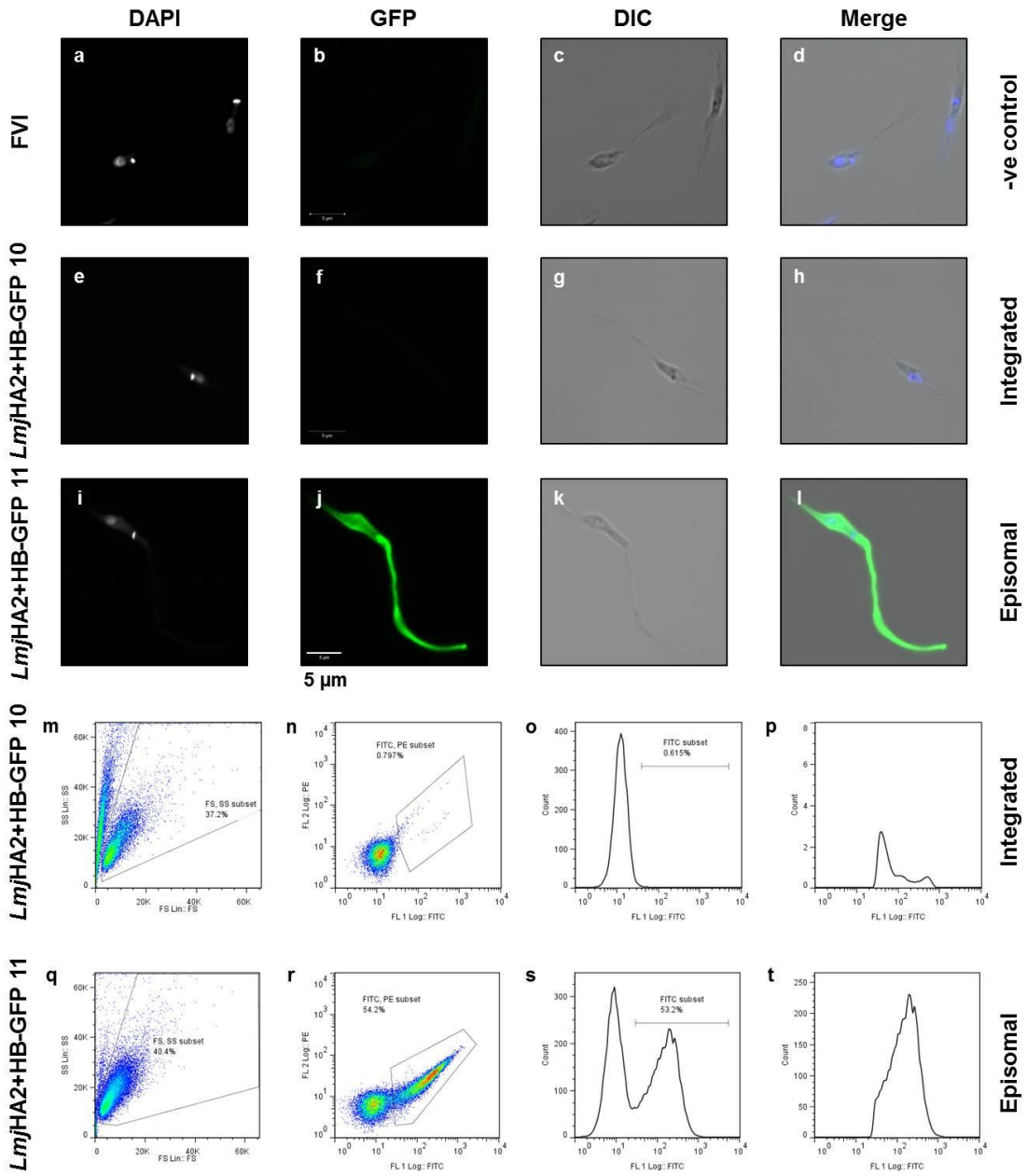


Fig.7.2 – Differences in HASPB-GFP fusion gene expression

The confocal images show the differences in fluorescence of the HASPB-GFP fusion protein expressed from a construct integrated into the cDNA16 locus (e-h) and the same construct transfected as an episome (i-l), while FVI (a-d) served as a negative control. Metacyclics are shown, as determined by measurement, of two clones of the *Lmj*HA2+HB-GFP sKI line as an example. While the clone with the integrated construct does not express the HASPB-GFP fusion gene, the episomal mutant has a strong positive signal from the flagellum and cell surface. FACS analysis of the same mutant confirmed that the *Lmj*HA2+HB-GFP sKI 10 (integrated construct) does not express HASPB-GFP. Graph n shows only a fluorescent negative cell population (fluorescent cells in the boxed area) and no second peak in graph o. Conversely, *Lmj*HA2+HB-GFP sKI 11 (episomal expressor) had a large fluorescent population (r; boxed area & s; right hand peak).

reporter genes into different loci of known regulated genes in *Leishmania* and see how that affects the reporter gene expression *in vitro* and *in vivo*. This may provide some insight into whether the differential expression from constructs in regulated loci is a common theme in *Leishmania*. Whether epigenetics are involved in the differences in HASP and SHERP gene expression has not been excluded yet. The observation that SHERP was detectable in *Lmj*S2+HB sKI mutants in midgut smears in the absence of HASPB expression, while SHERP was undetected in *Lmj*SHERP sKI, which served as a precursor for *Lmj*S2+HB sKI, does not suggest epigenetic involvement. The HASPB construct was integrated into the second allele and, since HASPB remained downregulated, it would not explain changes to the chromatin fold of the chromatid containing the SHERP construct.

Differences in parasite behaviour between natural and artificial systems have been described before for *Leishmania*. In laboratory settings, experimental infections in mammals are conveniently done by subcutaneous or intravenous injection of cultured late stationary stage parasites (needle inoculation). This approach has been questioned since considerable discrepancies in infection outcome have been observed between needle inoculation and transmission by sand fly bite (*54*, *392*, *393*). For instance, mice vaccinated with killed *Leishmania* parasites show protection to needle inoculation, but not to transmission by sand fly bite (*392*, *393*). Other studies have shown that components of sand fly saliva and the PSG enhance *Leishmania* infection establishment and disease progression (*106*, *254*, *255*, *394*). The problem here is the broad variety in the development of transmissible infections both within and between sand fly species that undermine the practicality and physiological relevance of sand fly transmission when studying disease outcomes (*32*, *146*, *373*, *395*). Inoculation by sand fly bite may transmit anything between 10 – 100,000 parasites per bite, although only 1 in 4 infected sand flies are likely to transmit significantly more than 1000 parasites per bite (*32*, *253*, *395*). High dosage transmission correlated to high parasite burdens in the respective vector (>30,000 parasite per midgut). Another key factor in successful transmission is the proportion of metacyclics of the total parasite burden at the point of insect blood feeding. It was shown that sand flies with ≥70% metacyclics in their midgut were more successful in transmission to a mammalian host (>70%) (*395*). These results need to be looked at cautiously, though, when making assumptions about transmission in nature, because under laboratory conditions, experiments are often performed

only on individual sand flies. *Lu. (Lu.) longipalpis* sand flies, however, were observed to feed cooperatively in aggregates rather than individually, which promoted their efficiency in blood uptake and reduced the individual's use of saliva (*396*). It is not known if this is a common trait of sand fly behaviour. There are no observations as to how this behaviour may impact on the transmission of *Leishmania*, whether less PSG is regurgitated too and whether the presence of uninfected sand flies is beneficial to transmission or not. This demonstrates, however, the importance of verifying *in vitro* observations in the natural system.

I have made the observation in this study that all my mutant lines can infect BALB/c mice and cause lesion formation, if the inoculum with late stationary stage parasites from culture is sufficiently high. Considering that PSG secretion, SV colonization and metacyclic generation in the replacement mutant lines of this study are impaired in the sand fly vector, it is questionable if transmission from a sand fly to a new mammalian host can occur. The investigation of this question is currently underway with our collaborators at the NIH, but the expectation is that the mutant lines are not transmissible in a natural setting. This again suggests that it is important to be critical towards *in vitro* data and confirm them in *in vivo* settings.

7.1.4. Using parasite morphology for promastigote stage identification

The problems in using parasite morphology as a tool for promastigote stage determination were already addressed in section 4.2.6. The natural plasticity of *Leishmania* morphology makes a clear division between developmental stages by morphology difficult. Due to undefined intermediate promastigote stages, there is always a gradient in the defining cell body and flagellum measurements, which shows no clear break-off points (Fig.4.9). Instead the majority of measurements appear to accumulate around the artificially introduced break off point, like for example ≥14 µm in cell body length to determine a nectomonad (Fig.7.3). The range in cell body lengths of nectomonads, however, is large reaching >30 µm in length and the longer the cell body, the greater the confidence in the identification of the nectomonad stage becomes. However, how can we judge with confidence that parasites measuring 14µm are any different in their developmental state than those measuring 13.99 µm? The same goes for the distinction between leptomonads and metacyclics, where the defining factor is a two time longer flagellum than the cell body. Fig.7.3.C & E shows two examples where the

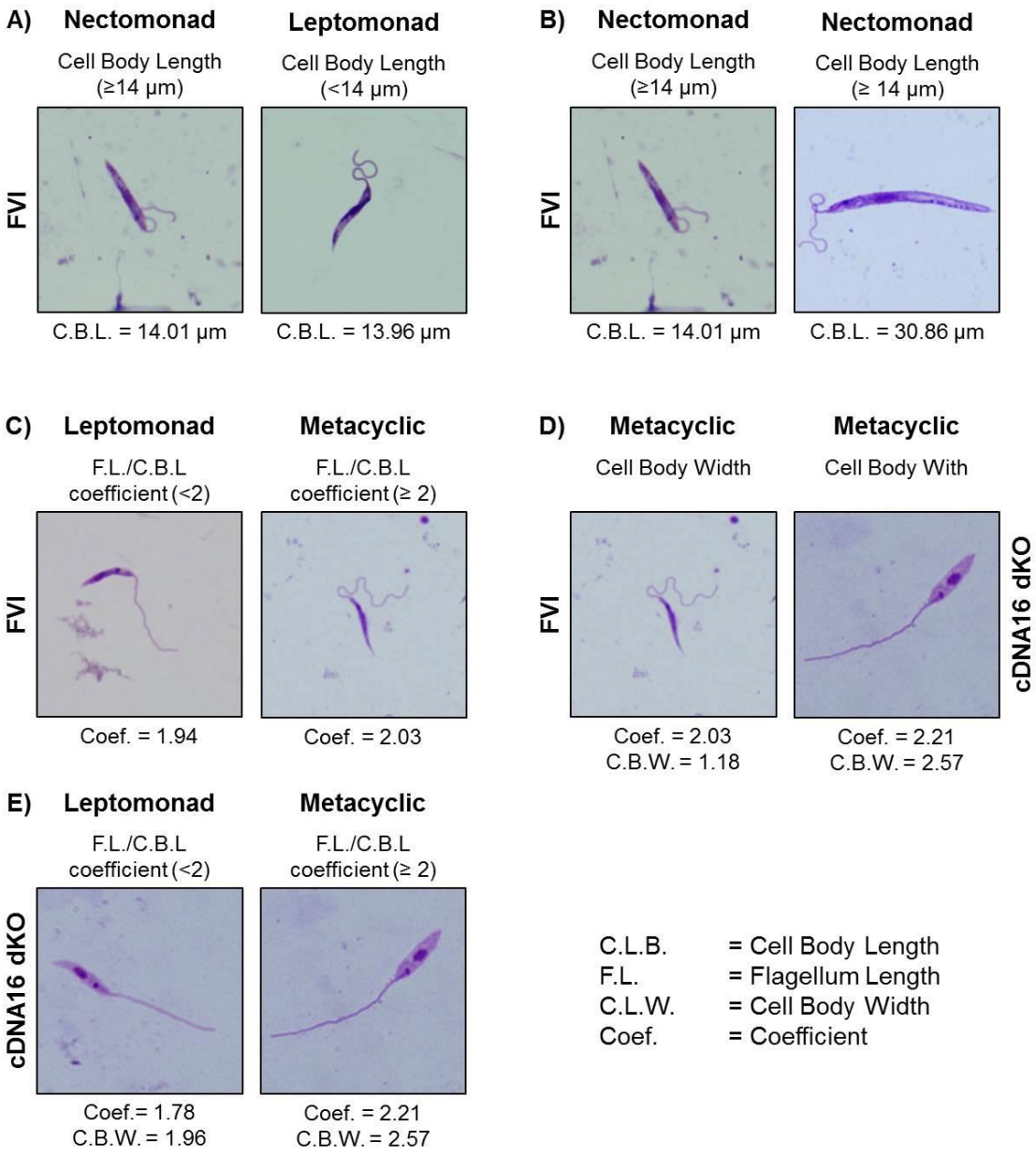


Fig.7.3 – Problems with plasticity of parasite morphology

A) compares two FVI parasite cells that are very similar in length. One just below and the other just above the 14 µm threshold that distinguishes between nectomonads and leptomonads. B) looks at the extreme range of nectomonad cell body lengths. C) compares an FVI leptomonad with a metacyclic which is just below and above the 2x flagella length compared to cell body length threshold, respectively. D) compares two metacyclics by measurement from FVI and *Lmj*cDNA16 sdKO. E) looks at a *Lmj*cDNA16 dKO leptomonad and metacyclic.

leptomonad is just below and the metacyclic just above the threshold. Whether in fact both cells are truly distinct forms molecularly or not is impossible to judge just by measurement. Also, parasite forms, which we have to classify as the same form, may in fact be different from one another. Fig.7.3.D shows a metacyclic by measurement from FVI and one from *Lmj*cDNA16 dKO. The cell body shape of these two metacyclics is distinct as the cell body of the *Lmj*cDNA16 dKO metacyclic is broad with a rounded posterior, while the FVI metacyclic is narrow with a pointed posterior. The only way to overcome this problem is by identifying more promastigote stage specific markers. To date the only accepted stage-specific promastigote markers are SHERP and HASPB, which are specific to the metacyclic stage (*333*, *341*). But there is no way of distinguishing procyclics, nectomonads and leptomonads by marker.

7.1.5. Model of HASP and SHERP regulation mechanism

The current model suggests that HASPB is the key gene for metacyclogenesis completion. Although this could not be proven directly in this study, the qPCR data on mRNA extracted from sand fly midgut and culture derived parasites suggested HASPB to be important. Whether HASPB expression is dependent on prior SHERP or HASPA2 expression in the sand fly midgut could not be shown due to the unexpected differences in HASP and SHERP regulation between sand fly midgut and culture conditions. What is clear is that HASPB function is not essential in culture for metacyclogenesis completion. However, it is feasible the HASPA2 expression and/or SHERP expression supports HASPB upregulation *in vivo*. The surface exposure of HASPB and its disordered structure may suggest that it binds to a ligand, which allows HASPB to fold properly. Since, despite previous efforts, no HASPB binding partner could be identified in the mammalian host, it remains possible that HASPB detects something in the sand fly midgut, which may cue the parasites development. Conversely, being a true metacyclic could be a prerequisite *in vivo* for HASPB expression, since HASPB is metacyclic specific in the sand fly midgut. SHERP may support metacyclogenesis by inducing autophagy in differentiating parasites (*348*), potentially supporting HASPB upregulation indirectly by supporting metacyclic generation via autophagy *in vivo*. However, this could not be shown in this study.

So far, it has not been possible to find any evidence that HASPB is shed in the sand fly midgut, which suggests that shedding may only occur after entry

into the mammalian host skin, perhaps due to HASPB recognition by B-1 B cell-derived natural antibodies (*377*). HASPA2 and HASPA1 promastigote and amastigote specific expression, respectively, may bear some sort of switch function, defining promastigote and amastigote stage by their expression, respectively, *in vivo*. The HASPAs are hypothesised to be dual acylated like HASPB, because of the conserved SH4 domain at the *N*-terminus of the HASPs, and trafficked to the cell surface too. Although they are missing the central repeats of HASPB, they could also be binding to some ligand to perform their function. The hypothesised HASPA and demonstrated HASPB surface exposure in particular on the flagellum would suggest a function in sensing, although the lack of transmembrane domains does not suggest how the signal may be internalized. Perhaps the HASPs act as a co-factor in a signalling complex on ligand binding.

I hypothesis that the differentiation from procyclics into nectomonads occurs independently from the HASPs and SHERP *in vivo*, because nectomonads were always generated efficiently in all mutant lines in the sand fly midgut, including *Lmj*cDNA16 dKO (Fig.7.4). Leptomonad generation may be influenced by the HASPs and SHERP, as was shown by the improved leptomonad generation by the replacement of multiple HASP and SHERP genes. However, it is not clear which HASP and/or SHERP gene(s) may support leptomonad generation. It needs to be considered that low level leptomonad generation also occurs in the *Lmj*cDNA16 dKO line. The generation of true metacyclics is then dependent on HASPB and potentially SHERP, since it was shown that HASPB and SHERP replacement was sufficient to recover metacyclic cell body shape. Also, episomal expression of HASPB seemed to rescue metacyclogenesis completion in Sádlová *et al.* (2010). The lack of HASPB upregulation in mutant lines *in vivo*, could explain why the metacyclic generation was so inefficient in all the mutant lines, assuming that HASPB is the driver for the final step in metacyclogenesis. The transformation into amastigotes may not be dependent on the HASPs, because *Lmj*cDNA16 dKO can infect BALB/c mice on needle inoculation and form amastigotes. However, HASPA2 expression in the absence of HASPA1 influences parasite infectivity in BALB/c mice. HASPA1 function may be relevant to complement HASPA2 function in amastigotes to counteract the negative effect HASPA2 has in the mutant lines once they are infected into BALB/c mice. Interestingly, not having any HASPA gene at all is just as beneficial as having HASPA1 and better than having HASPA2 in the mouse

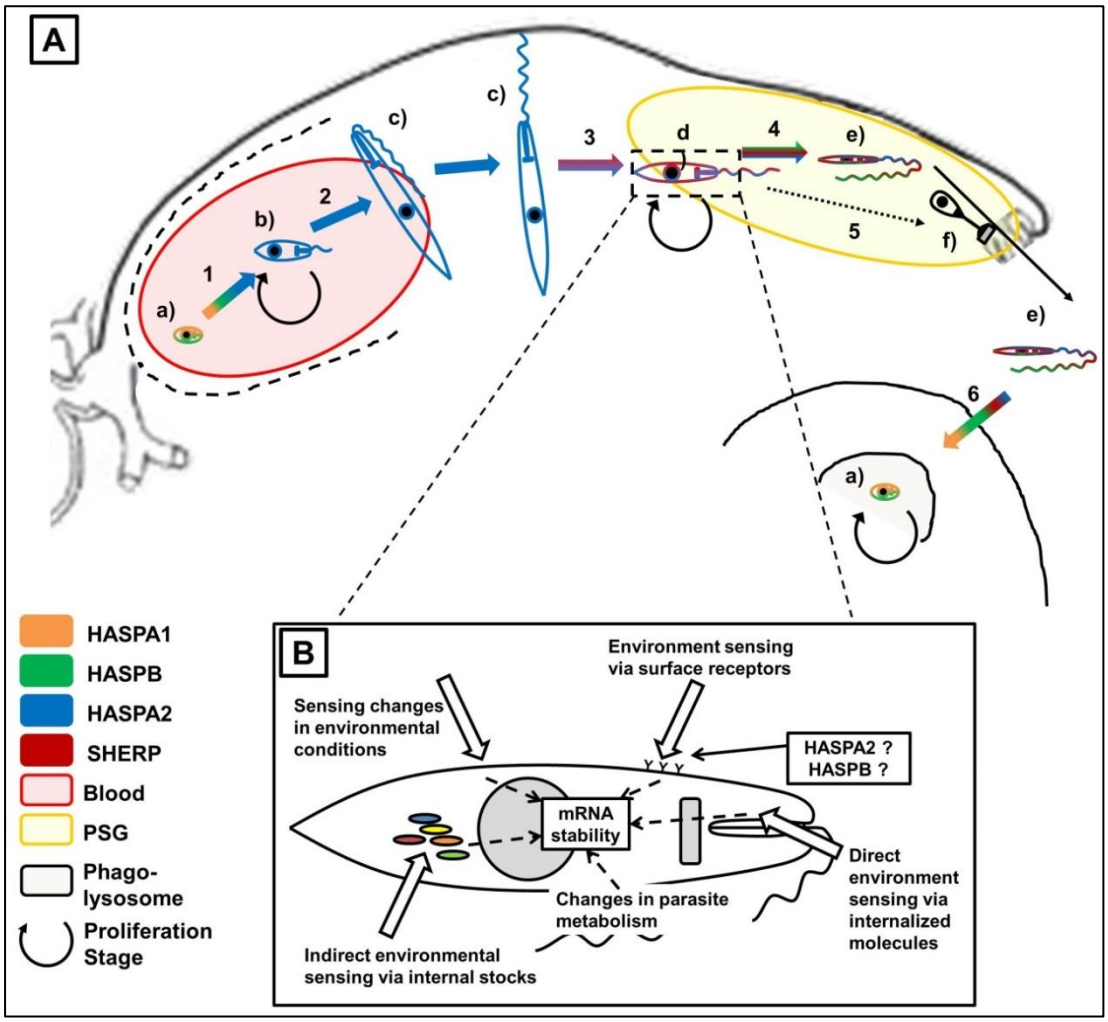


Fig.7.4 – Model for HASP & SHERP regulation *in vivo* during metacyclogenesis

[A] (a) Amastigotes expressing HASPA1 and HASPB are internalized with the blood meal (pale red). 1) As amastigotes differentiate into (b) procyclics HASPA1 and HASPB are downregulated to undetectable protein levels, while HASPA2 beginning to be upregulated. 2) HASPA2 continues to be expressed in (c) nectomonads before and after escape from the PM and during midgut epithelium attachment. 3) SHERP begins to be up-regulated in late (d) leptomonads prior to (e) metacyclic generation *in vivo*. 4) HASPB is then upregulated in (e) metacyclics, which can detach from the midgut wall and are transmitted with the PSG into the mammalian host skin during the next sand fly blood meal. 5) (f) Haptomonads are also formed from (d) leptomonads, but it is not known whether these express SHERP or HASPB. 6) (e) Metacyclics entering the mammalian host skin infect resident macrophages and transform back into (a) amastigotes. During this process HASPA2 and SHERP are downregulated and HASPA1 is upregulated, while HASPB continuous to be expressed. [B] The stage-specific upregulation of the HASPs and SHERP may be governed by a set of mechanisms *in vivo*, which react to different stimuli, like changes in temperature and pH, midgut molecules and metabolites, internal nutrient stores, etc.

model. Whether the switch from HASPA1 to HASPA2 on the transformation from amastigotes into procyclic promastigote is essential *in vivo* is not known, but at least in culture, they are not required for this process. Then again the cDNA16 locus is also not essential in culture for metacyclogenesis, but only in the sand fly midgut.

Why the difference in HASP and SHERP construct expression occurred between *in vitro* and *in vivo* conditions is not clear. It is possible that the problem is mRNA stability. At least in the case of HASPB, a difference in mRNA levels of 2.5 – 5-fold were observed at day 6 PBM *in vivo*. Since gene regulation is primarily post-transcriptionally, this reduction in HASPB mRNA could prevent the genes stage-specific upregulation in the construct mutants. Potentially, locus organization is also important for gene regulation in *Leishmania*.

7.1.6. Perspective on future studies

The differences in parasite behaviour between sand fly midgut and culture settings shown in this study emphasise the need for researchers in this field to verify their *in vitro* findings in the *in vivo* settings. Parasites are designed to adapt to their given environment and it appears that the adaptation of *Leishmania* to culture obscures natural phenotypes. Naturally, this does not mean that culture work should be abandoned altogether, considering the technical limitations of parasite work in the sand fly vector. But culture work on *Leishmania* parasites should always be complemented with *in vivo* work. It becomes also increasingly important to step away from the promastigote/amastigote paradigm and accept that there are at least five morphologically and functionally distinct promastigote forms: procyclic, nectomonad, leptomonad, metacyclic and haptomonad promastigotes (*145*, *397*). More than one amastigote form may exist too (*141*, *142*).

For the work on the HASPs and SHERP, it would be interesting to investigate if the HASPA2 and HASPB have a binding partner in the sand fly midgut and if HASPA1 has one in mammalian macrophages by pull-down assays. The generation of a specific anti-HASPA antibody would be very beneficial to investigate these relatively little investigated proteins. In this study, two attempts have been made to generate an anti-HASPA antibody via two distinct protocols, one using a recombinant-expressed full HASPA protein for immunization and the other using only the HASPA-specific central 19 amino

acid sequence. Neither approach generated an antibody specific for the native HASPA protein, although the recombinant HASPA protein was detected. Further investigation into this is required before it will be possible to isolate an anti-HASPA antibody.

It may be interesting to try to investigate the essentiality of the gene order within the cDNA16 locus for HASP and SHERP upregulation *in vivo*. One could try to randomly delete parts of the locus or integrate stretches of sequence into the intergenic regions and verify in the sand fly vector how that would influences parasite differentiation and expression of the remaining genes. The investigation whether mRNA length is influenced by coupled 5' *trans*-splicing and 3' polyadenylation is still underway and may reveal difference in HASP and SHERP mRNA lengths expressed from recombinant constructs *in vivo*.

7.1.7. Conclusion

Transmission of mammalian-infective *Leishmania* metacyclics by the bite of the sand fly vector is one of the key events in the carefully timed life-cycle of mammalian-infective *Leishmania spp*. Prior to transmission, the generation of metacyclic parasite forms is essential to re-establish parasite infectivity to a mammalian host. Parasite metacyclogenesis is the process in which metacyclics are generated from procyclic promastigotes via the intermediate nectomonad and leptomonad forms. The HASP and SHERP genes are essential to metacyclogenesis within in the sand fly vector, although their functions remain unknown. The studies presented here on the involvement of the HASPs and SHERPs in *L. (L.) major* metacyclogenesis showed various differences in HASP and SHERP mutant behaviour between culture and the sand fly midgut, showing the necessity to observe metacyclogenesis *in vivo* rather than *in vitro*. The data presented here suggest a function for HASPB, SHERP and HASPA2 in parasite differentiation in the sand fly midgut, while HASPA1 function is likely to be amastigote-specific. The problems encountered in this study with the stable upregulation of HASP and SHERP genes from homologous recombination constructs *in vivo*, were suggested to be a problem with mRNA stability pointing towards yet unknown regulation mechanisms of HASP and SHERP genes only active *in vivo*. These studies may offer a base for further investigations into the cDNA16 locus, its regulation *in vivo* and the stage specific functions of the HASPs and SHERP. New techniques will be required to deal with limitations encountered in

working within a sand fly midgut, but these may offer a new opportunity to investigate vector-borne parasites in the natural environment, which could offer brand new insights into the natural behaviour of parasites.

7.2. Part Two: The Orthologous HASP Locus of *L. (V.) braziliensis*

7.2.1. Data Summary

The data generated on the *L. (V.) braziliensis* OHL in this study confirmed the misassembly due to repeat collapses previously suggested by Depledge *et al.* (2010) (*294*). By subcloning, few polymorphisms were identified in the LbrM.23.1110 gene and the data suggested tandem repeats of these genes, while the LbrM.23.1120 gene was shown to be single heterozygous gene. Intraspecies variation in the digestion profile of the OHL suggested variation in the OHL content between *L. (V.) braziliensis* strains (Fig.6.6).

Several attempts to produce a full *L. (V.) braziliensis* OHL deletion mutant failed for yet unknown reasons. Fig.6.17 suggested that all three expected OHL alleles may have been targeted. However, a double deletion of OHL was not possible. This could mean that the OHL is in fact on a diploid chromosome and not a triploid one as suggested by Rogers *et al.* (2011) (*109*). The OHL could be essential for the parasites survival, which would be an important difference to the cDNA16 locus in *L. (Leishmania) spp.*, which could be deleted in *L. (L.) major* without any fitness disadvantage and significant phenotype in culture for the parasites.

7.2.2. The *L. (V.) braziliensis* orthologous HASP locus

The orthologous HASP locus is specific to members of the *L. (Viannia)* subgenus and is found in the same chromosomal region on chromosome 23 as the cDNA16 locus of *L. (Leishmania) spp.* (*294*). The currently available sequence shows the presence of two proteins of unknown function, LbrM.23.1110 and LbrM.23.1120. These genes were shown to bear remarkable structural and biochemical similarities to HASPB in *L. (Leishmania) spp*., both containing central highly antigenic and variable sequence repeats (*294*). It was proposed that *L. (Leishmania)* HASPs and the *L. (Viannia)* orthologous HASP genes (oHASP) have a common origin based on genome comparison data to the closely related monogenetic *Leptomonas seymouri*, which parasitizes insects, ciliates and nematodes (*294*). Therefore, it was proposed that HASPB and the two *L. (Viannia)* oHASPs have

conserved functions, despite the sequence differences (*294*). Hybridization analysis of the OHL in *L. (V.) braziliensis*, *L. (V.) guyanensis* and *L. (V.) peruviana* showed inter- and intraspecies variations of the length of OHL, which was also observed for the cDNA16 locus in *L. (Leishmania) spp.* (*294*, *331*, *334*). The observed numbers of distinct 30nt repeat domains per strain lay between 2 to 6 with 6-15 repeats per oHASP. The observed variation of the OHL *Sac*I digestion profile between tested *L. (V.) braziliensis* strains in this thesis suggested intraspecies variation of this locus too (Fig.6.6). Perhaps the function of the OHL genes invites repeat variations due to antigenic pressure, like in the case of VSG in *Trypanosoma spp.* (*398*). Sequence comparisons of this region of several *L. (V.) braziliensis* strains may reveal variations of 30 nt repeat numbers in the Lbr M.23.1110. Fig.6.13 showed a case of repeat loss in one of the sequenced 1110 copies from *Lbr*2904.

Depledge showed in his thesis that the currently available *L. (V.) braziliensis* OHL sequence suffers from repeat collapses due to highly conserved tandem repeat sequences. The OHL was mapped against the *L. (L.) major* cDNA16 locus sequence during the assembly. However, this region on chromosome 23 was already misassembled in *L. (L.) major*, because of sequence repetitiveness in the cDNA16 locus (*107*), and has not been corrected yet, despite the submission of rectified assembly by the Smith lab (accession no. AJ237587). *Leishmania* genomes (*107*), just as the related *T. brucei* (*285*) and *T. cruzi* genomes (*286*), have a high level of sequence repetition, making genome assemblies difficult (*109*). Manual sequencing approaches may be needed for such regions to provide an accurate assembled sequence.

While Depledge suggested in his thesis that the ORF of LbrM.23.1120 was longer than suggested in the annotated sequence on GeneDB and TriTrypDB, this study showed that LbrM.23.1120 exists in two heterozygous copies and that both the currently available ORF of 1120 and the adjusted 1120 ORF from Depledge are in fact correct. This study also showed that LbrM.23.1110 occurred as multiple copies with distinct SNPs per OHL. 1110 may occur as 4-5 sequentially repeated copies downstream of 1120. Interestingly, the intergenic region, which appears to be repeated between the 1110 copies, is highly conserved. This is unusual for non-coding regions, and it suggests that some selective pressure may rest on the intergenic region. Potentially, a yet-to-be-identified SHERP orthologue is present in the intergenic region. Although it was out of the scope of this study to perform an in depth analysis

of the intergenic region to confirm this hypothesis, a brief analysis, identified 4 potential ORFs of 1-2.5x the length of the SHERP ORF in the cDNA16 locus (Appendix 12). Further analysis is required to determine, whether an actual gene exists in the intergenic region of the OHL.

7.2.3. Current Orthologous HASP Locus model

The data generated in this study together with the work done by Depledge *et al.* (2009 & 2010), would suggest a heterozygous copy of LbrM.23.1120 and 4-5 copies of LbrM.23.1110 per allele with few polymorphism between them (Fig.6.19). All LbrM.23.1110 copies are suggested to be flanked up- and downstream by the conserved intergenic region with the exception of the most downstream 1110 gene copy, which is flanked downstream by the unique 3' flanking region of the OHL. A SHERP orthologue may be contained within the intergenic regions explaining the high level of sequence conservation.

7.2.4. Perspective on future studies

Future work may focus on the manual sequence assembly of the OHL. An accurate OHL map may help with the generation of a full OHL deletion mutant line and the PCR amplification of the whole locus for replacement construct generation, which had both failed in this study. It would be interesting to conclusively prove the essentiality of the OHL for *L. (V.) braziliensis* survival, which would be a significant difference to the cDNA16 locus in *L. (Leishmania) spp*. If, however, the full deletion of OHL should be possible, then the infection of the full OHL deletion mutant into sand flies would be very interesting to observe is metacyclogenesis is going to get stalled too. If yes, the cDNA16 locus and the OHL may prove to be the reason for the distinct supra- and peripylarian development of *L. (Leishmania) spp.* and *L. (Viannia) spp.*, respectively. The identification of a SHERP orthologue would be of high interest to establish further functional similarities between the OHL and the cDNA16 locus and to determine where their differences lie. Commonly, genes found in *Leishmania* genomes in the same chromosomal region have often a similar function (*107*, *109*) and it would be interesting to see if it is true for the OHL and the cDNA16 locus.

7.2.5. Conclusion

A lot less is known about the development of members of the *L. (Viannia)* subgenus than the *L. (Leishmania)* subgenus. While difference in the development of these two species have been shown in the past,

metacyclogenesis in *L. (Viannia) spp.* is much less well studied than in *L. (Leishmania) spp*. Since the cDNA16 locus has been shown to be essential for metacyclogenesis of *L. (L.) major*, it remains a possibility that the OHL is involved in metacyclogenesis too. Potentially, the OHL could promote the peripylarian development of *L. (Viannia) spp.*, while the cDNA16 locus promotes the suprapylarian development of *L. (Leishmania) spp.*, since genetic differences between *L. (Leishmania)* and *L. (Viannia) spp.* are only few (*108*). However, Depledge *et al.* (2010) showed that LbrM.23.1110 and 1120 are preferentially expressed in amastigotes, like HASPA1, which does not suggest a function in the sand fly vector. The failure to generate a full deletion mutant for the OHL suggested, however, that the OHL may be essential for parasite survival even in the promastigote stage. The correction of the sequence of the OHL will be important for the generation of full OHL deletion and replacement mutant lines in *L. (V.) braziliensis*. The data generated in this study may help as a base for future investigations into the OHL assembly.

Appendices

(Appendices on the CD-Rom attached to the back of the thesis)

Appendix 1 – List of human infective Leishmania spp.
Appendix 2 – Primers used for qPCR
Appendix 3 – Primers used for construct generation
Appendix 4 – Primers used DIG-probe generation
Appendix 5 – Primers used for OHL re-sequencing
Appendix 6 – Other primers used in this study
Appendix 7 – Raw data of growth assay (parasites - ml)
Appendix 8 – Raw Data Parasite Localization in the Sand Fly Midgut
Appendix 9 – Raw Data of Midgut Infection Loads (from Light Microscopy)
Appendix 10 – Raw Data of Midgut Infection Loads (from qPCR)
Appendix 11 – Measurements of sand fly midgut derived Leishmania parasites
Appendix 12 – Hypothetical ORFs for SHERP homologue in intergenic region
Appendix 13 – Leishmania (Viannia) braziliensis orthologous HASP locus (Version 1)
Appendix 14 – Leishmania (Viannia) braziliensis orthologous HASP locus (Version 2)
Appendix 15 – Complete LbrM.23.1110 sequences flanked up- and downstream by intergenic region
Appendix 16 – Complete LbrM.23.1110 sequences flanked upstream by the intergenic region and downstream by OHL 3' flanking region

List of abbreviations

A	Adenosine
AAMΦ	Alternatively activating macrophage
AIDS	Acquired Immunodeficiency Syndrome
AMB	Amphotericin B
AMG	Abdominal Midgut
AMP	Antimicrobial peptide
APC	Antigen Presenting Cell
ATP	Adenosine Triphosphate
BLAST	Basis Local Alignment Search Tool
BSD	Blasticidin
CAMΦ	Classically Activating macrophage
cDNA	Complementary DNA
CL	Cutaneous Leishmaniasis
DAPI	4', 6-diamidino-2-phenylindole
DCL	Diffuse Cutaneous Leishmaniasis
DIG	Digoxigenin
DL	Disseminated Cutaneous Leishmaniasis
DNA	Deoxyribonucleic Acid
dNTPs	Deoxyribonucleic Acid Triphosphates
DPMS	Dolichol-phosphate-mannose Synthase
dKI	Double Knock-in (= double replacement of the same gene/locus)
dKO	Double Knock-out (= double deletion of the same gene/locus)
ELISA	Enzyme Linked Immunosorbent Assay
EnS	Endoperitrophic space
ER	Endoplasmatic Reticulum
EU	European Union
FCS	Foetal calf serum
FG	Foregut
fPPG	filamentous proteophosphoglycan
Gal	Galaktose
gDNA	Genomic DNA
Glc	Glucose
GPI	Glycosylphosphatidylinositol
H^+	Hydrogen cation
HAART	Highly Active Anti-Retroviral Therapy
HASP	Hydrophilic Acylated Surface Protein

HG	Hindgut
HIV	Human Immunodeficiency Virus
HSPs	High-scoring Sequence Pairs
HTH	Helix-turn-helix
HYG	Hygromycin
iC3b	Inactive C3b
IFN	Interferon
IgG	Immunoglobulin G
IL	Interleukin
ISP	*S*erine Protease Inhibitors
IVDU	Intravenous Drug User
Kb	Kilo base pair
KDa	Kilo Daltons
KI	Knock-in (= replacement mutant)
KO	Knock-out (= deletion mutant)
L.	*Leishmania*
LACK	*Leishmania* homologue for receptors of activated C kinase
L-AMB	Lipid formulation of Amphotericin B
LAMP	Loop-mediated Isothermal Amplification
LB	Luria-Bertani medium
LCL	Localized Cutaneous Leishmaniasis
LPG	Lipophosphoglycan
LRC	Leishmaniasis Recidiva Cutis
Lu.	*Lutzomyia*
M199	Medium 199
Man	Mannose
MCL	Mucocutaneous Leishmaniasis
miRNA	micro-RNA
ML	Mucosal Leishmaniasis
MoFlo	Modular Flow Cytometer
MPL-SE	Monophosphoryl Lipid and Squalene in a stable Emulsion
mRNA	messenger RNA
MQ H_2O	Milli-Q water
MYA	Million Years Ago
NEO	Neomycin (=Geneticin)
NHS	*N*-hydroxysulfosuccinimide
NMT	*N*-myristoyl transferase
NO	Nitric Oxide

NOS2	Nitric Oxide Synthase 2
nt	Nucleotides
OC	Oligochromatography
oHASP	Orthologous HASP
OHL	Orthologous HASP Locus
ORF	Open Reading Frame
PAB1	Poly(A)-Binding Protein 1
PAC	Puromycin
PAGE	Polyacrylamide Gel Electrophoresis
PBM	Post Blood Meal
PBS	Phosphate Buffered Saline Solution
PCR	Polymerase Chain Reaction
PCR-OC	PCR-oligochromatography
PGCs	Polycistronic Gene Clusters
PGRP	Peptidoglycan Recognition Proteins
Ph.	*Phlebotomus*
p.i.	post inoculum
PKDL	Post Kala-azar Dermal Leishmaniasis
PM	Peritrophic Matrix or Peritrophic Membrane
PPG	Proteophosphoglycan
PSG	Parasite Secretory Gel
qPCR	Quantitative Real Time – PCR
RFLP	Restriction Fragment Length Polymorphism
RNA	Ribonucleic Acid
RT-PCR	Reverse Transcription – Polymerase Chain Reaction
SDS	Sodium Dodecyl Sulphate
SHERP	Small Hydrophilic Endoplasmatic Reticulum Associated Protein
S.	*Sauroleishmania*
sKI	Single Knock-in (stands for single replacement of a gene/locus)
sKO	Single Knock-out (stands for single replacement of a gene/locus)
SAT	Streptothricin
SIF	Stumpy Induction Factor
SL	Splice-leader
SNPs	Single-Nucleotide Polymorphisms
spp.	Species
SSRs	Strand Switch Regions
SV	Stomodeal Valve
TAE	Tris acetate EDTA buffer

TBS	Tris Buffered Saline
TEP1	Thioester-containing Protein 1
TGF	Transforming Growth Factor
Th	T helper
TMG	Thoracic Midgut
TNF	Tumour Necrosis Factor
T_{reg}	T regulator
U	Units
UTR	Untranslated Region
V.	*Viannia*
V-ATPase	Vacuole H^+-ATP Synthase
VL	Visceral Leishmaniasis
VSG	Variable Surface Glycoproteins
WHO	World Health Organisation

References

1. P. Desjeux, *Information on the epidemiology and control of the leishmaniases by country and or territory* (Geneva, 1991), p. 48.

2. P. Desjeux, Leishmaniasis: Public health aspects and control, *Clin. Dermatol.* **14**, 417–423 (1996).

3. J. Alvar *et al.*, Leishmaniasis worldwide and global estimates of its incidence., *PLoS One* **7**, e35671 (2012).

4. K. Rose *et al.*, Cutaneous leishmaniasis in red kangaroos: isolation and characterisation of the causative organisms., *Int. J. Parasitol.* **34**, 655–64 (2004).

5. P. Desjeux, Leishmaniasis: current situation and new perspectives, *Comp. Immunol. Microbiol. Infect. Dis.* **27**, 305–318 (2004).

6. F. Chappuis *et al.*, Visceral leishmaniasis: what are the needs for diagnosis, treatment and control?, *Nat. Rev. Microbiol.* **5**, 873–82 (2007).

7. R. Reithinger *et al.*, Cutaneous leishmaniasis, *Lancet Infect. Dis.* **7**, 581–96 (2007).

8. I. M. Mosleh, E. Geith, L. Natsheh, M. Abdul-Dayem, N. Abotteen, Cutaneous leishmaniasis in the Jordanian side of the Jordan Valley: severe under-reporting and consequences on public health management., *Trop. Med. Int. Heal.* **13**, 855–60 (2008).

9. V. P. Singh *et al.*, Estimation of under-reporting of visceral leishmaniasis cases in Bihar, India., *Am. Soc. Trop. Med. Hyg.* **82**, 9–11 (2010).

10. C. D. Mathers, M. Ezzati, A. D. Lopez, Measuring the burden of neglected tropical diseases: the global burden of disease framework., *PLoS Negl. Trop. Dis.* **1**, e114 (2007).

11. B. Rotureau, Are New World leishmaniases becoming anthroponoses?, *Med. Hypotheses* **67**, 1235–41 (2006).

12. L. A. Acosta *et al.*, Ampliación de la distribución de Lutzomyia longipalpis (Lutz & Neiva, 1912) (Diptera: Psychodidae) en el departamento de Caldas: potencial aumento del riesgo de leishmaniasis visceral, *Biomédica* **33**, 1–22 (2013).

13. D. Campbell-Lendrum, J.-C. Dujardin, Domestic and peridomestic transmission of American cutaneous leishmaniasis: changing epidemiological patterns present new control opportunities, *Mem. Inst. Oswaldo Cruz* **96**, 159–162 (2001).

14. A. J. Lysenko, Distribution of leishmaniasis in the Old World., *Bull. World Health Organ.* **44**, 515–20 (1971).

15. H. Reyburn, M. Rowland, M. Mohsen, B. Khan, C. R. Davies, The prolonged epidemic of anthroponotic cutaneous leishmaniasis in Kabul, Afghanistan: “bringing down the neighbourhood”., *Trans. R. Soc. Trop. Med. Hyg.* **97**, 170–6 (2003).

16. R. Molina *et al.*, Infection of sand flies by humans coinfected with Leishmania infantum and human immunodeficiency virus., *Am. Soc. Trop. Med. Hyg.* **60**, 51–3 (1999).

17. I. Cruz *et al.*, Leishmania in discarded syringes from intravenous drug users, *Lancet* **359**, 1124–1125 (2002).

18. C. B. Palatnik-de-Sousa *et al.*, Transmission of visceral leishmaniasis by blood transfusion in hamsters., *Brazilian J. Med. Biol. Res.* **29**, 1311–5 (1996).

19. Y. Le Fichoux *et al.*, Occurrence of Leishmania infantum parasitemia in asymptomatic blood donors living in an area of endemicity in southern France., *J. Clin. Microbiol.* **37**, 1953–7 (1999).

20. R. W. Ashford, The leishmaniases as emerging and reemerging zoonoses, *Int. J. Parasitol.* **30**, 1269–1281 (2000).

21. A. Pavli, H. C. Maltezou, Leishmaniasis, an emerging infection in travelers., *Int. J. Infect. Dis.* **14**, e1032–9 (2010).

22. I. Arevalo *et al.*, Successful Treatment of Drug-Resistant Cutaneous Leishmaniasis in Humans by Use of Imiquimod, an Immunomodulator, *Clin. Infect. Dis.* **33**, 1847–1851 (2001).

23. M. Calvopina *et al.*, Leishmania isoenzyme polymorphisms in Ecuador: relationships with geographic distribution and clinical presentation., *BMC Infect. Dis.* **6**, 139 (2006).

24. M. Calvopina, E. a Gomez, H. Sindermann, P. J. Cooper, Y. Hashiguchi, Relapse of new world diffuse cutaneous leishmaniasis caused by Leishmania (Leishmania) mexicana after miltefosine treatment., *Am. Soc. Trop. Med. Hyg.* **75**, 1074–7 (2006).

25. M. L. Turetz *et al.*, Disseminated leishmaniasis: a new and emerging form of leishmaniasis observed in northeastern Brazil., *J. Infect. Dis.* **186**, 1829–34 (2002).

26. M. M. Reeder, E. Al., in *The Imaging of Tropical Diseases*,.

27. M. CALVOPINA *et al.*, Atypical clinical variants in New World cutaneous leishmaniasis: disseminated, erysipeloid, and recidiva cutis due to Leishmania (V.) panamensis, *Am. Soc. Trop. Med. Hyg.* **73**, 281–284 (2005).

28. H. W. Murray, J. D. Berman, C. R. Davies, N. G. Saravia, Advances in leishmaniasis, *Lancet* **366**, 1561–1577 (2005).

29. A. Barral *et al.*, Leishmaniasis in Bahia, Brazil: Evidence that Leishmania amazonensis produces a wide spectrum of clinical disease., *Am. J. Trop. Med. Hyg.* **44**, 536–546 (1991).

30. U. Sharma, S. Singh, Immunobiology of leishmaniasis., *Indian J. Exp. Biol.* **47**, 412–23 (2009).

31. J. Alexander, F. Brombacher, T helper1/t helper2 cells and resistance/susceptibility to leishmania infection: is this paradigm still relevant?, *Front. Immunol.* **3**, 80 (2012).

32. N. Kimblin *et al.*, Quantification of the infectious dose of Leishmania major transmitted to the skin by single sand flies, *PNAS* **105**, 10125–10130 (2008).

33. U. Lambertz *et al.*, Secreted virulence factors and immune evasion in visceral leishmaniasis., *J. Leukoc. Biol.* **91**, 887–99 (2012).

34. S. M. Puentes, R. P. da Silva, D. L. Sacks, C. H. Hammer, K. A. Joiner, Serum resistance of metacyclic stage Leishmania major promastigotes is due to release of C5b-9., *J. Immunol.* **145**, 4311–6 (1990).

35. D. M. Mosser, T. a Springer, M. S. Diamond, Leishmania promastigotes require opsonic complement to bind to the human leukocyte integrin Mac-1 (CD11b/CD18)., *J. Cell Biol.* **116**, 511–20 (1992).

36. WHO, Control of the Leishmaniases - Report of a meeting of the WHO Expert Committee on the Leishmaniases, *WHO Tech. Rep. Ser.* **949**, 1 – 201 (2010).

37. A. Isnard, M. T. Shio, M. Olivier, Impact of Leishmania metalloprotease GP63 on macrophage signaling., *Front. Cell. Infect. Microbiol.* **2**, 72 (2012).

38. M. S. Faria, F. C. G. Reis, A. P. C. a Lima, Toll-like receptors in leishmania infections: guardians or promoters?, *J. Parasitol. Res.* **2012**, 930257 (2012).

39. J. M. Ehrchen *et al.*, The absence of cutaneous lymph nodes results in a Th2 response and increased susceptibility to Leishmania major infection in mice., *Infect. Immun.* **76**, 4241–50 (2008).

40. D. Agnello *et al.*, Cytokines and transcription factors that regulate T helper cell differentiation: new players and new insights., *J. Clin. Immunol.* **23**, 147–61 (2003).

41. P. Kropf *et al.*, Toll-like receptor 4 contributes to efficient control of infection with the protozoan parasite Leishmania major., *Infect. Immun.* **72**, 1920–8 (2004).

42. P. Kropf *et al.*, Arginase and polyamine synthesis are key factors in the regulation of experimental leishmaniasis in vivo., *FASEB J.* **19**, 1000–2 (2005).

43. B.-S. Choi *et al.*, Differential impact of L-arginine deprivation on the activation and effector functions of T cells and macrophages., *J. Leukoc. Biol.* **85**, 268–77 (2009).

44. Y. Belkaid, C. A. Piccirillo, S. Mendez, E. M. Shevach, D. L. Sacks, CD4+ CD25+ regulatory T cells control Leishmania major persistence and immunity, *Nature* **420**, 502–507 (2002).

45. S. Nylén, D. L. Sacks, Interleukin-10 and the pathogenesis of human visceral leishmaniasis., *Trends Immunol.* **28**, 378–84 (2007).

46. M. Barral-Netto *et al.*, Transforming growth factor-beta in leishmanial infection: a parasite escape mechanism., *Science (80-.).* **257**, 545–8 (1992).

47. G. van Zandbergen *et al.*, Leishmania disease development depends on the presence of apoptotic promastigotes in the virulent inoculum., *PNAS* **103**, 13837–42 (2006).

48. J. L. M. Wanderley, M. a Barcinski, Apoptosis and apoptotic mimicry: the Leishmania connection., *Cell. Mol. life Sci.* **67**, 1653–9 (2010).

49. K. R. Gantt *et al.*, Activation of TGF-beta by Leishmania chagasi: importance for parasite survival in macrophages., *J. Immunol.* **170**, 2613–20 (2003).

50. I. Follador *et al.*, Epidemiologic and immunologic findings for the subclinical form of Leishmania braziliensis infection., *Clin. Infect. Dis.* **34**, E54–8 (2002).

51. A. Caldas *et al.*, Balance of IL-10 and interferon-γ plasma levels in human visceral leishmaniasis: implications in the pathogenesis, *BMC Infect. Dis.* **5**, 113 (2005).

52. G. F. Cota, M. R. de Sousa, A. Rabello, Predictors of visceral leishmaniasis relapse in HIV-infected patients: a systematic review., *PLoS Negl. Trop. Dis.* **5**, e1153 (2011).

53. R. Rotondo *et al.*, Exocytosis of azurophil and arginase 1-containing granules by activated polymorphonuclear neutrophils is required to inhibit T lymphocyte proliferation., *J. Leukoc. Biol.* **89**, 721–7 (2011).

54. M. E. Rogers, T. Ilg, A. V. Nikolaev, M. A. J. Ferguson, P. A. Bates, Transmission of cutaneous leishmaniasis by sand flies is enhanced by regurgitation of fPPG, *Nature* **430**, 463–467 (2004).

55. C. R. Teixeira *et al.*, Saliva from Lutzomyia longipalpis induces CC chemokine ligand 2/monocyte chemoattractant protein-1 expression and macrophage recruitment., *J. Immunol.* **175**, 8346–53 (2005).

56. WHO, *GLOBAL HIV/AIDS RESPONSE: Epidemic update and health sector progress towards Universal Access* (Geneva: Switzerland, 2011), pp. 1–22.

57. J. Alvar *et al.*, The relationship between leishmaniasis and AIDS: the second 10 years., *Clin. Microbiol. Rev.* **21**, 334–59 (2008).

58. I. Cruz *et al.*, Leishmania/HIV co-infections in the second decade., *Indian J. Med. Res.* **123**, 357–88 (2006).

59. V. Pintado, P. Martín-Rabadán, M. L. Rivera, S. Moreno, E. Bouza, Visceral leishmaniasis in human immunodeficiency virus (HIV)-infected and non-HIV-infected patients. A comparative study., *Medicine (Baltimore).* **80**, 54–73 (2001).

60. M. Olivier, R. Badaró, F. J. Medrano, J. Moreno, The pathogenesis of Leishmania/HIV co-infection: cellular and immunological mechanisms., *Ann. Trop. Med. Parasitol.* **97 Suppl 1**, 79–98 (2003).

61. WHO, Leishmania/HIV co-infection. Epidemiological analysis of 692 retrospective cases, *Wkly. Epidemiol. Rec.* **72**, 49–54 (1997).

62. S. Lyons, H. Veeken, J. Long, Visceral leishmaniasis and HIV in Tigray, Ethiopia., *Trop. Med. Int. Heal.* **8**, 733–9 (2003).

63. R. López-Vélez, The impact of highly active antiretroviral therapy (HAART) on visceral leishmaniasis in Spanish patients who are co-infected with HIV., *Ann. Trop. Med. Parasitol.* **97 Suppl 1**, 143–7 (2003).

64. R. Paredes *et al.*, Leishmaniasis in HIV infection., *J. Postgrad. Med.* **49**, 39–49 (2003).

65. WHO, Leishmaniasis: background information (2010) (available at http://www.who.int/leishmaniasis/burden/en/).

66. B. L. Herwaldt, Leishmaniasis., *Lancet* **354**, 1191–9 (1999).

67. F. J. Andrade-Narváez, S. Medina-Peralta, A. Vargas-Gonzalez, S. B. Canto-Lara, S. Estrada-Parra, The histopathology of cutaneous leishmaniasis due to Leishmania (Leishmania) mexicana in the Yucatan peninsula, Mexico, *Rev. Inst. Med. Trop. Sao Paulo* **47**, 191–194 (2005).

68. F. Chappuis, S. Rijal, A. Soto, J. Menten, M. Boelaert, A meta-analysis of the diagnostic performance of the direct agglutination test and rK39 dipstick for visceral leishmaniasis., *BMJ* **333**, 723 (2006).

69. K. A. Weigle, L. Valderrama, A. L. Arias, C. Santrich, N. G. Saravia, Leishmanin skin test standardization and evaluation of safety, dose, storage, longevity of reaction and sensitization., *Am. Soc. Trop. Med. Hyg.* **44**, 260–71 (1991).

70. M. Ameen, Cutaneous leishmaniasis: advances in disease pathogenesis, diagnostics and therapeutics., *Clin. Exp. Dermatol.* **35**, 699–705 (2010).

71. P. Srivastava, A. Dayama, S. Mehrotra, S. Sundar, Diagnosis of visceral leishmaniasis., *Trans. R. Soc. Trop. Med. Hyg.* **105**, 1–6 (2011).

72. R. Reithinger, J.-C. Dujardin, Molecular diagnosis of leishmaniasis: current status and future applications., *J. Clin. Microbiol.* **45**, 21–5 (2007).

73. S. Deborggraeve *et al.*, Molecular dipstick test for diagnosis of sleeping sickness., *J. Clin. Microbiol.* **44**, 2884–9 (2006).

74. T. Notomi *et al.*, Loop-mediated isothermal amplification of DNA., *Nucleic Acids Res.* **28**, E63 (2000).

75. A. Schubach *et al.*, Detection of Leishmania DNA by polymerase chain reaction in scars of treated human patients., *J. Infect. Dis.* **178**, 911–4 (1998).

76. R. Maurya *et al.*, Evaluation of PCR for diagnosis of Indian kala-azar and assessment of cure, *J. Clin. Microbiol.* **43**, 3038 (2005).

77. U. N. Bramachari, Treatment of Kala-Azar with Intramuscular Injections of Hyper-Acid Antimonyl Tartrate (+ Urethane)., *Ind. Med. Gaz.* **55**, 176–177 (1920).

78. S. Sundar *et al.*, Failure of pentavalent antimony in visceral leishmaniasis in India: report from the center of the Indian epidemic., *Clin. Infect. Dis.* **31**, 1104–7 (2000).

79. J. Seaman *et al.*, Liposomal amphotericin B (AmBisome) in the treatment of complicated kala-azar under field conditions., *Clin. Infect. Dis.* **21**, 188–93 (1995).

80. R. Russo *et al.*, Visceral leishmaniasis in HIV infected patients: treatment with high dose liposomal amphotericin B (AmBisome)., *J. Infect.* **32**, 133–7 (1996).

81. S. Sundar, P. L. Olliaro, Miltefosine in the treatment of leishmaniasis: Clinical evidence for informed clinical risk management., *Ther. Clin. Risk Manag.* **3**, 733–40 (2007).

82. D. O. Santos *et al.*, Leishmaniasis treatment--a challenge that remains: a review., *Parasitol. Res.* **103**, 1–10 (2008).

83. P. Borrelli, A. Imperato, G. Murdaca, M. Scudeletti, Liposomal amphotericin B as first line and secondary prophylactic treatment for visceral leishmaniasis in a patient infected with HIV., *Ann. Ital. Med. Int.* **15**, 169–71 (2000).

84. M. Grögl, A. M. J. Oduola, L. D. C. Cordero, D. E. Kyle, Leishmania spp.: Development of pentostam-resistant clones in vitro by discontinuous drug exposure, *Exp. Parasitol.* **69**, 78–90 (1989).

85. P. E. Manson-Bahr, East African kala-azar with special reference to the pathology, prophylaxis and treatment, *Trans. R. Soc. Trop. Med. Hyg.* **53**, 123–126 (1959).

86. R. Porrozzi, A. Teva, V. F. Amaral, M. V Santos da Costa, G. Grimaldi, Cross-immunity experiments between different species or strains of Leishmania in rhesus macaques (Macaca mulatta)., *Am. Soc. Trop. Med. Hyg.* **71**, 297–305 (2004).

87. A. Khamesipour *et al.*, Leishmanization: use of an old method for evaluation of candidate vaccines against leishmaniasis., *Vaccine* **23**, 3642–8 (2005).

88. S. Noazin *et al.*, Efficacy of killed whole-parasite vaccines in the prevention of leishmaniasis: a meta-analysis., *Vaccine* **27**, 4747–53 (2009).

89. K. J. Evans, L. Kedzierski, Development of Vaccines against Visceral Leishmaniasis., *J. Trop. Med.* **2012**, 892817 (2012).

90. M. Hommel, C. L. Jaffe, B. Travi, G. Milon, Experimental models for leishmaniasis and for testing anti-leishmanial vaccines., *Ann. Trop. Med. Parasitol.* **89 Suppl 1**, 55–73 (1995).

91. N. Rachamim, C. L. Jaffe, Pure protein from Leishmania donovani protects mice against both cutaneous and visceral leishmaniasis., *J. Immunol.* **150**, 2322–31 (1993).

92. A. Ghosh, W. W. Zhang, G. Matlashewski, Immunization with A2 protein results in a mixed Th1/Th2 and a humoral response which protects mice against Leishmania donovani infections., *Vaccine* **20**, 59–66 (2001).

93. J. A. Streit, T. J. Recker, J. E. Donelson, M. E. Wilson, BCG expressing LCR1 of Leishmania chagasi induces protective immunity in susceptible mice., *Exp. Parasitol.* **94**, 33–41 (2000).

94. A. Maroof *et al.*, Therapeutic vaccination with recombinant adenovirus reduces splenic parasite burden in experimental visceral leishmaniasis., *J. Infect. Dis.* **205**, 853–63 (2012).

95. J. Chakravarty *et al.*, A clinical trial to evaluate the safety and immunogenicity of the LEISH-F1+MPL-SE vaccine for use in the prevention of visceral leishmaniasis., *Vaccine* **29**, 3531–7 (2011).

96. E. Nascimento *et al.*, A clinical trial to evaluate the safety and immunogenicity of the LEISH-F1+MPL-SE vaccine when used in combination with meglumine antimoniate for the treatment of cutaneous leishmaniasis., *Vaccine* **28**, 6581–7 (2010).

97. J. Trigo *et al.*, Treatment of canine visceral leishmaniasis by the vaccine Leish-111f+MPL-SE., *Vaccine* **28**, 3333–40 (2010).

98. F. Dantas-Torres, Leishmune vaccine: the newest tool for prevention and control of canine visceral leishmaniosis and its potential as a transmission-blocking vaccine., *Vet. Parasitol.* **141**, 1–8 (2006).

99. T. Okuno, M. Takeuchi, Y. Matsumoto, H. Otsuka, Y. Matsumoto, Pretreatment of leishmania homologue of receptors for activated C kinase (LACK) promotes disease progression caused by Leishmania amazonensis., *Exp. Anim.* **51**, 335–41 (2002).

100. K. Maasho *et al.*, A Leishmania homologue of receptors for activated C-kinase (LACK) induces both interferon-gamma and interleukin-10 in natural killer cells of healthy blood donors., *J. Infect. Dis.* **182**, 570–8 (2000).

101. E. A. Marques-da-Silva *et al.*, Intramuscular immunization with p36(LACK) DNA vaccine induces IFN-gamma production but does not protect BALB/c mice against Leishmania chagasi intravenous challenge., *Parasitol. Res.* **98**, 67–74 (2005).

102. S. L. Croft *et al.*, LeishDNAVAX (2013) (available at www.leishdnavax.org).

103. N. Collin *et al.*, Sand Fly Salivary Proteins Induce Strong Cellular Immunity in a Natural Reservoir of Visceral Leishmaniasis with Adverse Consequences for Leishmania, *PLoS Pathog.* **5**, e1000441 (2009).

104. R. Gomes *et al.*, Immunity to a salivary protein of a sand fly vector protects against the fatal outcome of visceral leishmaniasis in a hamster model., *PNAS* **105**, 7845–50 (2008).

105. F. Oliveira, P. G. Lawyer, S. Kamhawi, J. G. Valenzuela, Immunity to distinct sand fly salivary proteins primes the anti-Leishmania immune response towards protection or exacerbation of disease., *PLoS Negl. Trop. Dis.* **2**, e226 (2008).

106. R. G. Titus, J. M. C. Ribeiro, Salivary gland lysates from the sand fly Lutzomyia longipalpis enhance Leishmania infectivity., *Science (80-.).* **239**, 1306–8 (1988).

107. A. C. Ivens *et al.*, The genome of the kinetoplastid parasite, Leishmania major, *Science (80-.).* **309**, 436–442 (2005).

108. C. S. Peacock *et al.*, Comparative genomic analysis of three Leishmania species that cause diverse human disease., *Nat. Genet.* **39**, 839–47 (2007).

109. M. B. Rogers *et al.*, Chromosome and gene copy number variation allow major structural change between species and strains of Leishmania, *Genome Res.* **21**, 2129–2142 (2011).

110. E. Dumonteil, Vaccine development against Trypanosoma cruzi and Leishmania species in the post-genomic era., *Infect. Genet. Evol.* **9**, 1075–82 (2009).

111. V. Saf'janova, in *The Leishmanias, Protozoology*, (Academy of Sciences; All Union Society of Protozoologists, Leningrad, 1982), pp. 95–101.

112. R. Lainson, J. J. Shaw, in *The leishmaniases in biology and medicine. Volume I. Biology and epidemiology*, W. Peters, R. Killick-Kendrick, Eds. (Academic Press, London, 1987), pp. 1–120.

113. R. Lainson, R. D. Ward, J. J. Shaw, Leishmania in Phlebotomid Sandflies: VI. Importance of Hindgut Development in Distinguishing between Parasites of the Leishmania mexicana and L. braziliensis Complexes, *Proc. R. Soc. London. Ser. B, Biol. Sci.* **199**, 309–320 (1977).

114. D. G. Croan, D. A. Morrison, J. T. Ellis, Evolution of the genus Leishmania revealed by comparison of DNA and RNA polymerase gene sequences1, *Mol. Biochem. Parasitol.* **89**, 149–159 (1997).

115. H. A. Noyes *et al.*, A previously unclassified trypanosomatid responsible for human cutaneous lesions in Martinique (French West Indies) is the most divergent member of the genus Leishmania ss., *Parasitology* **124**, 17–24 (2002).

116. L. L. Walters, G. L. Chaplin, G. B. Modi, R. B. Tesh, Ultrastructural biology of Leishmania (Viannia) panamensis (= Leishmania braziliensis panamensis) in Lutzomyia gomezi (Diptera: Psychodidae): a natural host-parasite association, *Am. Soc. Trop. Med. Hyg.* **40**, 19–39 (1989).

117. L. L. Walters, G. B. Modi, G. L. Chaplin, R. B. Tesh, Ultrastructural development of Leishmania chagasi in its vector, Lutzomyia longipalpis (Diptera: Psychodidae)., *Am. Soc. Trop. Med. Hyg.* **41**, 295–317 (1989).

118. M. D. S. Do Brasil, *Atlas de leishmaniose tegumentar americana* L. Araújo, A. Campos, G. Leitão, C. Profeta, V. Medeiros, Eds. (EDITORA MS, Brasilia, ed. 1st, 2006; http://bvsms.saude.gov.br/bvs/publicacoes/atlas_lta.pdf), pp. 1–135.

119. A.-L. Bañuls, M. Hide, F. Prugnolle, Leishmania and the Leishmaniases: A Parasite Genetic Update and Advances in Taxonomy, Epidemiology and Pathogenicity in Humans, *Adv Parasitol* **64**, 1–109 (2007).

120. R. Lainson, The Neotropical Leishmania species: a brief historical review of their discovery, ecology and taxonomy, *Rev. Pan-Amazônica Saúde* **1**, 13–32 (2010).

121. R. Lainson, On Leishmania enriettii and other enigmatic Leishmania species of the Neotropics., *Mem. Inst. Oswaldo Cruz* **92**, 377–87 (1997).

122. M. Tibayrenc, S. Ben Abderrazak, F. Guerrini, A. Bañuls, Leishmania and the clonal theory of parasitic protozoa., *Arch. Inst. Pasteur Tunis* **70**, 375–82 (1993).

123. J. Lukeš *et al.*, Evolutionary and geographical history of the Leishmania donovani complex with a revision of current taxonomy, *PNAS* **104**, 9375–9380 (2007).

124. H. A. Noyes, Implications of a Neotropical origin of the genus Leishmania, *Mem. Inst. Oswaldo Cruz* **93**, 657–661 (1998).

125. S. F. Kerr, R. Merkelz, C. MacKinnon, Further support for a palaearctic origin of Leishmania, *Mem. Inst. Oswaldo Cruz* **95**, 579–581 (2000).

126. S. F. Kerr, Palaearctic origin of Leishmania, *Mem. Inst. Oswaldo Cruz* **95**, 75–80 (2000).

127. H. Momen, E. Cupolillo, Speculations on the origin and evolution of the genus Leishmania, *Mem. Inst. Oswaldo Cruz* **95**, 583–588 (2000).

128. C. B. C. Cox, P. D. P. Moore, *Biogeography: an ecological and evolutionary approach* (Blackwell Science, Oxford, ed. 8th, 2010; http://books.google.co.uk/books?hl=en&lr=&id=GP5HeCwkV2IC&oi=fnd&pg=PR11&dq=Biogeography+and+ecological+and+evolutionary+approaches+2000&ots=KcdsM0SnwD&sig=tZkAnqQOvyZX4FR4LqfxAAUqs8Q), p. 223 pp.

129. J. A. Wolfe, An analysis of Neogene climates in Beringia, *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **108**, 207–216 (1994).

130. R. Killick-Kendrick, Some epidemiological consequences of the evolutionary fit between Leishmaniae and their phlebotomine vectors., *Bull. Soc. Pathol. Exot. Filiales* **78**, 747–755 (1985).

131. D. J. Lewis, A taxonomic review of the genus Phlebotomus (Diptera: Psychodidae), *Bull. Br. Museum (Natural Hist. Entomol. Ser.* **45**, 121–209 (1982).

132. P. D. Ready, Biology of phlebotomine sand flies as vectors of disease agents., *Annu. Rev. Entomol.* **58**, 227–50 (2013).

133. P. D. Ready, Should sand fly taxonomy predict vectorial and ecological traits?, *J. vector Ecol.* **36 Suppl 1**, S17–22 (2011).

134. R. M. Novak, *Walker's Mammals of the World* (The Johns Hopkins University Press, Baltimore and London, ed. 5th, 1991), pp. 643–1629.

135. H. A. Noyes, B. A. Arana, M. L. Chance, R. Maingon, The Leishmania hertigi (Kinetoplastida; Trypanosomatidae) complex and the lizard Leishmania: their classification and evidence for a neotropical origin of the Leishmania-Endotrypanum clade., *J. Eukaryot. Microbiol.* **44**, 511–7 (1997).

136. N. Añez, E. Nieves, D. Cazorla, The validity of the developmental pattern in the sandfly gut for classification of Leishmania., *Trans. R. Soc. Trop. Med. Hyg.* **83**, 634–5 (1989).

137. A. P. Fernandes, K. Nelson, S. M. Beverley, Evolution of nuclear ribosomal RNAs in kinetoplastid protozoa: perspectives on the age and origins of parasitism., *PNAS* **90**, 11608–12 (1993).

138. WHO, *The leishmaniasis, technical report series 701* (1984), p. 140.

139. R. W. Ashford, Leishmaniasis reservoirs and their significance in control., *Clin. Dermatol.* **14**, 523–32 (1996).

140. D. L. Sacks, S. Kamhawi, Molucular aspects of Parasite-Vector and Vector-Host Interactions in leishmaniasis, *Annu. Rev. Microbiol.* **55**, 453–483 (2001).

141. K. Aoun, M. K. Chahed, M. Mokni, Z. Harrat, A. Bouratbine, Importance of amastigote forms morphology to differentiate Leishmania infantum and Leishmania major species., *Arch. Inst. Pasteur Tunis* **80**, 53–6 (2003).

142. M. W. Daboul, Is the Amastigote Form of *Leishmania* the Only Form Found in Humans Infected With Cutaneous Leishmaniasis?, *Lab. Med.* **39**, 38–41 (2008).

143. D. L. Sacks, N. Noben-Trauth, The immunology of susceptibility and resistance to Leishmania major in mice., *Nat. Rev. Immunol.* **2**, 845–58 (2002).

144. S. Kamhawi, Phlebotomine sand flies and Leishmania parasites: friends or foes?, *Trends Parasitol.* **22**, 439–445 (2006).

145. S. M. Gossage, M. E. Rogers, P. A. Bates, Two separate growth phases during the development of Leishmania in sand flies: implications for understanding the life cycle, *Int. J. Parasitol.* **33**, 1027–1034 (2003).

146. J. Sádlová *et al.*, The stage-regulated HASPB and SHERP proteins are essential for differentiation of the protozoan parasite Leishmania major in its sand fly vector, Phlebotomus papatasi., *Cell. Microbiol.* **12**, 1765–79 (2010).

147. S. Méndez *et al.*, Partial anaerobiosis induces infectivity of Leishmania infantum promastigotes., *Parasitol. Res.* **85**, 507–9 (1999).

148. M. L. Cunningham, R. G. Titus, S. J. Turco, S. M. Beverley, Regulation of Differentiation to the Infective Stage of the Protozoan Parasite Leishmania major by Tetrahydrobiopterin, *Science (80-.).* **292**, 285–287 (2001).

149. P. A. Bates, L. Tetley, Leishmania mexicana: Induction of Metacyclogenesis by Cultivation of Promastigotes at Acidic pH, *Exp. Parasitol.* **76**, 412–423 (1993).

150. P. A. Bates, M. E. Rogers, New insights into the developmental biology and transmission mechanisms of Leishmania., *Curr. Mol. Med.* **4**, 601–9 (2004).

151. P. F. P. Pimenta, G. B. Modi, S. T. Pereira, M. Shahabuddin, D. L. Sacks, A novel role for the peritrophic matrix in protecting Leishmania from the hydrolytic activities of the sand fly midgut, *Parasitology* **115**, 359–369 (1997).

152. A. Warburg, The structure of the female sand fly (Phlebotomus papatasi) alimentary canal, *Trans. R. Soc. Trop. Med. Hyg.* **102**, 161–166 (2008).

153. P. A. Bates, Transmission of Leishmania metacyclic promastigotes by phlebotomine sand flies, *Int. J. Parasitol.* **37**, 1097–1106 (2007).

154. R. P. P. Soares *et al.*, Differential midgut attachment of Leishmania (Viannia) braziliensis in the sand flies Lutzomyia (Nyssomyia) whitmani and Lutzomyia (Nyssomyia) intermedia., *J. Biomed. Biotechnol.* **2010**, 439174 (2010).

155. R. Wilson *et al.*, Stage-specific adhesion of leishmania promastigotes to sand fly midguts assessed using an improved comparative binding assay., *PLoS Negl. Trop. Dis.* **4**, 1–9 (2010).

156. J. Sádlová, P. Volf, Peritrophic matrix of Phlebotomus duboscqi and its kinetics during Leishmania major development., *Cell Tissue Res.* **337**, 313–25 (2009).

157. J. Cihláková, P. Volf, Development of different Leishmania major strains in the vector sandflies Phlebotomus papatasi and P. duboscqi., *Ann. Trop. Med. Parasitol.* **91**, 267–79 (1997).

158. D. L. Sacks, Leishmania-sand fly interactions controlling species-specific vector competence., *Cell. Microbiol.* **3**, 189–96 (2001).

159. R. P. P. Soares *et al.*, Leishmania chagasi: lipophosphoglycan characterization and binding to the midgut of the sand fly vector Lutzomyia longipalpis., *Mol. Biochem. Parasitol.* **121**, 213–24 (2002).

160. R. Vaidyanathan, Isolation of a Myoinhibitory Peptide from Leishmania major (Kinetoplastida: Trypanosomatidae) and Its Function in the Vector Sand Fly Phlebotomus papatasi (Diptera: Psychodidae), *J. Med. Entomol.* **42**, 142–152 (2005).

161. M. E. Rogers, M. L. Chance, P. A. Bates, The role of promastigote secretory gel in the origin and transmission of the infective stage of Leishmania mexicana by the sandfly Lutzomyia longipalpis., *Parasitology* **124**, 495–507 (2002).

162. P. Volf, M. Hajmová, J. Sádlová, J. Votýpka, Blocked stomodeal valve of the insect vector: similar mechanism of transmission in two trypanosomatid models., *Int. J. Parasitol.* **34**, 1221–7 (2004).

163. D. L. Sacks, P. V. Perkins, Development of infective stage Leishmania promastigotes within phlebotomine sand flies, *Am. Soc. Trop. Med. Hyg.* **34**, 456–459 (1985).

164. E. Nieves, P. F. P. Pimenta, Development of Leishmania (Viannia) braziliensis and Leishmania (Leishmania) amazonensis in the sand fly Lutzomyia migonei (Diptera: Psychodidae), *J. Med. Entomol.* **37**, 134–140 (2000).

165. R. P. P. Soares *et al.*, Leishmania braziliensis: a novel mechanism in the lipophosphoglycan regulation during metacyclogenesis., *Int. J. Parasitol.* **35**, 245–53 (2005).

166. G. Winter, M. Fuchs, M. J. McConville, Y.-D. Stierhof, P. Overath, Surface antigens of Leishmania mexicana amastigotes: characterization of glycoinositol phospholipids and a macrophage-derived glycosphingolipid., *J. Cell Sci.* **107 (Pt 9**, 2471–82 (1994).

167. B. J. Mengeling, S. M. Beverley, S. J. Turco, Designing glycoconjugate biosynthesis for an insidious intent: phosphoglycan assembly in Leishmania parasites., *Glycobiology* **7**, 873–80 (1997).

168. T. Ilg, Lipophosphoglycan is not required for infection of macrophages or mice by Leishmania mexicana., *EMBO J.* **19**, 1953–62 (2000).

169. D. L. Sacks *et al.*, The role of phosphoglycans in Leishmania-sand fly interactions., *PNAS* **97**, 406–11 (2000).

170. S. Kamhawi, G. B. Modi, P. F. P. Pimenta, E. D. Rowton, D. L. Sacks, The vectorial competence of Phlebotomus sergenti is specific for Leishmania tropica and is controlled by species-specific, lipophosphoglycan-mediated midgut attachment., *Parasitology* **121 (Pt 1)**, 25–33 (2000).

171. T. Ilg *et al.*, Structure of Leishmania mexicana lipophosphoglycan., *J. Biol. Chem.* **267**, 6834–40 (1992).

172. S. C. Ilgoutz, K. A. Mullin, B. R. Southwell, M. J. McConville, Glycosylphosphatidylinositol biosynthetic enzymes are localized to a stable tubular subcompartment of the endoplasmic reticulum in Leishmania mexicana., *EMBO J.* **18**, 3643–54 (1999).

173. K. A. Mullin *et al.*, Regulated degradation of an endoplasmic reticulum membrane protein in a tubular lysosome in Leishmania mexicana., *Mol. Biol. Cell* **12**, 2364–77 (2001).

174. S. J. Turco, A. Descoteaux, The lipophosphoglycan of Leishmania parasites., *Annu. Rev. Microbiol.* **46**, 65–94 (1992).

175. E. M. B. Saraiva, P. F. P. Pimenta, Changes in lipophosphoglycan and gene expression associated with the development of Leishmania major in Phlebotomus papatasi, *Parasitology* **111**, 275–287 (1995).

176. R. Killick-Kendrick, Phlebotomine vectors of the leishmaniases: a review., *Med. Vet. Entomol.* **4**, 1–24 (1990).

177. S. Adler, M. Ber, The transmission of Leishmania tropica by the bite of Phlebotomus papatasi, *Indian J. Med. Res* **29**, 803–809 (1941).

178. C. S. Swaminath, H. E. Shortt, L. A. P. Anderson, Transmission of Indian kala-azar to man by the bites of Phlebotomus argentipes, *Ann Brun Indian J med Res* **30**, 473–477 (1942).

179. Z. Berdjane-Brouk *et al.*, First detection of Leishmania major DNA in Sergentomyia (Spelaeomyia) darlingi from cutaneous leishmaniasis foci in Mali., *PLoS One* **7**, e28266 (2012).

180. M. J. Mutinga *et al.*, Cutaneous leishmaniasis in Kenya: Sergentomyia garnhami (Diptera Psychodidae), a possible vector of Leishmania major in Kitui District: a new focus of the disease., *East Afr. Med. J.* **71**, 424–8 (1994).

181. P. Parvizi, P. D. Ready, Nested PCRs and sequencing of nuclear ITS-rDNA fragments detect three Leishmania species of gerbils in sandflies from Iranian foci of zoonotic cutaneous leishmaniasis., *Trop. Med. Int. Heal.* **13**, 1159–71 (2008).

182. A. K. Seccombe, P. D. Ready, L. M. Huddleston, A catalogue of Old World phlebotomine sandlfies (Diptera: Psychodidae: Phlebotominae), *Occas. Pap. Syst. Entomol.* **8**, 1–57 (1993).

183. J. Sádlová *et al.*, Sergentomyia schwetzi is not a competent vector for Leishmania donovani and other Leishmania species pathogenic to humans., *Parasit. Vectors* **6**, 186 (2013).

184. A. M. Dougall *et al.*, Evidence incriminating midges (Diptera: Ceratopogonidae) as potential vectors of Leishmania in Australia., *Int. J. Parasitol.* **41**, 571–9 (2011).

185. R. Killick-Kendrick, The biology and control of phlebotomine sand flies., *Clin. Dermatol.* **17**, 279–89 (1999).

186. R. P. Lane, Sandflies (Phlebotominae), *Med. Insects Arachn. Chapman Hall, London* , 78–119 (1993).

187. Y. Schlein, A. Warburg, Phytophagy and the feeding cycle of Phlebotomus papatasi (Diptera: Psychodidae) under experimental conditions, *J. Med. Entomol.* **23**, 11–15 (1986).

188. M. D. Feliciangeli, Natural breeding places of phlebotomine sandflies., *Med. Vet. Entomol.* **18**, 71–80 (2004).

189. N. T. Davis, Leishmaniasis in the Sudan republic. 28. Anatomical studies on Phlebotomus orientalis Parrot and P. papatasi Scopoli (Diptera: Psychodidae), *J. Med. Entomol.* **4**, 50–65 (1967).

190. R. P. P. Soares, S. J. Turco, Lutzomyia longipalpis (Diptera: Psychodidae: Phlebotominae): a review., *An Acad Bras Cienc* **75**, 301–30 (2003).

191. Y. Tang, R. D. Ward, Stomodaeal valve ultrastructure in the sandfly Lutzomyia longipalpis (Diptera: Psychodidae)., *Med. Vet. Entomol.* **12**, 132–5 (1998).

192. Y. Tang, R. D. Ward, Sugar feeding and fluid destination control in the phlebotomine sandfly Lutzomyia longipalpis (Diptera: Psychodidae)., *Med. Vet. Entomol.* **12**, 13–9 (1998).

193. W. Rudin, H. Hecker, Functional morphology of the midgut of a sandfly as compared to other hematophagous nematocera, *Tissue Cell* **14**, 751–758 (1982).

194. R. F. Chapman, in *Comprehensive insect physiology, biochemistry and pharmacology*, G. A. Kerkut, L. I. Gilbert, Eds. (Pergamon Press, Oxford, 1985), pp. 165–205.

195. J. Sádlová, The life history of Leishmania (Kinetoplastida: Trypanosomatidae), *Acta Soc Zool Bohem* **63**, 331–366 (1999).

196. R. J. Dillon, E. El-Kordy, Carbohydrate digestion in sandflies: alpha-glucosidase activity in the midgut of Phlebotomus langeroni., *Comp Biochem Physiol B* **116B**, 35–40 (1997).

197. R. Dillon, Introduction to Sand Flies (2008) (available at http://pcwww.liv.ac.uk/leishmania/life_cycle__habitats.htm).

198. A. Warburg, Y. Schlein, The effect of post-bloodmeal nutrition of Phlebotomus papatasi on the transmission of Leishmania major., *Am. Soc. Trop. Med. Hyg.* **35**, 926–30 (1986).

199. V. C. Santos, R. N. Araujo, L. a D. Machado, M. H. Pereira, N. F. Gontijo, The physiology of the midgut of Lutzomyia longipalpis (Lutz and Neiva 1912): pH in different physiological conditions and mechanisms involved in its control., *J. Exp. Biol.* **211**, 2792–8 (2008).

200. N. F. Gontijo *et al.*, Lutzomyia longipalpis: pH in the gut, digestive glycosidases, and some speculations upon Leishmania development., *Exp. Parasitol.* **90**, 212–9 (1998).

201. W. R. Terra, C. Ferreira, J. E. Baker, in *Biology of the insect midgut*, M. J. Lehane, P. F. Bilingsley, Eds. (Chapman & Hall, London, 1996), pp. 235–236.

202. M. del Pilar Corena *et al.*, Carbonic anhydrase in the adult mosquito midgut., *J. Exp. Biol.* **208**, 3263–73 (2005).

203. V. C. Santos, C. a Nunes, M. H. Pereira, N. F. Gontijo, Mechanisms of pH control in the midgut of Lutzomyia longipalpis: roles for ingested molecules and hormones., *J. Exp. Biol.* **214**, 1411–8 (2011).

204. H. A. Zakai, M. L. Chance, P. A. Bates, In vitro stimulation of metacyclogenesis in Leishmania braziliensis, L. donovani, L. major and L. mexicana, *Parasitology* **116**, 305–309 (1998).

205. I. M. A. Felipe *et al.*, Leishmania infection in humans, dogs and sandflies in a visceral leishmaniasis endemic area in Maranhão, Brazil., *Mem. Inst. Oswaldo Cruz* **106**, 207–11 (2011).

206. A. Dostálová, P. Volf, Leishmania development in sand flies: parasite-vector interactions overview., *Parasit. Vectors* **5**, 276 (2012).

207. M. R. V Sant'Anna, H. Diaz-Albiter, M. Mubaraki, R. J. Dillon, P. A. Bates, Inhibition of trypsin expression in Lutzomyia longipalpis using RNAi enhances the survival of Leishmania., *Parasit. Vectors* **2**, 62 (2009).

208. M. Ramalho-Ortigão *et al.*, Exploring the midgut transcriptome of Phlebotomus papatasi: comparative analysis of expression profiles of sugar-fed, blood-fed and Leishmania-major-infected sandflies., *BMC Genomics* **8**, 300 (2007).

209. R. C. Jochim *et al.*, The midgut transcriptome of Lutzomyia longipalpis: comparative analysis of cDNA libraries from sugar-fed, blood-fed, post-digested and Leishmania infantum chagasi-infected sand flies., *BMC Genomics* **9**, 15 (2008).

210. L. Shao, M. Devenport, M. Jacobs-Lorena, The peritrophic matrix of hematophagous insects., *Arch. Insect Biochem. Physiol.* **47**, 119–25 (2001).

211. A. Dostálová *et al.*, The midgut transcriptome of Phlebotomus (Larroussius) perniciosus, a vector of Leishmania infantum: comparison of sugar fed and blood fed sand flies., *BMC Genomics* **12**, 223 (2011).

212. A. Zaidman-Rémy *et al.*, The Drosophila amidase PGRP-LB modulates the immune response to bacterial infection., *Immunity* **24**, 463–73 (2006).

213. S. Meister *et al.*, Anopheles gambiae PGRPLC-mediated defense against bacteria modulates infections with malaria parasites., *PLoS Pathog.* **5**, e1000542 (2009).

214. L. S. Morrison *et al.*, Ecotin-like serine peptidase inhibitor ISP1 of Leishmania major plays a role in flagellar pocket dynamics and promastigote differentiation., *Cell. Microbiol.* **14**, 1271–86 (2012).

215. S. C. P. Eschenlauer *et al.*, Influence of parasite encoded inhibitors of serine peptidases in early infection of macrophages with Leishmania major., *Cell. Microbiol.* **11**, 106–20 (2009).

216. N. C. Secundino *et al.*, Proteophosphoglycan confers resistance of Leishmania major to midgut digestive enzymes induced by blood feeding in vector sand flies., *Cell. Microbiol.* **12**, 906–18 (2010).

217. S. Kumar, A. Molina-Cruz, L. Gupta, J. Rodrigues, C. Barillas-Mury, A peroxidase/dual oxidase system modulates midgut epithelial immunity in Anopheles gambiae., *Science (80-.).* **327**, 1644–8 (2010).

218. R. H. G. Baxter *et al.*, Structural basis for conserved complement factor-like function in the antimalarial protein TEP1, *PNAS* **104**, 11615–11620 (2007).

219. S. Blandin, S. Shiao, L. Moita, C. Janse, Complement-Like Protein TEP1 Is a Determinant of Vectorial Capacity in the Malaria Vector< i> Anopheles gambiae</i>, *Cell* **116**, 661–670 (2004).

220. G. D. A. Oliveira, J. Lieberman, C. Barillas-Mury, Epithelial nitration by a peroxidase/NOX5 system mediates mosquito antiplasmodial immunity., *Science (80-.).* **335**, 856–9 (2012).

221. H. Diaz-Albiter, M. R. V Sant'Anna, F. A. Genta, R. J. Dillon, Reactive oxygen species-mediated immunity against Leishmania mexicana and Serratia marcescens in the sand phlebotomine fly Lutzomyia longipalpis., *J. Biol. Chem.* **287**, 23995–4003 (2012).

222. E. L. Telleria *et al.*, Caspar-like gene depletion reduces Leishmania infection in sand fly host Lutzomyia longipalpis., *J. Biol. Chem.* **287**, 12985–93 (2012).

223. N. Boulanger *et al.*, Characterization of a Defensin from the Sand Fly Phlebotomus duboscqi Induced by Challenge with Bacteria or the Protozoan Parasite Leishmania major, *Society* **72**, 7140–7146 (2004).

224. K. Blackburn, K. R. Wallbanks, D. H. Molyneux, D. R. Lavin, S. Winstanley, The peritrophic membrane of the female sandfly Phlebotomus papatasi, *Ann. Trop. Med. Parasitol.* **82**, 613–619 (1988).

225. W. Peters, *Peritrophic membranes.* (Springer-Verlag, 1992; http://www.cabdirect.org/abstracts/19920511645.html).

226. M. J. Lehane, Peritrophic matrix structure and function., *Annu. Rev. Entomol.* **42**, 525–50 (1997).

227. M. Jacobs-Lorena, O. MaungMaung, B. Beaty, W. Marquardt, The peritrophic matrix of insects., *Biol. Dis. vectors.* , 318–332 (1996).

228. V. Pascoa *et al.*, Aedes aegypti peritrophic matrix and its interaction with heme during blood digestion., *Insect Biochem. Mol. Biol.* **32**, 517–23 (2002).

229. L. L. Walters, K. P. Irons, G. B. Modi, R. B. Tesh, Refractory barriers in the sand fly Phlebotomus papatasi (Diptera: Psychodidae) to infection with Leishmania panamensis., *Am. Soc. Trop. Med. Hyg.* **46**, 211–28 (1992).

230. M. E. Rogers *et al.*, Leishmania chitinase facilitates colonization of sand fly vectors and enhances transmission to mice., *Cell. Microbiol.* **10**, 1363–72 (2008).

231. Y. Schlein, R. L. Jacobson, Haemoglobin inhibits the development of infective promastigotes and chitinase secretion in Leishmania major cultures., *Parasitology* **109**, 23–8 (1994).

232. Y. Schlein, R. L. Jacobson, J. Shlomai, Chitinase secreted by Leishmania functions in the sandfly vector., *Proc. R. Soc. London. Ser. B, Biol. Sci.* **245**, 121–6 (1991).

233. M. Shakarian, D. M. Dwyer, Pathogenic leishmania secrete antigenically related chitinases which are encoded by a highly conserved gene locus., *Exp. Parasitol.* **94**, 238–42 (2000).

234. M. Ramalho-Ortigão *et al.*, Characterization of a blood activated chitinolytic system in the midgut of the sand fly vectors Lutzomyia longipalpis and Phlebotomus papatasi., *Insect Mol. Biol.* **14**, 703–12 (2005).

235. R. Killick-Kendrick, Biology of Leishmania in phlebotomine sandflies, *Biol. Kinetoplastida* **2**, 395–460 (1979).

236. P. Volf, A. Kiewegová, M. Svobodová, Sandfly midgut lectin: effect of galactosamine on Leishmania major infections., *Med. Vet. Entomol.* **12**, 151–4 (1998).

237. R. J. Dillon, R. P. Lane, Detection of Leishmania lipophosphoglycan binding proteins in the gut of the sandfly vector, *Parasitology* **118**, 27–32 (1999).

238. S. Kamhawi *et al.*, A role for insect galectins in parasite survival., *Cell* **119**, 329–41 (2004).

239. P. Volf, J. Peckova, Sand flies and Leishmania: specific versus permissive vectors, *Trends Parasitol.* **23**, 91–92 (2007).

240. J. Myšková, M. Svobodová, S. M. Beverley, P. Volf, A lipophosphoglycan-independent development of Leishmania in permissive sand flies, *Microbes Infect.* **9**, 317–324 (2007).

241. R. Killick-Kendrick, K. R. Wallbanks, D. H. Molyneux, D. R. Lavin, The ultrastructure of Leishmania major in the foregut and proboscis of Phlebotomus papatasi, *Parasitol. Res.* **74**, 586–590 (1988).

242. R. Vaidyanathan, Leishmania parasites (Kinetoplastida: Trypanosomatidae) reversibly inhibit visceral muscle contractions in hemimetabolous and holometabolous insects., *J. Invertebr. Pathol.* **87**, 123–8 (2004).

243. R. Vaidyanathan, Reversible inhibition of contractions of mammalian cardiomyocytes and of smooth muscle by the protistan parasite Leishmania major, *Vet. Parasitol.* **134**, 53–60 (2005).

244. G. M. Coast, S. G. Webster, *Recent advances in arthropod endocrinology* G. M. Coast, S. G. Webster, Eds. (Cambridge Univ Press, Cambridge, 1998), pp. 227–354.

245. M. Ramalho-Ortigão, S. Kamhawi, E. D. Rowton, J. M. C. Ribeiro, J. G. Valenzuela, Cloning and characterization of trypsin- and chymotrypsin-like proteases from the midgut of the sand fly vector Phlebotomus papatasi, *Insect Biochem. Mol. Biol.* **33**, 163–171 (2003).

246. M. E. Rogers, The role of leishmania proteophosphoglycans in sand fly transmission and infection of the Mammalian host., *Front. Microbiol.* **3**, 223 (2012).

247. T. Ilg *et al.*, Purification and structural characterization of a filamentous, mucin-like proteophosphoglycan secreted by Leishmania parasites., *J. Biol. Chem.* **271**, 21583–96 (1996).

248. Y.-D. Stierhof *et al.*, Filamentous proteophosphoglycan secreted by Leishmania promastigotes forms gel-like three-dimensional networks that obstruct the digestive tract of infected sandfly vectors., *Eur. J. Cell Biol.* **78**, 675–89 (1999).

249. Y.-D. Stierhof, T. Ilg, D. G. Russell, H. Hohenberg, P. Overath, Characterization of polymer release from the flagellar pocket of Leishmania mexicana promastigotes., *J. Cell Biol.* **125**, 321–31 (1994).

250. T. Ilg, Proteophosphoglycans of Leishmania, *Parasitol. Today* **16**, 489–497 (2000).

251. M. E. Rogers, P. A. Bates, Leishmania manipulation of sand fly feeding behavior results in enhanced transmission., *PLoS Pathog.* **3**, e91 (2007).

252. R. Beach, G. Kiilu, J. Leeuwenburg, Modification of sand fly biting behavior by Leishmania leads to increased parasite transmission., *Am. Soc. Trop. Med. Hyg.* **34**, 278–82 (1985).

253. C. Maia, V. Šeblová, J. Sádlová, J. Votýpka, P. Volf, Experimental Transmission of Leishmania infantum by Two Major Vectors: A Comparison between a Viscerotropic and a Dermotropic Strain., *PLoS Negl. Trop. Dis.* **5**, e1181 (2011).

254. M. E. Rogers *et al.*, Proteophosophoglycans Regurgitated by Leishmania-Infected Sand Flies Target the L-Arginine Metabolism of Host Macrophages to Promote Parasite Survival, *PLoS Pathog.* **5**, e1000555 (2009).

255. M. E. Rogers, K. Corware, I. Müller, P. A. Bates, Leishmania infantum proteophosphoglycans regurgitated by the bite of its natural sand fly vector, Lutzomyia longipalpis, promote parasite establishment in mouse skin and skin-distant tissues., *Microbes Infect.* **12**, 875–9 (2010).

256. M. Shakarian, D. M. Dwyer, The Ld Cht1 gene encodes the secretory chitinase of the human pathogen Leishmania donovani., *Gene* **208**, 315–22 (1998).

257. P. Volf, A. Kiewegová, A. Nemec, Bacterial colonisation in the gut of Phlebotomus duboseqi (Diptera: Psychodidae): transtadial passage and the role of female diet., *Folia Parasitol. (Praha).* **49**, 73–7 (2002).

258. R. J. Dillon, E. el Kordy, M. Shehata, R. P. Lane, The prevalence of a microbiota in the digestive tract of Phlebotomus papatasi., *Ann. Trop. Med. Parasitol.* **90**, 669–73 (1996).

259. S. M. Oliveira *et al.*, [Prevalence of microbiota in the digestive tract of wild females of Lutzomyia longipalpis Lutz & Neiva, 1912) (Diptera: Psychodidae)]., *Rev. Soc. Bras. Med. Trop.* **33**, 319–22 (2000).

260. C. Gouveia, M. D. Asensi, V. Zahner, E. F. Rangel, S. M. P. de Oliveira, Study on the bacterial midgut microbiota associated to different Brazilian populations of Lutzomyia longipalpis (Lutz & Neiva) (Diptera: Psychodidae)., *Neotrop. Entomol.* **37**, 597–601 (2008).

261. H. Hillesland *et al.*, Identification of aerobic gut bacteria from the kala azar vector, Phlebotomus argentipes: a platform for potential paratransgenic manipulation of sand flies., *Am. Soc. Trop. Med. Hyg.* **79**, 881–6 (2008).

262. S. M. P. de Oliveira *et al.*, [Digestive tract microbiota in female Lutzomyia longipalpis (Lutz & Neiva, 1912) (Diptera: Psychodidae) from colonies feeding on blood meal and sucrose plus blood meal]., *Cad. Saude Publica* **17**, 229–32 (2001).

263. M. R. V Sant'Anna *et al.*, Investigation of the bacterial communities associated with females of Lutzomyia sand fly species from South America., *PLoS One* **7**, e42531 (2012).

264. R. J. Dillon, V. M. Dillon, The gut bacteria of insects: nonpathogenic interactions., *Annu. Rev. Entomol.* **49**, 71–92 (2004).

265. I. Hurwitz *et al.*, Paratransgenic control of vector borne diseases., *Int. J. Biol. Sci.* **7**, 1334–44 (2011).

266. I. Hurwitz, H. Hillesland, A. Fieck, P. Das, R. Durvasula, The paratransgenic sand fly: a platform for control of Leishmania transmission., *Parasit. Vectors* **4**, 82 (2011).

267. C. B. Pumpuni, M. S. Beier, J. P. Nataro, L. D. Guers, J. R. Davis, Plasmodium falciparum: inhibition of sporogonic development in Anopheles stephensi by gram-negative bacteria., *Exp. Parasitol.* **77**, 195–9 (1993).

268. L. Gonzalez-Ceron, F. Santillan, M. H. Rodriguez, D. Mendez, J. E. Hernandez-Avila, Bacteria in midguts of field-collected Anopheles albimanus block Plasmodium vivax sporogonic development., *J. Med. Entomol.* **40**, 371–4 (2003).

269. Y. Dong, F. Manfredini, G. Dimopoulos, Implication of the mosquito midgut microbiota in the defense against malaria parasites., *PLoS Pathog.* **5**, e1000423 (2009).

270. C. M. Cirimotich, J. L. Ramirez, G. Dimopoulos, Native microbiota shape insect vector competence for human pathogens., *Cell Host Microbe* **10**, 307–10 (2011).

271. S. Adler, O. Theodor, Transmission of disease agents by phlebotomine sand flies, *Annu. Rev. Entomol.* **2**, 203–226 (1957).

272. N. S. Akopyants *et al.*, Demonstration of genetic exchange during cyclical development of Leishmania in the sand fly vector, *Science (80-.).* **324**, 265–268 (2009).

273. M. Tibayrenc, F. Kjellberg, F. J. Ayala, A clonal theory of parasitic protozoa: the population structures of Entamoeba, Giardia, Leishmania, Naegleria, Plasmodium, Trichomonas, and Trypanosoma and their medical and taxonomical consequences., *PNAS* **87**, 2414–8 (1990).

274. K. Victoir, J.-C. Dujardin, How to succeed in parasitic life without sex? Asking Leishmania., *Trends Parasitol.* **18**, 81–5 (2002).

275. M. Y. Youssef, M. M. Eissa, S. T. el Mansoury, Evidence of sexual reproduction in the protozoan parasite Leishmania of the Old World., *J. Egypt. Soc. Parasitol.* **27**, 651–7 (1997).

276. S. Odiwuor, S. De Doncker, I. Maes, J.-C. Dujardin, G. Van der Auwera, Natural Leishmania donovani/Leishmania aethiopica hybrids identified from Ethiopia., *Infect. Genet. Evol.* **11**, 2113–8 (2011).

277. A. A. Belli, M. A. Miles, J. M. Kelly, A putative Leishmania panamensis/Leishmania braziliensis hybrid is a causative agent of human cutaneous leishmaniasis in Nicaragua., *Parasitology* **109 (Pt 4**, 435–42 (1994).

278. A.-L. Bañuls *et al.*, Evidence for hybridization by multilocus enzyme electrophoresis and random amplified polymorphic DNA between Leishmania braziliensis and Leishmania panamensis/guyanensis in Ecuador, *J. Eukaryot. Microbiol.* **44**, 408–411 (1997).

279. D. Nolder, N. Roncal, C. R. Davies, A. Llanos-Cuentas, M. A. Miles, Multiple hybrid genotypes of Leishmania (viannia) in a focus of mucocutaneous Leishmaniasis., *Am. Soc. Trop. Med. Hyg.* **76**, 573–8 (2007).

280. R. Lainson, J. J. Shaw, in *Topley & Wilson's Microbiology and Microbial Infections: Parasitology*, F. E. G. Cox, D. Waklin, S. H. Gillespie, D. D. Despommier, Eds. (Hodder Arnold ASM Press, London, 2005), pp. 313–49.

281. V. Rougeron *et al.*, Reproductive strategies and population structure in Leishmania: substantial amount of sex in Leishmania Viannia guyanensis., *Mol. Ecol.* **20**, 3116–27 (2011).

282. M. Z. Alam *et al.*, Multilocus microsatellite typing (MLMT) reveals genetic homogeneity of Leishmania donovani strains in the Indian subcontinent., *Infect. Genet. Evol.* **9**, 24–31 (2009).

283. R. Killick-Kendrick, D. H. Molyneux, M. Hommel, A. J. Leaney, E. S. Robertson, Leishmania in phlebotomid sandflies. V. The nature and significance of infections of the pylorus and ileum of the sandfly by leishmaniae of the braziliensis complex., *Proc. R. Soc. Lond. B. Biol. Sci.* **198**, 191–9 (1977).

284. R. Beach, G. Kiilu, L. D. Hendricks, C. N. Oster, J. Leeuwenburg, Cutaneous leishmaniasis in Kenya: transmission of Leishmania major to man by the bite of a naturally infected Phlebotomus duboscqi., *Trans. R. Soc. Trop. Med. Hyg.* **78**, 747–51 (1984).

285. M. Berriman *et al.*, The genome of the African trypanosome Trypanosoma brucei., *Science (80-.).* **309**, 416–22 (2005).

286. N. M. El-Sayed *et al.*, The genome sequence of Trypanosoma cruzi, etiologic agent of Chagas disease., *Science (80-.).* **309**, 409–15 (2005).

287. P. Wincker *et al.*, The Leishmania genome comprises 36 chromosomes conserved across widely divergent human pathogenic species., *Nucleic Acids Res.* **24**, 1688–94 (1996).

288. N. M. A. El-Sayed *et al.*, Comparative genomics of trypanosomatid parasitic protozoa., *Science (80-.).* **309**, 404–9 (2005).

289. J. Stevens, A. Rambaut, Evolutionary rate differences in trypanosomes., *Infect. Genet. Evol.* **1**, 143–50 (2001).

290. T. Downing *et al.*, Whole genome sequencing of multiple Leishmania donovani clinical isolates provides insights into population structure and mechanisms of drug resistance, *Genome Res.* **21**, 2143–2156 (2011).

291. D. F. Smith, C. S. Peacock, A. K. Cruz, Comparative genomics: from genotype to disease phenotype in the leishmaniases., *Int. J. Parasitol.* **37**, 1173–86 (2007).

292. W.-W. Zhang, G. Matlashewski, Screening Leishmania donovani-specific genes required for visceral infection., *Mol. Microbiol.* **77**, 505–17 (2010).

293. D. P. Depledge *et al.*, Comparative Expression Profiling of Leishmania: Modulation in Gene Expression between Species and in Different Host Genetic Backgrounds, *PLoS Negl. Trop. Dis.* **3**, e476 (2009).

294. D. P. Depledge *et al.*, Leishmania-specific surface antigens show sub-genus sequence variation and immune recognition, *PLoS Negl. Trop. Dis.* **4**, e829 (2010).

295. A. K. Cruz, R. G. Titus, S. M. Beverley, Plasticity in chromosome number and testing of essential genes in Leishmania by targeting., *PNAS* **90**, 1599–603 (1993).

296. S. Martínez-Calvillo, K. D. Stuart, P. J. Myler, Ploidy changes associated with disruption of two adjacent genes on Leishmania major chromosome 1., *Int. J. Parasitol.* **35**, 419–29 (2005).

297. P. Leprohon *et al.*, Gene expression modulation is associated with gene amplification, supernumerary chromosomes and chromosome loss in antimony-resistant Leishmania infantum., *Nucleic Acids Res.* **37**, 1387–99 (2009).

298. S. Besteiro, R. A. M. Williams, G. H. Coombs, J. C. Mottram, Protein turnover and differentiation in Leishmania., *Int. J. Parasitol.* **37**, 1063–75 (2007).

299. D. A. Campbell, S. Thomas, N. R. Sturm, Transcription in kinetoplastid protozoa: why be normal?, *Microbes Infect.* **5**, 1231–1240 (2003).

300. S. Martínez-Calvillo *et al.*, Transcription of Leishmania major Friedlin Chromosome 1 Initiates in Both Directions within a Single Region, *Mol. Cell* **11**, 1291–1299 (2003).

301. R. Evers *et al.*, Trypanosoma brucei contains two RNA polymerase II largest subunit genes with an altered C-terminal domain., *Cell* **56**, 585–97 (1989).

302. R. M. Saito, M. G. Elgort, D. A. Campbell, A conserved upstream element is essential for transcription of the Leishmania tarentolae mini-exon gene., *EMBO J.* **13**, 5460–9 (1994).

303. P. Respuela, M. Ferella, A. Rada-Iglesias, L. Aslund, Histone acetylation and methylation at sites initiating divergent polycistronic transcription in Trypanosoma cruzi., *J. Biol. Chem.* **283**, 15884–92 (2008).

304. R. Anish, M. B. Hossain, R. H. Jacobson, S. Takada, Characterization of transcription from TATA-less promoters: identification of a new core promoter element XCPE2 and analysis of factor requirements., *PLoS One* **4**, e5103 (2009).

305. N. R. Sturm, M. C. Yu, D. A. Campbell, Transcription termination and 3'-End processing of the spliced leader RNA in kinetoplastids., *Mol. Cell. Biol.* **19**, 1595–604 (1999).

306. S. Martínez-Calvillo, D. Nguyen, K. D. Stuart, P. J. Myler, Transcription initiation and termination on Leishmania major chromosome 3, *Eukaryot. Cell* **3**, 506 (2004).

307. T. N. Siegel *et al.*, Four histone variants mark the boundaries of polycistronic transcription units in Trypanosoma brucei., *Genes Dev.* **23**, 1063–76 (2009).

308. C. E. Clayton, Life without transcriptional control? From fly to man and back again, *EMBO J.* **21**, 1881–1888 (2002).

309. X. Liang, A. Haritan, S. Uliel, S. Michaeli, Trans and cis splicing in trypanosomatids: Mechanism, factors, and regulation, *Eukaryot. Cell* **2**, 830–840 (2003).

310. J. R. Zamudio, B. Mittra, D. A. Campbell, N. R. Sturm, Hypermethylated cap 4 maximizes Trypanosoma brucei translation., *Mol. Microbiol.* **72**, 1100–10 (2009).

311. T. N. Siegel, K. S. W. Tan, G. A. M. Cross, Systematic study of sequence motifs for RNA trans splicing in Trypanosoma brucei., *Mol. Cell. Biol.* **25**, 9586–94 (2005).

312. C. López-Estraño, C. Tschudi, E. Ullu, Exonic sequences in the 5' untranslated region of alpha-tubulin mRNA modulate trans splicing in Trypanosoma brucei., *Mol. Cell. Biol.* **18**, 4620–8 (1998).

313. J. H. LeBowitz, H. Q. Smith, L. Rusche, S. M. Beverley, Coupling of poly(A) site selection and trans-splicing in Leishmania., *Genes Dev.* **7**, 996–1007 (1993).

314. C. Benz, D. Nilsson, B. Andersson, C. Clayton, D. L. Guilbride, Messenger RNA processing sites in Trypanosoma brucei., *Mol. Biochem. Parasitol.* **143**, 125–34 (2005).

315. S. Haile, B. Papadopoulou, Developmental regulation of gene expression in trypanosomatid parasitic protozoa., *Curr. Opin. Microbiol.* **10**, 569–77 (2007).

316. S. M. Fernández-Moya, A. M. Estévez, Posttranscriptional control and the role of RNA-binding proteins in gene regulation in trypanosomatid protozoan parasites., *Wiley Interdiscip. Rev. RNA* **1**, 34–46 (2010).

317. J. D. Keene, RNA regulons: coordination of post-transcriptional events., *Nat. Rev. Genet.* **8**, 533–43 (2007).

318. C. Clayton, M. Shapira, Post-transcriptional regulation of gene expression in trypanosomes and leishmanias., *Mol. Biochem. Parasitol.* **156**, 93–101 (2007).

319. A. Hehl, E. Vassella, R. Braun, I. Roditi, A conserved stem-loop structure in the 3' untranslated region of procyclin mRNAs regulates expression in Trypanosoma brucei., *PNAS* **91**, 370–4 (1994).

320. K. K. Mishra, T. R. Holzer, L. L. Moore, J. H. LeBowitz, A negative regulatory element controls mRNA abundance of the Leishmania mexicana Paraflagellar rod gene PFR2., *Eukaryot. Cell* **2**, 1009–17 (2003).

321. S. Haile, A. Dupé, B. Papadopoulou, Deadenylation-independent stage-specific mRNA degradation in Leishmania., *Nucleic Acids Res.* **36**, 1634–44 (2008).

322. M. Engstler, M. Boshart, Cold shock and regulation of surface protein trafficking convey sensitization to inducers of stage differentiation in Trypanosoma brucei., *Genes Dev.* **18**, 2798–811 (2004).

323. G. Cohen-Freue, T. R. Holzer, J. D. Forney, W. R. McMaster, Global gene expression in Leishmania., *Int. J. Parasitol.* **37**, 1077–86 (2007).

324. M. Mayho, K. Fenn, P. Craddy, S. Crosthwaite, K. Matthews, Post-transcriptional control of nuclear-encoded cytochrome oxidase subunits in Trypanosoma brucei: evidence for genome-wide conservation of life-cycle stage-specific regulatory elements., *Nucleic Acids Res.* **34**, 5312–24 (2006).

325. M. David *et al.*, Preferential translation of Hsp83 in Leishmania requires a thermosensitive polypyrimidine-rich element in the 3' UTR and involves scanning of the 5' UTR., *RNA* **16**, 364–74 (2010).

326. M. Parsons, E. A. Worthey, P. N. Ward, J. C. Mottram, Comparative analysis of the kinomes of three pathogenic trypanosomatids: Leishmania major, Trypanosoma brucei and Trypanosoma cruzi., *BMC Genomics* **6**, 127 (2005).

327. T. R. Holzer, W. R. McMaster, J. D. Forney, Expression profiling by whole-genome interspecies microarray hybridization reveals differential gene expression in procyclic promastigotes, lesion-derived amastigotes, and axenic amastigotes in Leishmania mexicana., *Mol. Biochem. Parasitol.* **146**, 198–218 (2006).

328. K. Leifso, G. Cohen-Freue, N. Dogra, A. Murray, W. R. McMaster, Genomic and proteomic expression analysis of Leishmania promastigote and amastigote life stages: the Leishmania genome is constitutively expressed., *Mol. Biochem. Parasitol.* **152**, 35–46 (2007).

329. R. M. R. Coulson, D. F. Smith, Isolation of genes showing increased or unique expression in the infective promastigotes of Leishmania major., *Mol. Biochem. Parasitol.* **40**, 63–75 (1990).

330. H. M. Flinn, D. F. Smith, Genomic organisation and expression of a differentially-regulated gene family from Leishmania major, *Nucleic Acids Res.* **20**, 755–762 (1992).

331. T. M. Alce *et al.*, Expression of hydrophilic surface proteins in infective stages of Leishmania donovani, *Mol. Biochem. Parasitol.* **102**, 191–196 (1999).

332. P. G. McKean, P. W. Denny, E. Knuepfer, J. K. Keen, D. F. Smith, Phenotypic changes associated with deletion and overexpression of a stage-regulated gene family in Leishmania, *Cell. Microbiol.* **3**, 511–523 (2001).

333. H. M. Flinn, D. Rangarajan, D. F. Smith, Expression of a hydrophilic surface protein in infective stages of Leishmania major, *Mol. Biochem. Parasitol.* **64**, 259–270 (1994).

334. P. G. McKean, R. Delahay, P. F. P. Pimenta, D. F. Smith, Characterisation of a second protein encoded by the differentially regulated LmcDNA16 gene family of Leishmania major., *Mol. Biochem. Parasitol.* **85**, 221–31 (1997).

335. P. W. Denny, S. Gokool, D. G. Russell, M. C. Field, D. F. Smith, Acylation-dependent Protein Export in Leishmania, *J. Biol. Chem.* **275**, 11017–11025 (2000).

336. L. M. MacLean *et al.*, Trafficking and release of Leishmania metacyclic HASPB on macrophage invasion., *Cell. Microbiol.* **14**, 740–61 (2012).

337. P. F. P. Pimenta, P. Pinto da Silva, D. Rangarajan, D. F. Smith, D. L. Sacks, Leishmania major: association of the differentially expressed gene B protein and the

surface lipophosphogylcan as revealed by membrane capping, *Exp. Parasitol.* **79**, 468–479 (1994).

338. D. Rangarajan, S. Gokool, M. V McCrossan, D. F. Smith, The gene B protein localises to the surface of Leishmania major parasites in the absence of metacyclic stage lipophosphoglycan., *J. Cell Sci.* **108 (Pt 1**, 3359–66 (1995).

339. M. D. Resh, Regulation of cellular signalling by fatty acid acylation and prenylation of signal transduction proteins., *Cell. Signal.* **8**, 403–12 (1996).

340. S. Tournaviti *et al.*, Reversible Phosphorylation as a Molecular Switch to Regulate Plasma Membrane Targeting of Acylated SH4 Domain Proteins, *Traffic* **10**, 1047–1060 (2009).

341. E. Knuepfer, Y.-D. Stierhof, P. G. McKean, D. F. Smith, P. G. M. C. Kean, Characterization of a differentially expressed protein that shows an unusual localization to intracellular membranes in Leishmania major., *Biochem. J.* **356**, 335–344 (2001).

342. F. Oliveira, R. C. Jochim, J. G. Valenzuela, S. Kamhawi, Sand flies, Leishmania and transcriptome-borne solutions, *Mol. Biol.* **58**, 1–5 (2009).

343. K. Kuchler, Unusual routes of protein secretion: the easy way out., *Trends Cell Biol.* **3**, 421–6 (1993).

344. A. E. Cleves, Protein transport: The nonclassical ins and outs, *Curr. Biol.* **7**, R318–R320 (1997).

345. W. Nickel, Unconventional secretory routes: direct protein export across the plasma membrane of mammalian cells., *Traffic* **6**, 607–14 (2005).

346. C. Stegmayer *et al.*, Direct transport across the plasma membrane of mammalian cells of Leishmania HASPB as revealed by a CHO export mutant., *J. Cell Sci.* **118**, 517–27 (2005).

347. S. J. Parsons, J. T. Parsons, Src family kinases, key regulators of signal transduction., *Oncogene* **23**, 7906–9 (2004).

348. B. Moore *et al.*, Structural basis of molecular recognition of the Leishmania small hydrophilic endoplasmic reticulum-associated protein (SHERP) at membrane surfaces, *J. Biol. Chem.* **286**, 9246–9256 (2011).

349. S. Besteiro, R. A. M. Williams, L. S. Morrison, G. H. Coombs, J. C. Mottram, Endosome sorting and autophagy are essential for differentiation and virulence of Leishmania major., *J. Biol. Chem.* **281**, 11384–96 (2006).

350. A. Brennand *et al.*, Autophagy in parasitic protists: unique features and drug targets., *Mol. Biochem. Parasitol.* **177**, 83–99 (2011).

351. S. Rozen, H. Skaletsky, Primer3 on the WWW for general users and for biologist programmers, *Methods Mol Biol* **132**, 365–386 (2000).

352. A. Untergasser *et al.*, Primer3Plus, an enhanced web interface to Primer3., *Nucleic Acids Res.* **35**, W71–4 (2007).

353. S. F. Altschul, W. Gish, W. Miller, E. W. Myers, D. L. Lipman, Basic local alignment search tool, *J. Mol. Biol.* **215**, 403–410 (1990).

354. J. D. Thompson, T. J. Gibson, D. G. Higgins, Multiple sequence alignment using ClustalW and ClustalX., *Curr. Protoc. Bioinforma.* **Chapter 2**, Unit 2.3 (2002).

355. M. A. Larkin *et al.*, Clustal W and Clustal X version 2.0., *Bioinformatics* **23**, 2947–8 (2007).

356. X. Huang, A. Madan, CAP3: A DNA Sequence Assembly Program, *Genome Res.* **9**, 868–877 (1999).

357. G. Lu, E. Moriyama, Vector NTI, a balanced all-in-one sequence analysis suite, *Brief. Bioinform.* **5**, 378–388 (2004).

358. M. Heiman, Webcutter 2.0 (1997) (available at http://rna.lundberg.gu.se/cutter2/).

359. M. D. Abràmoff, P. J. Magalhães, S. J. Ram, Image processing with ImageJ, *Biophotonics Int.* **11**, 36–42 (2004).

360. J. Singh, Mendeley: A free research management tool for desktop and web., *J. Pharmacol. Pharmacother.* **1**, 62–3 (2010).

361. R. P. da Silva, D. L. Sacks, Metacyclogenesis is a major determinant of Leishmania promastigote virolence and attenuation, *Infect. Immun.* **55**, 2802–2806 (1987).

362. D. Paape, T. Aebischer, Contribution of proteomics of Leishmania spp. to the understanding of differentiation, drug resistance mechanisms, vaccine and drug development., *J. Proteomics* **74**, 1614–1624 (2011).

363. P. Chakraborty, S. Sturgill-Koszycki, D. G. Russell, Isolation and characterization of pathogen-containing phagosomes., *Methods Cell Biol.* **45**, 261–76 (1994).

364. J. A. Brannigan *et al.*, N-myristoyltransferase from Leishmania donovani: structural and functional characterisation of a potential drug target for visceral leishmaniasis., *J. Mol. Biol.* **396**, 985–99 (2010).

365. L. L. Walters, K. P. Irons, H. Guzman, R. B. Tesh, Formation and composition of the peritrophic membrane in the sand fly, Phlebotomus perniciosus (Diptera: Psychodidae), *J. Med. Entomol.* **30**, 179–198 (1993).

366. J. S. Oliveira, M. N. Melo, N. F. Gontijo, A sensitive method for assaying chemotaxic responses of Leishmania promastigotes., *Exp. Parasitol.* **96**, 187–9 (2000).

367. G. Leslie, M. Barrett, R. Burchmore, Leishmania mexicana: promastigotes migrate through osmotic gradients, *Exp. Parasitol.* **102**, 117–120 (2002).

368. J. Sambrook, D. W. Russell, *Molecular Cloning: A Laboratory Manual Vol.3* (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, ed. 3rd, 2001).

369. H. P. Price *et al.*, Myristoyl-CoA:protein N-myristoyltransferase, an essential enzyme and potential drug target in kinetoplastid parasites., *J. Biol. Chem.* **278**, 7206–14 (2003).

370. P. Volf, V. Volfova, Establishment and maintenance of sand fly colonies., *J. vector Ecol.* **36 Suppl 1**, S1–9 (2011).

371. I. Benkova, P. Volf, Effect of temperature on metabolism of Phlebotomus papatasi (Diptera: Psychodidae)., *J. Med. Entomol.* **44**, 150–4 (2007).

372. G. Madulo-Leblond, R. Killick-Kendrick, M. Killick-Kendrick, B. Pesson, [Comparison between Phlebotomus duboscqi Neveu-Lemaire, 1906, and Phlebotomus papatasi (Scopoli, 1786): morphological and isoenzymatic studies]., *Parassitologia* **33 Suppl**, 387–91 (1991).

373. J. Myšková, J. Votýpka, P. Volf, Leishmania in sand flies: comparison of quantitative polymerase chain reaction with other techniques to determine the intensity of infection., *J. Med. Entomol.* **45**, 133–8 (2008).

374. H. Price, C. Panethymitaki, D. Goulding, D. F. Smith, Functional analysis of TbARL1, an N-myristoylated Golgi protein essential for viability in bloodstream trypanosomes, *J. Cell Sci.* **118**, 831–841 (2005).

375. D. E. Dobson *et al.*, Leishmania major Survival in Selective Phlebotomus papatasi Sand Fly Vector Requires a Specific SCG-Encoded Lipophosphoglycan Galactosylation Pattern., *PLoS Pathog.* **6**, e1001185 (2010).

376. G. Dick, L. K. Akslen-Hoel, F. Grøndahl, I. Kjos, K. Prydz, Proteoglycan synthesis and Golgi organization in polarized epithelial cells., *J. Histochem. Cytochem.* **60**, 926–35 (2012).

377. S. Stäger *et al.*, Natural antibodies and complement are endogenous adjuvants for vaccine-induced CD8+ T-cell responses., *Nat. Med.* **9**, 1287–92 (2003).

378. G.-H. Sun-Wada, Y. Wada, M. Futai, Diverse and essential roles of mammalian vacuolar-type proton pump ATPase: toward the physiological understanding of inside acidic compartments., *Biochim. Biophys. Acta* **1658**, 106–14 (2004).

379. E. Vassella, B. Reuner, B. Yutzy, M. Boshart, Differentiation of African trypanosomes is controlled by a density sensing mechanism which signals cell cycle arrest via the cAMP pathway., *J. Cell Sci.* **110 (Pt 2**, 2661–71 (1997).

380. D. A. Baker, Malaria gametocytogenesis., *Mol. Biochem. Parasitol.* **172**, 57–65 (2010).

381. P. MacGregor, B. Szöőr, N. J. Savill, K. R. Matthews, Trypanosomal immune evasion, chronicity and transmission: an elegant balancing act., *Nat. Rev. Microbiol.* **10**, 431–8 (2012).

382. R. Brun, M. Schönenberger, Stimulating effect of citrate and cis-Aconitate on the transformation of Trypanosoma brucei bloodstream forms to procyclic forms in vitro., *Z. Parasitenkd.* **66**, 17–24 (1981).

383. S. Dean, R. Marchetti, K. Kirk, K. R. Matthews, A surface transporter family conveys the trypanosome differentiation signal, *Nature* **459**, 213–217 (2009).

384. J. M. Boitz, B. Ullman, A. Jardim, N. S. Carter, Purine salvage in Leishmania: complex or simple by design?, *Trends Parasitol.* **28**, 345–52 (2012).

385. C. M. Turner, The rate of antigenic variation in fly-transmitted and syringe-passaged infections of Trypanosoma brucei., *FEMS Microbiol. Lett.* **153**, 227–31 (1997).

386. C. M. Turner, J. D. Barry, High frequency of antigenic variation in Trypanosoma brucei rhodesiense infections., *Parasitology* **99 Pt 1**, 67–75 (1989).

387. L. He, G. J. Hannon, MicroRNAs: small RNAs with a big role in gene regulation., *Nat. Rev. Genet.* **5**, 522–31 (2004).

388. K. T. Militello, P. Refour, C. A. Comeaux, M. T. Duraisingh, Antisense RNA and RNAi in protozoan parasites: working hard or hardly working?, *Mol. Biochem. Parasitol.* **157**, 117–26 (2008).

389. Q. Liu, W. Tuo, H. Gao, X.-Q. Zhu, MicroRNAs of parasites: current status and future perspectives., *Parasitol. Res.* **107**, 501–7 (2010).

390. L.-F. Lye *et al.*, Retention and loss of RNA interference pathways in trypanosomatid protozoans., *PLoS Pathog.* **6**, e1001161 (2010).

391. V. D. Atayde *et al.*, The structure and repertoire of small interfering RNAs in Leishmania (Viannia) braziliensis reveal diversification in the trypanosomatid RNAi pathway., *Mol. Microbiol.* **87**, 580–93 (2013).

392. M. E. Rogers, O. V. O. Sizova, M. A. J. Ferguson, A. V. Nikolaev, P. A. Bates, Synthetic glycovaccine protects against the bite of Leishmania-infected sand flies, *J. Infect. Dis.* **194**, 512–518 (2006).

393. N. C. Peters *et al.*, Vector Transmission of Leishmania Abrogates Vaccine-Induced Protective Immunity, *PLoS Pathog.* **5**, e1000484 (2009).

394. J. G. Valenzuela *et al.*, Toward a defined anti-Leishmania vaccine targeting vector antigens: characterization of a protective salivary protein., *J. Exp. Med.* **194**, 331–42 (2001).

395. L. W. Stamper *et al.*, Infection Parameters in the Sand Fly Vector That Predict Transmission of Leishmania major., *PLoS Negl. Trop. Dis.* **5**, e1288 (2011).

396. F. Tripet, S. Clegg, D.-E. Elnaiem, R. D. Ward, Cooperative blood-feeding and the function and implications of feeding aggregations in the sand fly, Lutzomyia longipalpis (Diptera: Psychodidae)., *PLoS Negl. Trop. Dis.* **3**, e503 (2009).

397. P. A. Bates, Leishmania sand fly interaction: progress and challenges., *Curr. Opin. Microbiol.* **11**, 340–4 (2008).

398. D. Horn, G. A. . Cross, Analysis of Trypanosoma brucei vsg expression site switching in vitro, *Mol. Biochem. Parasitol.* **84**, 189–201 (1997).