Copyright is owned by the Author of the thesis. Permission is given for a copy to be downloaded by an individual for the purpose of research and private study only. The thesis may not be reproduced elsewhere without the permission of the Author.

Elimination of *Leishmania* RNA virus from *Leishmania* spp. via expression of homologous viral capsid proteins

A thesis presented in partial fulfilment of the requirements for the degree of

Master of Science

in

Biological Sciences

at

Massey University, Palmerston North,

New Zealand.

Donnamae Saavedra Klocek

2023

Somewhere, something incredible is waiting to be known.

~Carl Sagan

ACKNOWLEDGEMENTS

I would like to thank Professor Vyacheslav Yurchenko for introducing me to the fascinating world of trypanosomatid biology, for welcoming me into his lab despite bureaucratic impediments, and for being a wellspring of knowledge and advice for all things trypanosomatid-related.

I'm incredibly grateful to Dr. Dragana Gagic for accommodating my special circumstances, for taking the time to communicate with me at unusual hours, and for the invaluable advice and guidance in completing the requirements for this degree.

I would like to thank Alexandra Zakharova, Andreu Saura, and Anastasia Bykova for helping me get reacquainted with standard molecular biology techniques, and for introducing me to new techniques I haven't used before.

I'm grateful to my husband, kids, extended family, and in-laws for the boundless support and patience they've shown throughout this journey.

ABSTRACT

Leishmaniasis is a neglected tropical disease caused by the protozoan parasite *Leishmania* with complex pathology ranging from skin ulcers to fatal organ damage. There are no human vaccines for leishmaniasis, and resistance to the limited arsenal of drugs used for antileishmanial therapy emerges frequently among *Leishmania* parasites. Some strains of *Leishmania* have been found to harbour *Leishmania* RNA virus (LRV), which has been associated with higher rates of treatment failure, hyperinflammation, and relapse. Investigating the relationship between LRV and *Leishmania* may potentially reveal new targets or strategies for antileishmanial treatment.

LRV has a dsRNA genome encoding two genes – capsid protein (CP) and RNA-dependent RNA polymerase (RDRP). Previous studies on the closely related L-A virus in *Saccharomyces cerevisiae* demonstrated that overexpression of full or truncated versions of the viral CP led to elimination of the virus from the host. This study aimed to test the specificity of CP-induced viral elimination. CPs from 5 strains of LRV – LRV1-4, LRV1-Lg2014, LRV1-LbrLEM2700, LRV1-1, and LRV2 – were integrated into the genomic SSU loci of two LRV-harbouring *Leishmania* strains, *L. guyanensis* M4147 and *L. major* T44g, bearing LRV1-4 and LRV2, respectively. The expression of exogenous CPs was confirmed via Western blotting and RT-qPCR. The effects of CP overexpression on viral copy numbers of resident LRVs were evaluated via RT-qPCR. All 5 CPs significantly reduced or eliminated LRV1-4 from *L. guyanensis*, whereas only LRV2 CP diminished LRV2 copy number in *L. major*. Thus, *L. guyanensis* and *L. major* appear to have disparate relationships with, and dissimilar mechanisms for maintenance of, their resident LRV.

ABBREVIATIONS

°C	degrees Celsius
+ssRNA	positive-sense ssRNA
-ssRNA	negative-sense ssRNA
2CMA	2′C-methyladenosine triphosphate
3′	3 prime
5′	5 prime
α-	anti
µg	microgram
µL	microlitre
AGE	agarose gel electrophoresis
Amp^R	ampicillin resistance
bp	base pairs
BCA	bicinchoninic acid
BSA	bovine serum albumin
cDNA	complementary DNA
CCL	chemokine (C-C motif) ligand
CL	cutaneous leishmaniasis
CP	capsid protein
DCL	disseminated cutaneous leishmaniasis
ddH_2O	deionized distilled water
DMSO	dimethyl sulphoxide
DNA	deoxyribonucleic acid
DNase	deoxyribonuclease
dNTP	dinucleotide triphosphate
dsRNA	double-stranded RNA
EDTA	ethylenediaminetetraacetic acid
EV	extracellular vesicle
FBS	foetal bovine serum
g	gram

g	relative centrifugal force
HA	hemagglutinin
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HR	homologous recombination
IFN	interferon
IL	interleukin
kb	kilobase pairs
kDa	kilodalton
kmp11	kinetoplast membrane protein-11
L	litre
LB	Luria-Bertani bacteriological media
LRV	*Leishmania* RNA virus
mA	milliampere
MAK3	maintenance of killer 3
MCL	mucocutaneous leishmaniasis
mg	milligrams
min	minutes
mL	millilitre
mM	millimolar
mRNA	messenger RNA
MW	molecular weight
ng	nanogram
NLRP3	NLR Family Pyrin Domain Containing 3
nt	nucleotide
ORF	open reading frame
PAGE	polyacrylamide gel electrophoresis
PBS	phosphate buffered saline
PCR	polymerase chain reaction
Pen/Strep	penicillin + streptomycin solution
PKDL	post-kala-azar dermal leishmaniasis

PTM	post-translational modification
PTU	polycistronic transcription unit
PVDF	polyvinylidene difluoride
RBP	RNA binding protein
RDRP	RNA-dependent RNA polymerase
RF	ribosomal frameshift
RNA	ribonucleic acid
RNAi	RNA interference
RNase	ribonuclease
rRNA	ribosomal RNA
RT-qPCR	quantitative reverse transcription PCR
SDS	sodium dodecyl sulphate
SL	spliced leader
ssRNA	single-stranded RNA
SSU	small subunit ribosomal RNA
TAE	tris acetic acid EDTA
TBS	Tris buffered saline
TBST	TBS-Tween
TE	tris EDTA buffer
TEMED	N,N,N',N'-Tetramethylethylenediamine
TLR3	Toll-like receptor 3
TNF	tumour necrosis factor
Tris	2-amino-2-hydromethyl-1,3-propanidol
UTR	untranslated region
V	volts
VL	visceral leishmaniasis
VLP	virus-like particle
v/v	volume per volume
w/v	weight per volume
WT	wild type

LIST OF FIGURES

Figure 1.1 Taxonomic classification of *Leishmania* 6

Figure 1.2 Life cycle of *Leishmania* 7

Figure 1.3 LRV genome and replication cycle 13

Figure 1.4 Model of how LRV CP overexpression eliminates LRV 20

Figure 2.1 Antibiotic selection in 24-well plates 26

Figure 3.1 Plasmid map of pLEXSY-neo2.1 35

Figure 3.2 Cloning the LRV2 CP gene 37

Figure 3.3 PCR verification of pLEXSY-neo2.1-LRV2_CP ligation products 38

Figure 3.4 Integration of HA-CP harbouring pLEXSY-neo2.1 derivative into *Leishmania* genomic SSU locus 41

Figure 3.5 Western blot analysis of expression of 5 LRV CPs in *L. guyanensis* and *L. major* 43

Figure 3.6 PCR verification of genomic integration of pLEXSY-neo2.1-LRV1-1_CP in three clonal lines 45

Figure 3.7 Transcript levels of genome-integrated, exogenous LRV CPs in *L. guyanensis* and *L. major* 46

Figure 3.8 Effect of expression of LRV CPs on LRV1-4 copy number in *L. guyanensis* and LRV2 copy number in *L. major* 48

LIST OF TABLES

Table 2.1 Sources of consumable materials used in this work **22**

Table 2.2 Manufacturers of equipment used in this work **23**

Table 2.3 PCR settings, standard **28**

Table 2.4 PCR settings, high fidelity **28**

Table 3.1 LRV strains of CP genes used in this study **39**

Table 3.2 Newly constructed parasite lines expressing HA-tagged LRV CP from genomic SSU locus **40**

Table 4.1 Identity between pairs of CPs **54**

TABLE OF CONTENTS

ABSTRACT .... **i**

ABBREVIATIONS .... **ii**

LIST OF FIGURES .... **v**

LIST OF TABLES .... **vi**

CHAPTER 1. INTRODUCTION

1.1 Leishmaniasis **1**

1.1.1 Clinical features of leishmaniasis **1**

1.1.2 Prevention and treatment of leishmaniasis **3**

1.2 *Leishmania* **5**

1.2.1 Taxonomy **5**

1.2.2 *Leishmania* life cycle **7**

1.2.3 Molecular biology of *Leishmania* **8**

1.2.3.1 Regulation of gene expression in *Leishmania* **8**

1.2.3.1.1 The unique transcription system of *Leishmania* **9**

1.2.3.1.2 Post-transcriptional control of gene expression **9**

1.2.3.1.3 Aneuploidy and genome plasticity **10**

1.2.3.1.4 Homologous recombination **11**

1.3. *Leishmania* RNA virus **12**

1.3.1 History and taxonomy **12**

1.3.2 LRV structure and genome organization **13**

1.3.3 The *Leishmania*-LRV relationship **15**

1.3.4 LRV as a factor affecting severity of leishmaniasis **16**

1.3.5 Previous attempts at eliminating LRV **17**

1.3.5.1 Chemical intervention **18**

1.3.5.2 CP Overexpression **18**

1.4 Research Outline **19**

1.4.1 Hypothesis **20**

1.4.2 Research aims **21**

CHAPTER 2. MATERIALS AND METHODS

2.1 Materials **22**
2.2 Methods **24**
2.2.1 *Leishmania* culture **24**
2.2.1.1 Passaging of *Leishmania* cultures **24**
2.2.1.2 Freezing of *Leishmania* cultures **24**
2.2.1.3 Thawing of *Leishmania* cultures **24**
2.2.1.4 Growth kinetics analysis **25**
2.2.1.5 Making clonal *Leishmania* lines **25**
2.2.1.6 Transfection **25**
2.2.1.6.1 Electroporation........ **25**
2.2.1.6.2 Antibiotic selection **26**
2.2.1.6.3 Transfection of *Leishmania* parasites with constructs for genomic integration into *Leishmania* 18S SSU rRNA locus **26**
2.2.2 Bacteria **27**
2.2.2.1 Transformation **27**
2.2.2.2 Plasmid production **27**
2.2.3 DNA **27**
2.2.3.1 Polymerase chain reaction (PCR) **27**
2.2.3.1.1 Standard PCR **27**
2.2.3.1.2 High fidelity PCR **28**
2.2.3.1.3 Colony PCR **28**
2.2.3.2 Genomic DNA extraction **29**
2.2.3.3 Agarose gel electrophoresis **29**
2.2.3.4 Restriction endonuclease digestion **30**
2.2.3.5 Ligation **30**
2.2.3.6 Preparation of samples for sequencing **30**
2.2.4 RNA **30**
2.2.4.1 RNA extraction **30**
2.2.4.2 DNase treatment **31**
2.2.4.3 Complementary DNA (cDNA) synthesis **31**
2.2.4.4 Quantitative PCR (qPCR) **31**
2.2.5 Protein **32**

2.2.5.1 Protein extraction via sonication **32**
2.2.5.2 Protein quantification **32**
2.2.5.3 Sodium dodecyl sulphate - polyacrylamide gel electrophoresis (SDS-PAGE) **33**
2.2.5.2 Western blotting **33**
2.2.6 Statistical analysis **34**

CHAPTER 3. RESULTS

3.1 Construction of transfection vectors for genomic integration of LRV CP genes **35**
3.1.1 Vector selection **35**
3.1.2 Construction of LRV CP expression vectors derived from pLEXSY-neo2.1 ..... **37**
3.2 Integration of five LRV CP gene variants into *L. guyanensis* and *L. major* genomic DNA **40**
3.3 Confirmation of expression of genome-integrated LRV CPs **42**
3.4 Transcript levels of genome-integrated exogenous LRV CPs **45**
3.5 Effect of expression of LRV CPs on viral copy number of resident LRVs **47**

CHAPTER 4. DISCUSSION

4.1 Integration of HA-tagged LRV CPs into *L. guyanensis* and *L. major* genomic DNA **50**
4.2 The HA-tagged LRV CPs were expressed at dissimilar levels **50**
4.3 Expression of all types of CPs led to elimination or copy number diminution of LRV1-4 in *L. guyanensis* **53**
4.4 Expression of LRV2 CP led to significant copy number reduction, but not elimination, of LRV2 from *L. major*, while other CPs had no effect **54**
4.5 *L. major* and *L. guyanensis* have dissimilar relationships with their resident LRVs **56**

CHAPTER 5. SUMMARY AND FUTURE DIRECTIONS

5.1 Summary **58**
5.2 Future directions **60**
5.2.1 Expressing HA-tagged LRV1-1 CP in *L. guyanensis* CUMC1 ..... **60**
5.2.2 Re-tagging of LRV1-1 CP **60**
5.2.3 Immunofluorescence microscopy to detect formation of heterogeneous capsids . **61**
5.2.4 Expression of CP-RDRP from the genomic SSU locus **61**
5.2.5 Incorporation of full LRV1-4 cDNA into the genomic SSU locus **62**

5.2.6 Testing the in vivo effect of LRV1-4 CP on the virulence of other LRV-bearing *Leishmania* strains **62**
5.2.7 Using HA-CP expressing *Leishmania* lines to determine the effects of various conditions on EV-mediated LRV release **63**
5.2.8 Identification of *Leishmania* host factors involved in LRV maintenance **63**

REFERENCES 65

APPENDIX A. PRIMERS 74

APPENDIX B. PLASMIDS 75

APPENDIX C. PUBLICATION 87

CHAPTER 1. INTRODUCTION

1.1 Leishmaniasis

Leishmaniasis is a neglected, complex, vector-borne tropical and subtropical disease caused by the protozoan parasite *Leishmania.* 12 million people are currently suffering from leishmaniasis, and an estimated 1.2 million new cases occur annually (WHO, 2022). 350 million people in 98 *Leishmania*-endemic countries within Central and South America, the Indian subcontinent, the Middle East, Mediterranean Europe, and Northeast Africa are at constant risk of infection (Gradoni, 2018; WORLD HEALTH ORGANIZATION, 2022). Leishmaniasis is the second leading cause of parasite-related deaths after malaria (Torres-Guerrero et al., 2017). Because the insect vectors and reservoir hosts of *Leishmania* thrive in hot, humid conditions, global prediction models suggest that the higher temperatures brought about by the escalating climate crisis will broaden the geographical distribution of *Leishmania* vectors and hosts; consequently, the epidemiological risk of leishmaniasis to humans is expected to increase (González et al., 2010; Koch et al., 2017).

1.1.1 Clinical features of leishmaniasis

Leishmaniasis has a broad range of pathological manifestations, and is of two major clinical forms:

1. Cutaneous leishmaniasis (CL) manifests as skin ulcerations that, while self-limiting and localised, leave disfiguring, permanent scars (Torres-Guerrero et al., 2017). While most cases of CL spontaneously resolve within a span of 3-18 months, the psychological impacts of the prominent and unsightly scars are substantial, especially

due to lifelong social stigmatisation of the affected individuals (Malvolti et al., 2021; Yanik et al., 2004).

2. Visceral leishmaniasis (VL) involves accumulation of *Leishmania* parasites in macrophage-rich visceral organs (Gradoni, 2018; Steverding, 2017). VL, also known as kala-azar, causes enlargement of the liver and spleen, leading to death in >95% of cases if untreated (Torres-Guerrero et al., 2017).

CL that the host fails to resolve can disseminate to secondary sites and trigger damaging, often chronic, hyperinflammatory responses (Ronet et al., 2011; Torres-Guerrero et al., 2017). CL that has spread to mucosal tissue is called mucocutaneous leishmaniasis (MCL). MCL is a severe, aggressive and debilitating form of the disease that involves destruction of the soft tissues of the mouth, nose, palate, and throat (Gradoni, 2018; Ronet et al., 2011). MCL responds poorly to standard antileishmanial treatments, and renders the patient susceptible to opportunistic infections by bacteria and fungi, further compounding the disease burden on the patient (Ronet et al., 2011; Torres-Guerrero et al., 2017). Other forms of CL include disseminated cutaneous leishmaniasis (DCL), wherein dozens to thousands of polymorphic skin lesions appear in 2 or more non-contiguous parts of the body, sometimes including the nasal mucosa(Machado et al., 2019). Another form is post-kala-azar dermal leishmaniasis (PKDL), a persistent form of CL suffered by people who have recovered from VL (Zijlstra et al., 2003). Sequelae of CL have high rates of treatment failure (Gradoni, 2018; Kaye & Scott, 2011; Steverding, 2017).

While the clinical variability among leishmaniases is poorly understood at the molecular level, intrinsic features of the parasite play a major role. For any given patient with leishmaniasis, the particular species of the causative *Leishmania* is the principal predictor of disease phenotype (Gradoni, 2018). Thirty of the 53 known species of *Leishmania* infect mammals, and of these, 21 are pathogenic to humans. Only three species belonging to the

L. donovani complex – *L. donovani*, *L. infantum*, and *L. chagasi* – are associated with VL (Steverding, 2017; Torres-Guerrero et al., 2017). MCL predominantly occurs in Central and South America, typically after infection with species belonging to the *L. braziliensis* complex – *L. braziliensis, L. guyanensis,* and *L. panamensis* (Ronet et al., 2011).

In addition to parasite-related factors, extrinsic modulators of the severity of leishmaniasis include interactions of the *Leishmania* parasite with its insect vector, variations in geography that influence vector growth and behaviour, complex differences in the immune responses of mammalian hosts, and coinfection of the mammalian host with other pathogens such as HIV (Gradoni, 2018; Zangger et al., 2013).

1.1.2 Prevention and treatment of leishmaniasis

Leishmaniasis is most common in underdeveloped regions of the world with poor nutrition, housing, and sanitation (Gradoni, 2018). Epidemics of leishmaniasis often break out in makeshift encampments of refugees displaced by war and conflict (Steverding, 2017). A recent example is the upsurge of leishmaniasis in Syria and the surrounding Middle Eastern countries in the last decade, which is a consequence of decimation of health infrastructures, population displacement, food insecurity, and deterioration of living conditions brought about by the Syrian civil war (Du et al., 2016; Steverding, 2017). Use of insecticides to curtail *Leishmania* transmission is contentious due to unknown, possibly negative, environmental and agricultural consequences (Fernández-Prada et al., 2019).

Four antileishmanial vaccines have been approved for use in dogs (Velez & Gállego, 2020). Vaccination of dogs in regions where leishmaniasis is endemic reduces transmission of *Leishmania* to humans, as dogs kept as domestic pets are a major reservoir host for *Leishmania* (Dantas-Torres, 2007). None of these vaccines have been approved

for human use because they do not elicit adequate immunity and cause adverse reactions in humans (Moafi et al., 2019).

There are currently no vaccines or other chemoprophylactic agents for leishmaniasis that are suitable for humans, despite decades of attempts at developing them (Olías-Molero et al., 2021). While a strong humoral response is induced in leishmaniasis, the antibodies produced are not protective and have instead been shown to exacerbate the disease (Miles et al., 2005). Activation of CD8$^+$ T cells likewise aids disease progression, as their cytolytic activities on infected cells release *Leishmania* amastigotes without harming them, thereby facilitating parasite spread and engulfment by new hosts (Malvolti et al., 2021; Scott & Novais, 2016). Nonetheless, in the vast majority of cases, individuals who have fully recovered from leishmaniasis are invulnerable to reinfection, which suggests that vaccination against leishmaniasis is not unachievable (Malvolti et al., 2021). This may require the development of strategies that activate CD4$^+$ effector and memory T cells while keeping the CD8$^+$ T cell pathway at bay (Scott & Novais, 2016).

As *Leishmania* parasites in humans predominantly reside within phagocytic immune cells, the current arsenal of antileishmanial treatments consists of chemotherapeutic agents targeting infected cells; these agents are derived from antimony, miltefosine, and amphotericin B (Fernández-Prada et al., 2019; Olías-Molero et al., 2021). These drugs come with a range of issues, the most significant being their toxicity to non-target cells, serious side effects, and numerous contraindications (Aronson et al., 2017; Olías-Molero et al., 2021). Furthermore, these drugs are costly and have limited availability, with some countries relying on a single type of antileishmanial drug (WHO, 2022). Because of the clinical heterogeneity of leishmaniases, antileishmanial therapy is most effective when tailored to individual cases in terms of drug type, dosage, and treatment duration (Aronson

et al., 2017). Limited availability of chemotherapeutic drugs in some regions of the world means antileishmanial therapies in these regions are often suboptimal.

The emergence of drug resistance among *Leishmania* is also an ongoing problem (Fernández-Prada et al., 2019). Drug resistance emerges, in part, from the administration of sub-therapeutic doses of antileishmanial drugs, which selects for drug-resistant parasites (Ponte-Sucre et al., 2017). Sub-therapeutic dosing can result from faulty drug administration by inexperienced medical workers, low patient compliance, or the use of drugs that have been stored improperly, as leishmaniasis is endemic in numerous hot and humid regions with suboptimal storage facilities (Aronson et al., 2017; Ponte-Sucre et al., 2017).

1.2 *Leishmania*

1.2.1 Taxonomy

The phylum Euglenozoa, a speciose group of mono- and bi-flagellated protists with a wide range of lifestyles, branches into three classes: Euglenida, Diplonemea, and Kinetoplastea (Kostygov, Karnkowska, et al., 2021). Members of the class Kinetoplastea, referred to as kinetoplastids, possess a single large mitochondrion known as the kinetoplast, located near the base of their flagella (d'Avila-Levy et al., 2015). Among the kinetoplastids is the family Trypanosomatidae, which includes medically important genera such as *Trypanosoma* and *Leishmania* (Figure 1.1) (Kostygov, Karnkowska, et al., 2021). The genus *Leishmania* is further divided into 4 subgenera: *Viannia, Leishmania, Sauroleishmania,* and *Mundinia* (Figure 1.1B) (Kostygov, Karnkowska, et al., 2021). Among these, the first two subgenera, *Leishmania (Leishmania)* and *Leishmania (Viannia)*, are the most medically significant to humans (Gradoni, 2018).

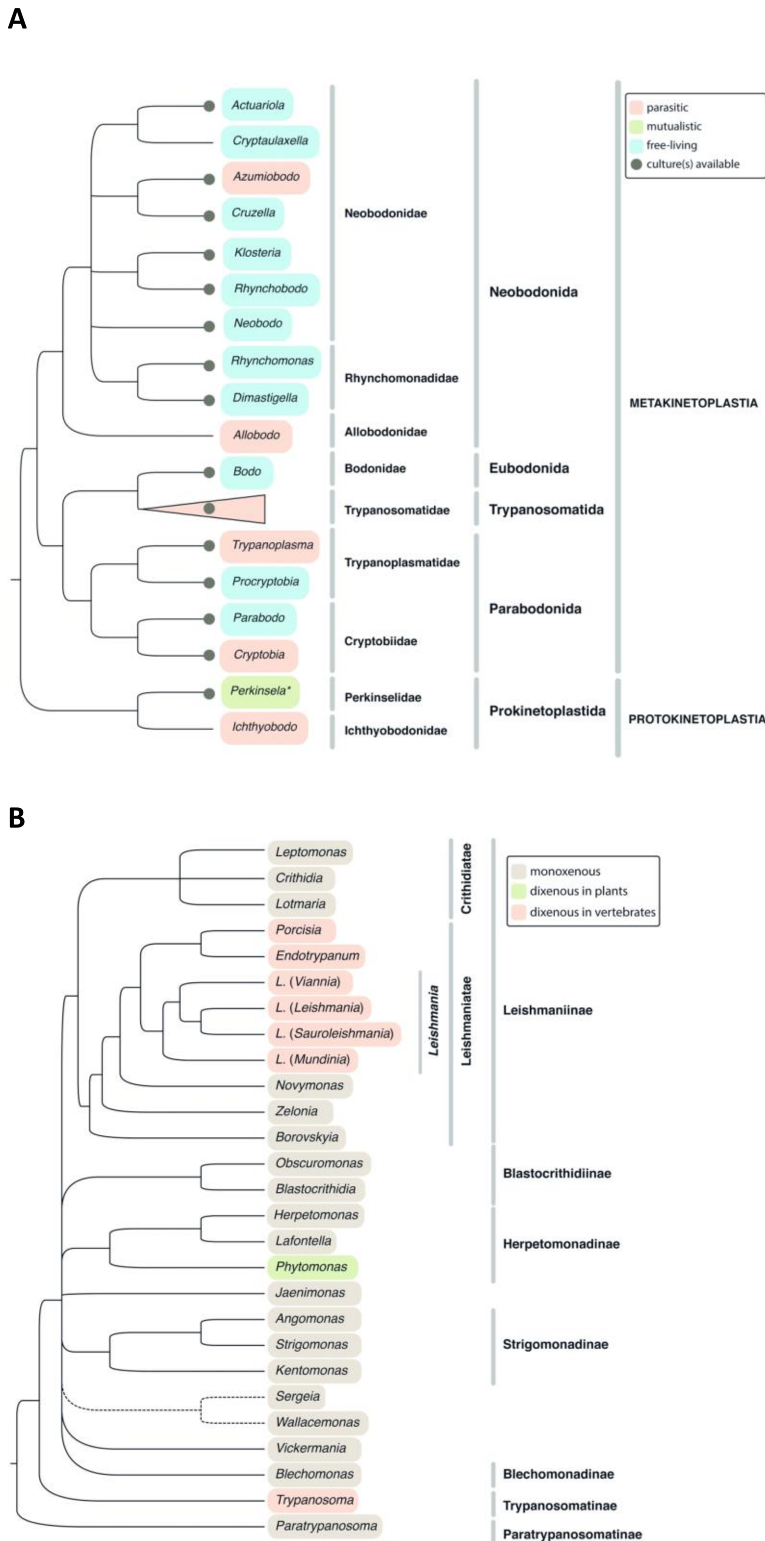


Figure 1.1 Taxonomic classification of *Leishmania*. **A** Phylogenetic tree of class Kinetoplastea constructed mostly using 18S rRNA gene. Genus *Leishmania* is classified under family Trypanosomatidae. **B** Phylogenetic tree of family Trypanosomatidae constructed mostly using 18S rRNA gene. *Leishmania* is comprised of four subgenera: *Viannia, Leishmania, Sauroleishmania,* and *Mundinia*. *Images reproduced without modification from: Kostygov, A. Y., Karnkowska, A., Votýpka, J., Tashyreva, D., Maciszewski, K., Yurchenko, V., & Lukeš, J. (2021). Euglenozoa: taxonomy, diversity and ecology, symbioses and viruses. Open Biology, 11(3).*

Leishmania (Leishmania) is mainly transmitted by sand flies of the genus *Phlebotomus* in the Old World, while *Leishmania (Viannia)* is mainly transmitted by *Lutzomyia* sand flies in the New World (Torres-Guerrero et al., 2017). Domestic and wild mammals from at least 8 mammalian orders serve as reservoir hosts for human-pathogenic species of *Leishmania* (Gradoni, 2018).

1.2.2 *Leishmania* life cycle

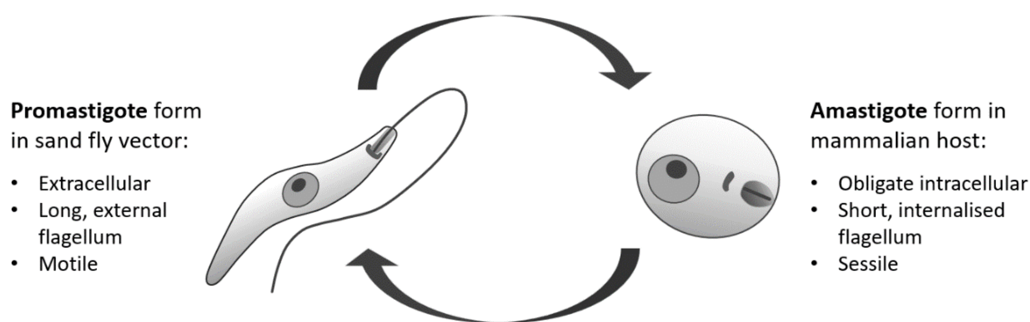


Figure 1.2 *Leishmania* life cycle. *Leishmania* cycle between two morphologically distinct forms in the sand fly and vertebrate host. *Image drawn by author using Microsoft PowerPoint.*

Leishmania are dixenous parasites – they alternate between two types of hosts (Figure 1.2). The flagellated, motile, extracellular, promastigote forms reside in the midgut of female phlebotomine sand flies (Gossage et al., 2003). They migrate to the vector's pharyngeal valve, from which they are deposited into the blood of a mammalian host when the sand fly takes a blood meal (Bates, 2007). These promastigotes are phagocytosed by local cells, primarily macrophages and other professional phagocytes that are recruited in response to the wound made by the sand fly's proboscis (Kaye & Scott, 2011). The shift to the harsh environment of phagolysosomes induces radical changes in *Leishmania* morphology and physiology: they differentiate into non-flagellated, non-motile, obligate intracellular amastigotes that are able to proliferate and thrive within the phagolysosomes (Handman & Bullen, 2002). When an uninfected sand fly draws blood from an infected mammalian

host, some *Leishmania*-harbouring macrophages are ingested in the process. The changes in temperature and pH in the sand fly's midgut induce differentiation of *Leishmania* from amastigotes to promastigotes, thus restarting the cycle. (Gradoni, 2018).

1.2.3 Molecular biology of *Leishmania*

Due to the absence of effective vaccines for leishmaniasis and the frequent emergence of drug resistance in *Leishmania*, significant effort has been made to understand their biology and elucidate the genetic factors and mechanisms that contribute to their pathogenicity, in order to find new antileishmanial drug targets or develop new schemes for curtailing disease progression (Cantanhêde et al., 2021). Decades of studies on *Leishmania* and other medically important trypanosomatids such as *Trypanosoma* – the etiological cause of African sleeping sickness and Chagas disease – have resulted in the establishment of a good *Leishmania* genetic toolbox.

Most studies on *Leishmania* are done on the promastigote form, which is fairly easy to maintain in host-free *in vitro* cultures. Strains of various *Leishmania* species have been maintained and studied *in vitro* for decades, with no apparent detriment to their fitness or infectivity (Roberts, 2011).

Leishmania have a number genetic features that are different from typical eukaryotes. Consequently, genetic manipulation of *Leishmania* usually necessitates amendments to standard molecular biology techniques.

1.2.3.1 Regulation of gene expression in *Leishmania*

As *Leishmania* are early-diverging eukaryotes, they lack most of the systems for transcriptional control of gene expression that the majority of eukaryotes evolved after the divergence (Curotto de Lafaille et al., 1992). Individual *Leishmania* genes do not have

canonical promoters or other *cis*-acting transcriptional regulatory elements, and the *Leishmania* genome does not encode *trans*-acting transcription factors which bind to these elements (Grünebast & Clos, 2020).

1.2.3.1.1 The unique transcription system of *Leishmania*

Leishmania have a remarkable transcription system that is a hybrid between the prokaryotic and eukaryotic systems. Like prokaryotes, *Leishmania* genes lack introns and are transcribed polycistronically (Curotto de Lafaille et al., 1992). But unlike prokaryotic operons, the genes in each polycistronic transcription unit (PTU) are not functionally related (Grünebast & Clos, 2020). Less than 5% of *Leishmania* genes are differentially transcribed between the promastigote and amastigote life cycle stages (Cohen-Freue et al., 2007). Thus, *Leishmania* transcription is constitutive: it is carried out at a fairly uniform rate genome-wide (Cohen-Freue et al., 2007; Karamysheva et al., 2020).

Despite lacking introns, *Leishmania* genes undergo *trans*-splicing, whereby a universal capped spliced leader (SL) RNA sequence is attached to the 5′-end of each gene (Curotto de Lafaille et al., 1992). The UTR between any two genes in a PTU contain both the polyadenylation signal for the upstream gene and the *trans*-splicing signal for the downstream gene (Grünebast & Clos, 2020). Thus, mature *Leishmania* mRNAs have 5′ caps and poly-A tails like other eukaryotic mRNAs. When introducing a gene of interest into *Leishmania*, whether into the genome or as an episome, it is necessary to flank the gene with the appropriate splicing and polyadenylation signals to ensure proper expression (Roberts, 2011).

1.2.3.1.2 Post-transcriptional control of gene expression

Because of the near absence of transcriptional regulation in *Leishmania*, post-transcriptional control is necessary to modulate the levels and functions of proteins,

especially during transitions between developmental stages of the parasite (Grünebast & Clos, 2020; Karamysheva et al., 2020; Shapira et al., 2001). Factors involved in post-transcriptional regulation include RNA-binding proteins (RBPs), microRNAs, retroposons, and RNA editing enzymes (Bringaud et al., 2007; Grünebast & Clos, 2020; Shapira et al., 2001).

As a general rule, stressors such as heat shock or nutrient dearth bring about global translational repression in *Leishmania* (Karamysheva et al., 2020; Shapira et al., 2001). Cytoplasmic "stress granules" form, which sequester inactive mRNAs and stalled ribosomes (Shrivastava et al., 2019). Some RNAs escape global repression by being bound to RBPs (Nandan et al., 2017; Nocua et al., 2017).

RBPs are groups of proteins that interact with RNAs and determine their structure, processing, stability, localisation, transport, translation, and degradation in response to certain signals or environmental conditions (Ferreira et al., 2020; Karamysheva et al., 2020; Nandan et al., 2017; Shrivastava et al., 2019). RBPs appear to recognise specific sequences or structures in RNA (Ferreira et al., 2020). Activities of RBPs are in turn modulated by post-translational modifications such as arginine methylation, via pathways that have not been fully elucidated (Ferreira et al., 2020). Most RBPs and RBP-modulating factors are yet to be identified, and the mechanisms behind their appearance and function have yet to be characterised (Nandan et al., 2017).

1.2.3.1.3 Aneuploidy and genome plasticity

A major means by which *Leishmania* modulate gene expression is by altering gene copy number. Gene expression is upregulated by making extra copies of genes, clusters of genes, or whole chromosomes (Grünebast & Clos, 2020). Conversely, gene expression is downregulated via deletion of genes, gene regions, or whole chromosomes. Circularisation

of large multi-gene fragments to create extrachromosomal amplicons also happens with some frequency in *Leishmania*, and is one of the means by which they develop drug resistance (Douanne et al., 2022; Ubeda et al., 2008). Hence, *Leishmania* are very loosely diploid; aneuploidy and local copy number variations are rife and used as means of adapting to the numerous stresses and uncertainties associated with switching between radically dissimilar hosts (Cohen-Freue et al., 2007; Roberts, 2011).

The likelihood of a gene being present in more than two copies is something to bear in mind when doing genetic manipulations in *Leishmania*.

1.2.3.1.4 Homologous recombination

Like most eukaryotes, *Leishmania* carry out homologous recombination (HR) as a means of repairing double-stranded breaks in genomic DNA. A remarkable feature of HR in *Leishmania* is that it is also used at extraordinarily high frequency for genomic rearrangements (Ubeda et al., 2014; Zirpel & Clos, 2019). Widely distributed across the *Leishmania* genome are direct or inverted repeated sequences (Ubeda et al., 2014). HR between these repeated elements enables *Leishmania* to amplify or delete regions between the repeats, or create ~20-70kb extrachromosomal amplicons (Ubeda et al., 2008, 2014). These amplicons have been shown to contain the gene targets of some antileishmanial drugs; a higher number of targets renders the parasites insusceptible to the effects of these drugs (Douanne et al., 2022; Ubeda et al., 2014).

The highly active HR system in *Leishmania* has been exploited for decades as the main means of introducing, knocking out, or mutating genes in *Leishmania* genomic DNA (Roberts, 2011). A popular genomic locus for HR-integration of exogenous genes into *Leishmania* is the 18S small subunit ribosomal RNA (SSU) locus. This locus is ideal for a number of reasons. The SSU is highly conserved, so constructs designed for integration

of a gene into one strain/species of *Leishmania* are therefore likely to be suitable for integration of the same gene into the SSU of other *Leishmania* strains/species (Espada et al., 2021). Additionally, the SSU is present in 20-40 copies on chromosome 27 (Martínez-Calvillo et al., 2001). This comes with two main advantages: 1, disruption of a copy of the 18S rRNA gene is unlikely to jeopardise the fitness of *Leishmania*; and 2, it is possible for more than one copy of the targeting construct to be incorporated into SSU loci, which could result in higher expression of an introduced gene.

1.3 *Leishmania* RNA virus

1.3.1 History and taxonomy

The first evidence of possible viruses in *Leishmania* was published in 1974, when virus-like particles (VLPs) were detected in the cytoplasm of *L. hertigi* (now reclassified as *Paraleishmania hertigi*) in electron microscopy images (Molyneux, 1974). Structural and molecular characterisation of *Leishmania* RNA virus (LRV) began more than a decade later (Tarr et al., 1988).

LRV is a double-stranded RNA (dsRNA) virus of the family *Totiviridae*. While totiviruses are most commonly found in yeast and filamentous fungi, they have also been found in plants, crustaceans, wild fish, bats, and protozoa (Hillman & Cohen, 2021). Apart from *Leishmania*, protozoan hosts of totiviruses that are pathogenic to humans include *Giardia lamblia* and *Trichomonas vaginalis*, some strains of which are infected with *Giardia lamblia* virus (GLV) and *Trichomonas vaginalis* virus (TVV), respectively (Stevens et al., 2021).

Two species of LRV infect *Leishmania*. LRV1 is rife in the New World subgenus *Leishmania (Viannia)*, especially in *L. guyanensis, L. braziliensis,* and *L. panamensis.* LRV1 has also been found, to lesser extents, in *L. lainsoni, L. naiffi, and L. shawi*. LRV2

infects subgenus *Leishmania (Leishmania)* – most prominently *L. major*, but has also been detected in a few strains of *L. aethiopica* and *L. infantum* (Kostygov, Grybchuk, et al., 2021).

1.3.2 LRV structure and genome organisation

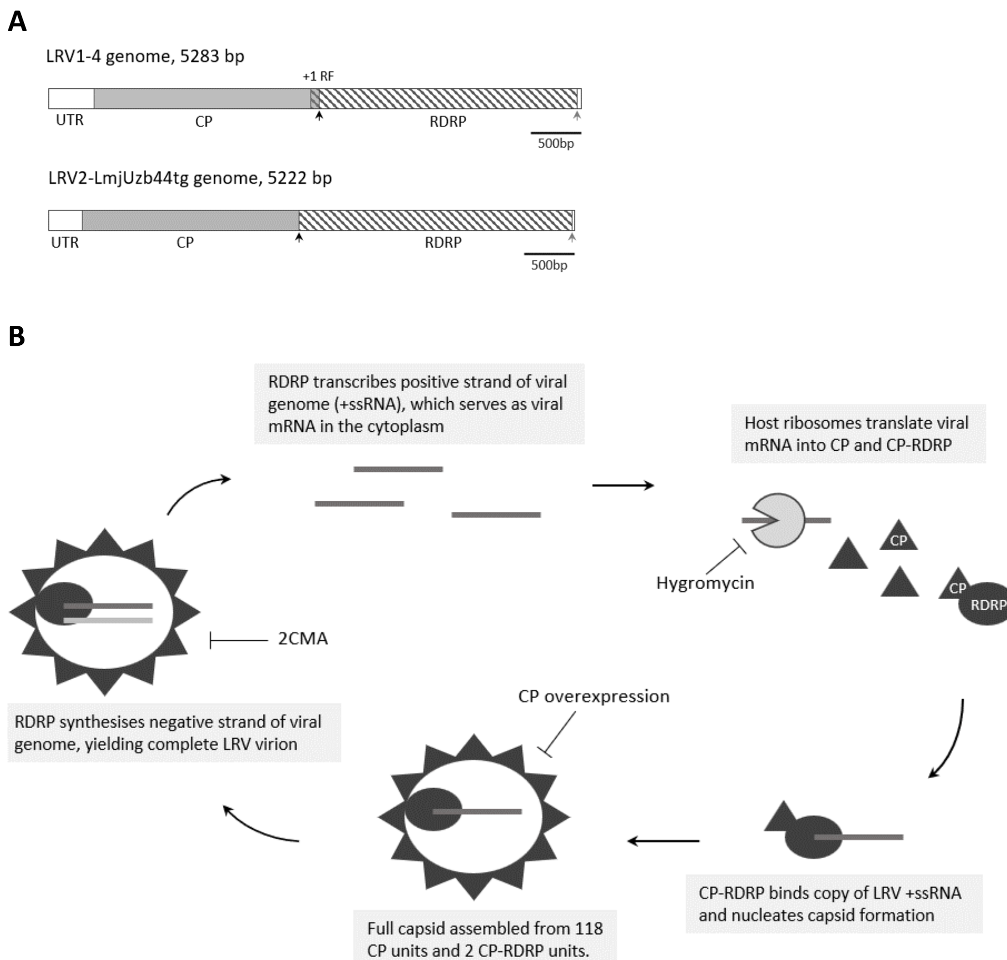


Figure 1.3 LRV genome and replication cycle. A Genome organisation of LRV1-4 and LRV2. Arrows indicate locations of frequent (black) and infrequent (grey) translation termini. RF, ribosomal frameshift. **B** The LRV replication cycle. Putative targets of interventions that have been found to eliminate LRV indicated by ⊥. *Diagrams drawn by author using Microsoft Powerpoint.*

The structures of various totiviruses, including LRVs, have been characterised using cryo-electron microscopy (Procházková et al., 2021). In the host cytoplasm, totiviruses exist as full virions with non-enveloped icosahedral capsids, roughly 40nm in diameter, made up of 120 capsid proteins arranged in 60 asymmetrical dimers (Hillman & Cohen, 2021;

Procházková et al., 2021). The capsid typically contains a single molecule of the totivirus genome, which it shields from host nucleases and the RNA interference system active in most totivirus hosts including *Leishmania* (*Viannia)* but missing in subgenus *Leishmania* (*Leishmania)* (Carrion et al., 2008; Lye et al., 2010). In general, totiviruses do not have an extracellular phase.

Totiviruses have linear, nonsegmented, dsRNA genomes between 4.6-7kb in length. LRV has a genome of size 5.2 to 5.3 kb, comprising 2 open reading frames (ORFs). ORF1 encodes the capsid protein (CP), while ORF2 encodes the RNA-dependent RNA polymerase (RDRP) (Wickner et al., 2012). In some strains of LRV, such as LRV1-4 (the resident virus of *L. guyanensis* M4147), the two ORFs overlap by 71 base pairs (Figure 1.3 A). In others, such as LRV2-LmjUzb44tg (the resident virus of *L. major* T44g), there is no overlap between CP and RDRP, however, the latter's coding sequence does not begin with a start codon (Hillman & Cohen, 2021).

The LRV replication cycle is similar to the model totivirus, L-A, in *Saccharomyces cerevisiae*. Copies of the positive strand of the viral genome (+ssRNA) are transcribed within the capsid by RDRP and extruded to the cytoplasm, where they serve as viral mRNAs (Figure 1.3 B) (Robinson & Beverley, 2018). Due to the lack of a 5′ cap and a poly-A tail, the viral mRNA is susceptible to degradation by host RNA endo- and exonucleases. CP translation is initiated at an internal ribosomal entry site located at the 5′ end of the +ssRNA (Hillman & Cohen, 2021; Lopinski et al., 2000). 98% of the time, translation terminates at the stop codon at end of the CP coding sequence (Figure 1.3 A black arrows) (Wickner et al., 2012). RNA secondary structures at the region of the overlap/junction of CP and RDRP cause stalling of the ribosome (Lopinski et al., 2000). 1.9% of the time, this stalling results in a -1 ribosomal frameshift (RF) in the case of L-A,

a +1 RF in the case of LRV1-4, or a bypass of the CP stop codon in the case of LRV2-LmjUzb44tg. Translation proceeds to the end of the RDRP ORF (Figure 1.3, grey arrows), producing a CP-RDRP fusion protein (Carrion et al., 2008). Two copies of the CP-RDRP fusion protein are incorporated into each 120-subunit capsid (Hillman & Cohen, 2021).

While CPs can polymerise into VLPs on their own, capsid assembly proceeds quickest when nucleated by CP-RDRP (Fujimura et al., 1992). The RDRP part of the fusion protein binds a copy of the viral mRNA, around which a capsid assembles (Esteban et al., 1988). Following encapsidation, RDRP synthesises the negative strand of the viral genome, completing the cycle (Carrion et al., 2008). Thus, the dsRNA genome of LRV is generally not exposed to the host, as it remains fully sequestered by the capsid throughout the viral replication cycle.

1.3.3 The *Leishmania*-LRV relationship

The phylogenetic relationships between LRV species correspond to the phylogenetic relationships between their parasite hosts (Widmer & Dooley, 1995). Furthermore, genetic similarities are highest between LRVs of *Leishmania* strains isolated from geographical locations closest to each other (Cantanhêde et al., 2018). These findings suggest an intricate and long-standing co-evolution of *Leishmania* and LRV. Initial infection of *Leishmania* with LRV is believed to have occurred prior to the divergence of *Leishmania* into different species in the Mesozoic or early Cenozoic era, ~90-36 million years ago (Cantanhêde et al., 2018; Lukeš et al., 2007; Steverding, 2017; Widmer & Dooley, 1995).

While *Leishmania* do not typically carry out conjugation or any other form of sexual reproduction, recent studies have demonstrated that *Leishmania* parasites frequently release extracellular vesicles (EVs) that contain proteins and linear or circular extrachromosomal DNA (Douanne et al., 2022). EVs are used by *Leishmania* to deliver

payloads of effector proteins to mammalian cells to modulate their function (da Silva Lira Filho et al., 2022; Silverman & Reiner, 2012). Neighbouring parasites may also internalise EVs to acquire drug resistance genes and other virulence factors (Atayde et al., 2019; Douanne et al., 2022). Full LRV virions have been detected in EVs derived from LRV-positive *Leishmania* parasites (Atayde et al., 2019; Olivier & Zamboni, 2020). However, experimental infections of LRV-negative strains of *Leishmania* with EVs derived from LRV-positive strains have invariably proven transient, lasting no more than a few weeks (Atayde et al., 2019)[*]. Moreover, while some LRV host switches have occurred over the multi-million-year course of *Leishmania*-LRV co-evolution, these events have been exceedingly rare (Kostygov, Grybchuk, et al., 2021). Therefore, LRV relies almost exclusively on vertical transmission for survival.

The relationship between *Leishmania* and LRV appears to be symbiotic rather than parasitic. There are no appreciable differences in morphology or growth kinetics between LRV-positive and LRV-negative *Leishmania* strains grown in laboratories (Cantanhêde et al., 2021; Kostygov, Grybchuk, et al., 2021). Unlike most viruses, LRV does not lyse its host and is maintained in very low copy number, around 10-16 virions per cell (Kuhlmann et al., 2017; Robinson & Beverley, 2018). This low copy number means the virus would neither overwhelm nor exact a heavy metabolic or fitness toll on the host. LRV may even confer survival advantages to *Leishmania* in the mammalian host stage of their life cycle (next section).

1.3.4 LRV as a factor affecting leishmaniasis severity

In the last decade, an association has been made between the presence of LRV and the progression of leishmaniasis from the relatively benign cutaneous form to more destructive

[*] In addition to these published experimental infections, infections of a range of other LRV-negative strains have been attempted by the same group, with similar results (from personal communication).

disseminated or mucocutaneous forms (Cantanhêde et al., 2015; de Carvalho et al., 2019; Ives et al., 2011). Roughly a quarter of *L. braziliensis* and *L. guyanensis* clinical isolates from patients in Brazil are LRV1-positive (Bourreau et al., 2016; Ives et al., 2011). Metastasising strains of *L. guyanensis* have been found to have higher LRV1 levels than non-metastasising strains (Ives et al., 2011). Studies in mice and humans have shown that LRV-harbouring *Leishmania* are more resistant to antileishmanial treatment, have higher capacity for metastatic spread, and are more likely to cause symptomatic relapse months or years after apparent recovery (Abtahi et al., 2020; Bourreau et al., 2016; Hartley et al., 2012).

LRV facilitates survival of *Leishmania* by modulating the immune response of the vertebrate host, partly via the innate recognition of LRV's dsRNA genome by the host's Toll-like receptor 3 (TLR3) (Hartley et al., 2012). TLR3 activation leads to production of proinflammatory cytokines IL-6, IL-10, IFN1β, TNF-α, CXCL10, CCL4, and CCL5 (Hartley et al., 2012; Ives et al., 2011; Rath et al., 2022). TLR3 activation also triggers the autophagic machinery in macrophages, which tempers the NLRP3 inflammasome activity normally involved in limiting *Leishmania* replication (de Carvalho et al., 2019; Hartley et al., 2012; Rath et al., 2022). These processes lead to the inordinate infiltration and destruction of cutaneous and mucosal tissue by immune cells, which aid parasite survival (Rath et al., 2022).

1.3.5 Previous attempts at eliminating LRV

Various strategies have been formulated for ridding LRV-harbouring *Leishmania* of their resident virus (Figure 1.3). Because they are isogenic to LRV-harbouring wild type lines, the "cured" parasite lines serve as better negative controls than naturally virus-free strains for studies interrogating the roles played by LRV in *Leishmania* virulence (Castiglioni et

al., 2017; Kuhlmann et al., 2017). Parasite responses to the process of LRV elimination (e.g., changes in growth, morphology, motility) may also illuminate critical aspects of the parasite-virus symbiosis and reveal possible targets for antileishmanial therapy (Hartley et al., 2012). Treatments that eliminate LRV from *Leishmania* in patients may ameliorate the clinical severity of leishmaniasis and potentiate standard antileishmanial treatment.

1.3.5.1 Chemical interventions

In 1997, the M4147 strain of *L. guyanensis* transfected with a plasmid vector bearing a hygromycin resistance gene and treated with 50µg/mL hygromycin B for several weeks appeared to have been cured of LRV1-4 (Ro et al., 1997). This became a benchmark for subsequent LRV1 studies (Castiglioni et al., 2017; Ives et al., 2011; Lye et al., 2010). The authors of the study speculated that hygromycin B specifically inhibited translation of viral transcripts, leading to viral elimination (Ro et al., 1997).

Another study targeted RDRP using 2′-C-methyladenosine triphosphate (2CMA) to inhibit viral transcription and replication, which led to complete viral elimination without any observable effect on *Leishmania* growth kinetics and morphology (Kuhlmann et al., 2017; Robinson & Beverley, 2018). This method has since been used to create isogenic virus-negative versions of LRV-harbouring *Leishmania* strains (de Carvalho et al., 2019)[*].

1.3.5.2 CP overexpression

In 1995, Widmer overexpressed episomally-encoded LRV2-1 CP in LRV2-1's natural host, *L. major* 5-ASKH. This led to a marked reduction (but not elimination) of the level of endogenous LRV2-1, which remained stable over time (Widmer, 1995). The overexpressed CPs were able to assemble into VLPs, and had a dominant negative effect

[*] 2CMA was also used to create isogenic LRV-negative *L. guyanensis* M4147 and *L. major* T44g lines in another part of the journal article of which this study is a part (Appendix C).

on the assembly of endogenous capsids. A subsequent study showed that inoculation of mice with adjuvanted recombinant LRV1-4 CPs protected them from developing complications of leishmaniasis when later challenged with LRV1-4-infected *L. guyanensis* (Castiglioni et al., 2017). Viral copy numbers of the L-A totivirus and its satellite M virus have also been reduced or eliminated from their host *S. cerevisiae* after overexpression of full-length or truncated versions of the L-A virus coat protein (Aitmanaitė et al., 2021; Lukša et al., 2017; Yao et al., 1995).

1.4 Research Outline

In two previous studies in totiviruses, expression of full-length or truncated versions of the viral capsid protein (CP) led to copy number reduction or complete elimination of the virus from the host. This dominant negative effect is speculated to be due to the ability of CPs to assemble into virus-like particles (VLPs) even in the absence of CP-RDRP. Theoretically, overabundance of CPs disrupts the CP:CP-RDRP ratio in the cytoplasm (normally 59:1), leading to the formation of VLPs with one or zero copies of CP-RDRP (Figure 1.4). As CP on its own has no binding sites for ssRNA or dsRNA, capsids devoid of RDRP cannot encapsidate the viral genome. Every round of host cell division yields daughter cells with fewer copies of properly formed virions, ultimately leading to viral elimination.

This study is undertaken to assess the specificity of CP-induced LRV elimination from *L. guyanensis* M4147 and *L. major* T44g. Five types of LRV CPs have been selected for overexpression: LRV1-4 CP, derived from the virus harboured by *L. guyanensis* M4147; LRV2- LmjUzb44tg (hereafter referred to as LRV2 as it is the only strain of LRV2 used in this study) CP derived from *L. major* T44g; and CPs from three other LRV1 strains. The ability of CPs from these five strains to bring about copy number reduction or

elimination of LRV1-4 from *L. guyanensis* M4147 or LRV2 from *L. major* T44g is speculated to be contingent on the similarity between the exogenous CP and the endogenous CP of the resident LRV.

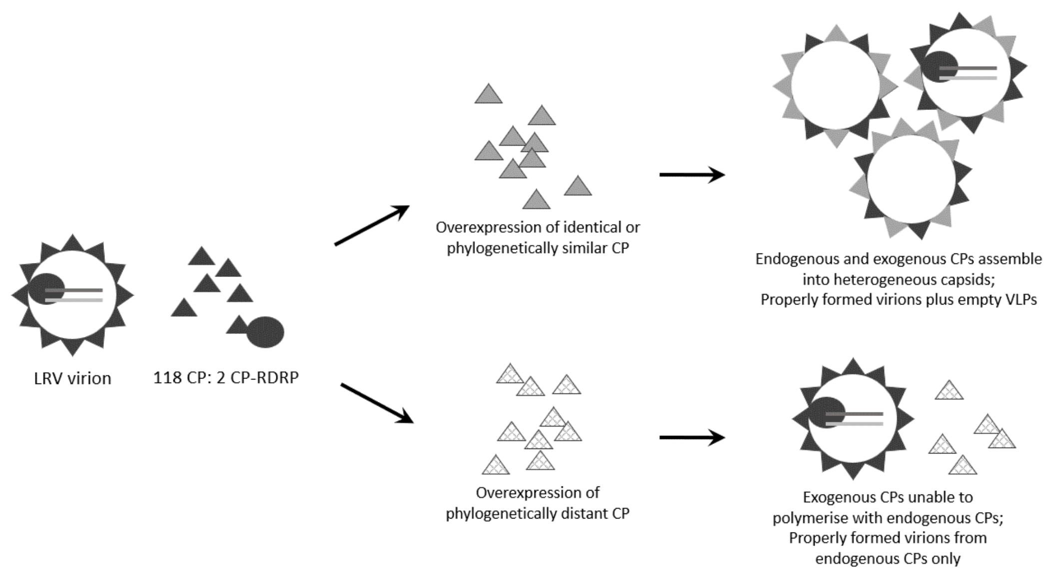


Figure 1.4 Model of how LRV CP overexpression eliminates LRV. Overexpressed exogenous CPs form heterogeneous capsids with CPs of the resident LRV, diluting the number of properly formed virions that harbour the viral genome. This is contingent on similarity between exogenous and native CPs. *Diagram drawn by author using Microsoft Powerpoint.*

As four of the CPs are derived from LRV strains of the same phylogenetic group, LRV1, it is speculated that overexpression of these CPs would eliminate LRV1-4 from *L. guyanensis* M4147, whereas overexpression of the more phylogenetically distant LRV2 CP would have no effect on LRV1-4 copy number.

Conversely, overexpression of LRV2 CP is expected to eliminate LRV2 from *L. major* T44g, whereas CPs from the four LRV1 strains are expected to have little to no effect on LRV2 copy number.

1.4.1 Hypothesis:

Overexpression of CPs from LRV1-4, LRV1-Lg2014, LRV1-LBrLEM2700, or LRV1-1 eliminates LRV1-4 from *L. guyanensis* M4147, but does not eliminate LRV2 from *L.*

major T44g. Overexpression of LRV2 CP in *L. major* eliminates LRV2 from *L. major*, but does not eliminate LRV1-4 from *L. guyanensis*.

1.4.2 Research Aims:

1. Construct vectors for integrating HA-tagged capsid genes of LRV1-4, LRV1-Lg2014, LRV1-LBrLEM2700, LRV1-1, and LRV2 into the 18S small subunit ribosomal RNA (SSU) locus of *L. guyanensis* M4147 and *L. major* T44g via homologous recombination.
2. Transfect *L. guyanensis* and *L. major* with vectors from Aim 1.
3. Ascertain transcription of recombinant CPs from SSU via quantitative RT-PCR.
4. Ascertain translation of HA-tagged CPs via Western blotting.
5. Assess effect of expression of 5 different capsids on LRV1-4 copy number in *L. guyanensis* and LRV2 copy number in *L. major* via quantitative RT-PCR.

CHAPTER 2. MATERIALS AND METHODS

2.1 Materials

Table 2.1 Sources of consumable materials used in this work.

Manufacturer/Source	Materials
Bemis Company, Inc., Neenah, WI, USA	Parafilm
BioSera Europe, Nuaillé, France	Foetal Bovine Serum
Eppendorf, Hamburg, Germany	0.2mL PCR tubes, 1.5mL microcentrifuge tubes, 2mL microcentrifuge tubes, LoBind 1.5mL microcentrifuge tubes, LoBind 2mL microcentrifuge tubes
Jena Bioscience, Jena, Germany	pLEXSY-neo2.1
Kapa Biosystems, Wilmington, MA, USA	KAPA SYBR® FAST
MilliporeSigma, Burlington, MA, USA	M199 medium, α-HA antibodies, α-tubulin antibodies
Molecular Research Center Inc., OH, USA	TRI Reagent
New England Biolabs, Ipswich, MA, USA	BglII restriction enzyme, DNAse I, glycogen, NcoI restriction enzyme, NotI restriction enzyme, Q5 High-Fidelity DNA Polymerase, SwaI restriction enzyme
Nippon Genetics Europe, Düren, Germany	Midori Green Advance
PCR Biosystems, London, UK	PCRBIO Taq Mix Red, PCRBIO VeriFi Polymerase, PCRBIO HiFi Polymerase
Qiagen, Hilden, Germany	PCR QIAquick PCR Purification Kit
Roche Life Science, Basel, Switzerland	cOmplete™ EDTA-free Protease Inhibitor Cocktail, Transcriptor High Fidelity cDNA Synthesis Kit
Serva, Heidelberg, Germany	Acrylamide/Bis (29:1), Tris
Sigma-Aldrich, St. Louis, MO, USA	25cm^2 vented culture flasks
Thermo Fisher Scientific, Waltham, MA, USA	0.2mL thin-walled PCR tubes, 384-well qPCR plates, 96-well qPCR plates, APS, biopterin, GeneJET gel extraction kit, GeneJET plasmid miniprep kit, haemin, HEPES, hygromycin, penicillin, Pierce BCA Protein Assay, Pierce ECL Western blotting substrate, Pierce RIPA buffer,

	puromycin, PVDF membrane, proteinase K, RNase A, streptomycin, T4 DNA ligase, 50× TAE buffer, TEMED, TURBO DNA-free Kit
VWR, Radnor, PA, USA	6-well cell culture plates, 12-well cell culture plates, 15mL centrifuge tubes, 24-well cell culture plates, 50mL centrifuge tubes, 96-well cell culture plates, agar, agarose, EDTA, G418, glycine, KCl, KH_2PO_4, KOH, NaCl, Na_2HPO_4, NH_4OAc, petri plates, pipette tips, SDS, Tween-20

Table 2.2 Manufacturers of equipment used in this work.

Manufacturer	Equipment
Bioer, Zhejiang, China	Mixing/heating block
Bio-Rad, Hercules, CA, USA	Electrophoresis power source, thermocycler, Trans-Blot Turbo transfer system
BioSan, Riga, Latvia	Microspin FV-2400 vortex, UVT-S-AR RNA box
BioLogics, Cary, NC, USA	Ultrasonic homogenizer
Eppendorf, Hamburg, Germany	Thermocycler
ESCO, St. Louis, MO, USA	Airstream class II biological safety cabinet
Hanil Science Industrial, Incheon, South Korea	Smart R17 Micro Refrigerated Centrifuge
Lonza, Basel, Switzerland	Nucleofector 2b electroporator
Nuve, Ankara, Turkey	EN 055 incubator
Roche Life Science, Basel, Switzerland	LightCycler 480 II real-time PCR cycler
Tecan, Männedorf, Switzerland	Infinite M Nano plate reader
Thermo Scientific, Waltham, MA, USA	Heraeus Multifuge X1R refrigerated centrifuge, NanoDrop Lite spectrophotometer, Smart2Pure milliQ water purifier
Velp Scientifica, Lombardia, Italy	AREC Heating magnetic stirrer
VWR, Radnor, PA, USA	Thermocycler

2.2. Methods

2.2.1 *Leishmania* culture

All *Leishmania* parasite culture procedures were carried out in an ESCO Airstream class II biological safety cabinet, using the aseptic technique. Cultures were incubated at 23-26 °C following each procedure unless otherwise stated.

2.2.1.1 Passaging of *Leishmania* cultures

Leishmania promastigotes – wild type strains *Leishmania guyanensis* MHOM/BR/78/M4147 (hereafter referred to as *L. guyanensis* M4147 or *L. guyanensis*), *Leishmania major* MRHO/UZ/2003/IsvT44g (hereafter referred to as *L. major* T44g or *L. major*), and genetically modified derivatives of these two strains – were grown in 25 cm^2 filter-capped cell culture flasks in M199 growth medium supplemented with 2 µg/mL haemin, 10% heat-inactivated foetal bovine serum, 2 µg/mL biopterin, 100 U/mL penicillin, 100 µg/mL streptomycin and 40 mM HEPES (pH 7.4). Parasites were passaged weekly to fresh medium at a 1:20 to 1:50 dilution when they reached stationary phase.

2.2.1.2 Freezing of *Leishmania* cultures

50 µL of DMSO was added to 950 µL of a *Leishmania* culture at late log to early stationary phase in a 2 mL cryovial. Cryovials were kept at -80 °C for a few days, then transferred to a liquid nitrogen tank for long-term storage.

2.2.1.3 Thawing of *Leishmania* cultures

Cryovials were thawed at room temperature. Contents of each cryovial were transferred to a 15mL tube containing 9 mL of fresh M199 growth medium, and mixed via gentle

inversion. Parasite suspensions were centrifuged at 1,000 × *g* for 10 minutes at 4 °C. Pellets were resuspended in 5 mL of M199 in 25 cm^2 flasks.

2.2.1.4 Growth kinetics analysis

Triplicates of each strain were seeded at an initial density of 5×10^5 parasites mL^{-1}. Parasite densities were measured every 24 hours using a haemocytometer.

2.2.1.5 Making clonal *Leishmania* lines

Cultures were diluted to 2×10^4 parasites mL^{-1}. 50 µL of the parasite suspension was spread onto 25 mL 1% agar M199 plates with the appropriate antibiotics. Plates were sealed with parafilm, inverted, and incubated at 23-26 °C until colonies were visible. Single colonies were picked with sterile pipette tips and transferred to 200 µL liquid M199 in 96-well plates. Clonal parasite suspensions were transferred to standard vented 25 cm^2 flasks after one week.

2.2.1.6 Transfection

2.2.1.6.1 Electroporation

Leishmania parasites were grown to mid-late log phase (1-3 × 10^7 parasites mL^{-1}). Parasites were centrifuged at 1000 × *g* for 10 minutes at 4 °C, washed with cytomix (67 mM Na_2HPO_4, 23 mM NaH_2PO_4, 5 mM KCl, 50 mM HEPES [pH 7.4], 0.15 mM $CaCl_2$, recipe adapted from Beneke [2017]), centrifuged at 1000 × *g* for 5 minutes at 4 °C, then resuspended in cytomix at a density of 1.5×10^8 parasites mL^{-1}. 100 µL of the parasite suspension was mixed with 60 µL DNA solution (linearised vector and/or PCR products) or cytomix (mock transfection) and transferred into a 2 mm electroporation cuvette. Parasites were subjected to electroporation using Lonza Nucleofector 2b, programme X-

001, 1 pulse (as per Beneke, 2017). Parasites were transferred to 25 cm^2 flasks containing 5 mL of prewarmed M199 without antibiotics for 16-20 hours.

2.2.1.6.2 Antibiotic selection

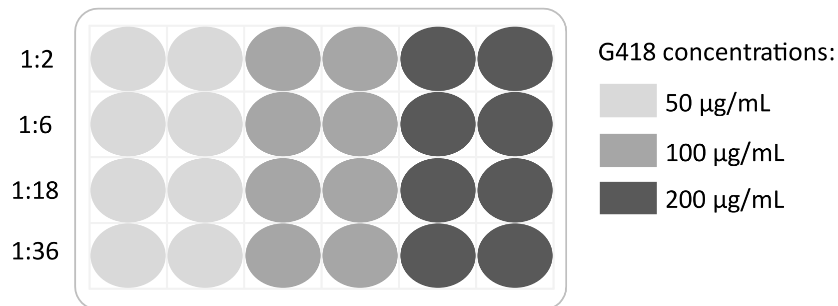


Figure 2.1 Antibiotic selection in 24-well plates. Parasites in recovery medium (3×10^6 parasites mL^{-1} M199) were seeded into a 24-well plate at four dilutions (rows), and treated with three concentrations (columns) of G418 sulphate. Each well contained 1 mL of parasite suspension. *Diagram drawn by author using Microsoft Powerpoint.*

Transfectants that had successfully integrated antibiotic resistance markers into their genomes were selected via treatment with three concentrations of the appropriate antibiotic (G418 sulphate) in 24-well plates (Fig. 2.1). Successful transfectants became apparent 10-21 days after transfection, when all the parasites in the mock-transfected control setup had died.

2.2.1.6.3 Transfection of *Leishmania* parasites with constructs for genomic integration into *Leishmania* SSU locus

20 µg of a pLEXSY vector derivative was digested overnight with SwaI restriction enzyme. Digestion products were separated via agarose gel electrophoresis. The band of interest was excised from the gel, purified using GeneJET gel extraction kit, then quantified using NanoDrop Lite spectrophotometer.

5 µg of linearised vector was used for each transfection. Transfection was carried out via electroporation (see 2.2.1.4.1 and 2.2.1.4.2). Genomic DNA of G418-resistant parasites were checked for correct integration of transfection constructs.

2.2.2 Bacteria

All procedures involving *Escherichia coli* were carried out following the aseptic technique.

2.2.2.1 Transformation

Five ng of purified plasmid DNA, or 2 µL of a 10 µL ligation mix, was added to 100 µL of competent *E.coli* XL-Blue cells. The transformation mixture was kept on ice for 30 minutes, then subjected to heat shock at 42 °C for 45 seconds. 35 µL of the transformation mix was spread onto an LB-agar plate (1% agar) supplemented with 100 µg mL^{-1} ampicillin and incubated at 37 °C overnight.

2.2.2.2 Plasmid production

Colonies were picked from LB-agar plates and inoculated into 5mL of LB supplemented with 100 µg mL^{-1} ampicillin. Cultures were incubated overnight, with shaking, at 37 °C. Plasmids were extracted using GeneJET plasmid miniprep kit.

2.2.3 DNA

2.2.3.1 Polymerase Chain Reaction (PCR)

2.2.3.1.1 Standard PCR (Taq)

Each 50 µL PCR reaction contained 0.5 µM of forward primer, 0.5 µM reverse primer, 10 ng plasmid DNA or 100 ng genomic DNA as template, and 25 µL of 2× PCRBIO Taq Mix Red (Taq DNA polymerase, 6 mM $MgCl_2$, 2 mM dNTPs). PCR was carried out using a thermocycler under the following conditions (per manufacturer’s recommendations):

Table 2.3 PCR settings, standard

Step		Temperature, °C	Time, seconds	Cycles
Preamplification denaturation		95	60	1
Thermal cycling	Denaturation	95	15	35
	Annealing	58	15	
	Extension	72	20 per kb	
Final Extension		72	120	1

2.2.3.1.2 High fidelity PCR

Each 50µL PCR reaction contained 0.5 µM of forward primer, 0.5 µM reverse primer, 10 ng plasmid DNA or 100 ng genomic DNA as template, 200 µM dNTPs, 1U Q5 Hot Start High Fidelity DNA Polymerase, and 10 µL 5× Q5 Reaction Buffer (125 mM TAPS-HCl, 250 mM KCl, 10 mM MgCl2, 5 mM β-mercaptoethanol).

Hot-start PCR was carried out using a thermocycler under the following conditions (per product documentation and manufacturer's product-specific Tm calculator*):

Table 2.4 PCR settings, high fidelity

Step		Temperature, °C	Time, seconds	Cycles
Preamplification denaturation		98	30	1
Thermal cycling	Denaturation	98	10	35
	Annealing	62	20	
	Extension	72	30 per kb	
Final Extension		72	180	1

2.2.3.1.3 Colony PCR

Bacterial colonies were picked from LB agar plates and suspended in 50µL ddH_2O. 1µL of each resulting suspension was used as template for a standard 10µL PCR reaction (2.2.3.1.1).

* https://tmcalculator.neb.com/

2.2.3.2 Genomic DNA extraction

1mL of *Leishmania* cultures at mid- to late-log phase were centrifuged at 1,000 × *g* for 1 minute. The pellets were resuspended in 100 µL PBS (137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 1.8 mM KH_2PO_4). 100 µL of cell lysis buffer (10 mM Tris pH 8.0, 100 mM NaCl, 10 mM EDTA pH 8.0, 0.50% SDS) was added to each cell suspension, along with 2 µL proteinase K and 3 µL RNase A. The samples were incubated at 65 °C for 30 minutes with agitation. An equal volume of phenol/chloroform/isoamyl alcohol (25:24:1) was added to each sample. Samples were mixed by inversion, then centrifuged at maximum speed for 5 minutes. The upper aqueous phase of each sample was transferred to a fresh tube. 100 µL of 7.5 M NH_4OAc, 750 µL 96% ethanol, and 1 µL glycogen (20 µg/µL) were added to each sample. The samples were kept at -20 °C for 1-16 hours to precipitate DNA. The samples were centrifuged at 16,000 × *g* for 20 minutes at 4 °C. The supernatants were discarded. The pellets were washed twice with 150 µL 70% ethanol, then centrifuged at 16,000 × *g* for 2 minutes. The pellets were air-dried for 10 minutes, then resuspended in 50 µL TE buffer (10mM Tris, 1mM EDTA, pH 8.0).

2.2.3.3 Agarose gel electrophoresis

1% (w/v) agarose gel in 1× TAE (40 mM Tris, 20 mM acetic acid, 1 mM EDTA), prestained with Midori Green Advance, was placed into a horizontal electrophoresis tank, and submerged in 1× TAE. DNA MW markers were loaded into one well. DNA samples, mixed (5:1) with NEB 6× Purple Gel Loading Dye were loaded into designated wells. Electrophoresis was carried out at 10-15 V/cm for 15 to 60 minutes. Bands were visualised and photographed using the Syngene G:BOX gel imaging system and GeneSys software.

2.2.3.4 Restriction endonuclease digestion

In general, 1-3 µL restriction endonuclease (10U/µL) and 6 µL of the 10× concentrated recommended buffer were added to 5-25 µg plasmid DNA, in a total reaction volume of 60 µL. The digestion mix was incubated at the recommended temperature for 2-16 hours, then analysed by AGE (2.2.3.3).

2.2.3.5 Ligation

In general, a vector:insert molar ratio between 1:5 to 1:10 was used, totalling 50 ng of DNA per reaction. 1 µL 10× T4 DNA ligase buffer and 0.5 µL T4 DNA ligase were added to the vector+insert DNA mixture in a total reaction volume of 10 µL. The ligation mix was incubated at room temperature for 10 minutes. 3 µL of the ligation mix was used for transforming bacteria (section 2.2.2.1).

2.2.3.6 Preparation of samples for sequencing

Each 10 µL sequencing mix contained 2.5 µM of a sequencing primer and either 400-500 ng of purified plasmid or 10-50 ng of purified PCR product. Samples were labelled with barcode stickers and sent to Eurofins Genomics (Ebersberg, Germany) for Sanger sequencing.

2.2.4 RNA

2.2.4.1 RNA extraction

5-10 × 10^7 parasites from a *Leishmania* culture at mid- to late-log phase were centrifuged at 1,000 × *g* for 15 minutes at 4 °C. The pellet was dissolved in 1 mL Tri Reagent and incubated at room temperature for 10 minutes. The cell suspension was mixed with 200 µL of chloroform, incubated at room temperature for 10 minutes, then centrifuged at 12,000 × *g* for 15 minutes at 4 °C. The upper aqueous layer was transferred to a new tube,

mixed with 500 µL of isopropanol, then kept at -80 °C for 2 hours. The solution was centrifuged at 12,000 × *g* for 10 minutes at 4 °C. The pellet was washed twice with 500 µL of 75% ethanol, centrifuged at 12,000 × *g* for 5 minutes at 4 °C after each wash, then air dried for 10 minutes. The RNA pellet was dissolved in 25 µL of nuclease-free water, then quantified using NanoDrop Lite.

2.2.4.2 DNase treatment

21.5 µL of RNA suspension from 2.2.4.1 was mixed with 2.5 µL 10× TURBO DNase Buffer and 1 µL TURBO DNase. The mixture was incubated at 37 °C for 1 hour. 2.5 µL of TURBO DNase Inactivation Reagent was added to the mixture. After 5-minute incubation at room temperature, the mixture was centrifuged for 5 minutes using a rotor centrifuge. The supernatant was transferred to a fresh tube, then quantified using NanoDrop Lite.

2.2.4.3 cDNA synthesis

For each sample (from 2.2.4.2), 1-4 µg of DNase-treated RNA was converted to cDNA using Transcriptor High Fidelity cDNA Synthesis Kit (Roche). Because this protocol was principally used for assessing changes in LRV copy number, and neither strand of the dsRNA genome of LRV is polyadenylated, random hexamer primers (instead of poly-T primers) in the kit were used.

2.2.4.4 Quantitative PCR

Samples (from 2.2.4.3) were diluted 3- to 5-fold with ddH_2O. A standard solution was made from a mixture of 5 µL aliquots of each diluted sample. The standard was serially diluted 5-fold 5 times for construction of a standard curve.

Each 10µL reaction mix contained 2.5 pmol forward primer, 2.5 pmol reverse primer, 5 µL 2× KAPA SYBR FAST mix (Kapa Biosystems), and 1 µL sample. Reaction mixes, prepared in triplicate for each standard/sample, were loaded into designated wells of a 96- or 384-well qPCR plate. Quantitative PCR was carried out using a LightCycler 480 II real-time PCR cycler.

2.2.5 Protein

2.2.5.1 Protein extraction via sonication

5mL cultures of parasites at log-phase were centrifuged at 1,000 × *g* for 10 minutes at 4 °C. The pellets were washed once with 5 mL PBS, then centrifuged at 1,000 × *g* for 10 minutes at 4 °C. The pellets were resuspended in 600 µL lysis buffer (10 mL RIPA buffer [50 mM Tris HCl, 150 mM NaCl, 1.0% (v/v) NP-40, 0.5% (w/v) sodium deoxycholate, 1.0 mM EDTA, 0.1% (w/v) SDS and 0.01% (w/v) sodium azide, pH 7.4] plus 1 tablet cOmplete™ EDTA-free Protease Inhibitor Cocktail) and transferred to 2 mL microcentrifuge tubes set on ice. Each sample was subjected to 3 rounds of 4 pulses of sonication at 40% intensity, with 10 second intervals between rounds. The samples were centrifuged at 16,000 × *g* for 10 minutes at 4 °C. The supernatants containing the extracted proteins were transferred to new tubes and quantified.

2.2.5.2 Protein quantification

2,000 µg/mL bovine serum albumin (BSA) was mixed with RIPA buffer to generate the following protein concentration standards: 2,000 µg/mL, 1,000 µg/mL, 800 µg/mL, 600 µg/mL, 400 µg/mL, 200 µg/mL, 100 µg/mL, 0 µg/mL. Each sample was diluted 2-fold, 5-fold, and 10-fold. 10 µl of each standard/sample was loaded into a designated well of a 96-well plate, in duplicate. 200 µL of Pierce working reagent (50 parts BCA reagent A: 1 part

BCA reagent B, Thermo Scientific) was added to each well. Samples were mixed using a plate shaker, then incubated at 37 °C for 30 minutes. Absorbance of each sample at 562 nm was measured using a plate reader (Tecan Infinite M Nano). A standard curve was constructed by plotting the blank-corrected 562 nm reading of each BSA standard vs its concentration. Concentrations of samples were interpolated from the standard curve.

2.2.5.3 Sodium dodecyl sulphate – polyacrylamide gel electrophoresis (SDS-PAGE)

10% polyacrylamide separating gels (10% [v/v] acrylamide/bis, 390 mM Tris [pH 8.8], 0.1% [w/v] ammonium persulphate, 0.1% [w/v] SDS, 0.01% [v/v] TEMED) overlaid with 4% polyacrylamide stacking gels (4% acrylamide/bis, 125 mM Tris [pH 6.8], 0.1% APS, 0.1% SDS, 0.01% TEMED) were cast in 1.5 mm width mini-PROTEAN plates (Bio-Rad). Gel cassettes were placed into electrophoresis tank containing running buffer (25 mM Tris, 192 mM glycine, 0.1% SDS). 15 µg protein samples in loading dye (50 mM Tris-HCl [pH 6.8], 2% SDS, 1.5 mM bromophenol blue, 10% [v/v] glycerol, 0.6 M β-mercaptoethanol) were heated at 95°C for 5-10 minutes. Samples were spun briefly, then loaded into designated wells of the stacking gel. One well was loaded with prestained protein size markers. Electrophoresis was carried out at 45 V for 40 minutes, then increased to 140-180 V for 1.5 hours or until the dye reached the bottom of the gel.

2.2.5.4 Western blotting

A PVDF membrane was activated by soaking in methanol, then equilibrated with transfer buffer (25 mM Tris-Cl [pH 8.3], 192 mM glycine, 20% [v/v] methanol). Proteins separated by SDS-PAGE were electrotransferred to the PVDF membrane using the Trans-blot Turbo transfer system (Bio-Rad) at 25V for 30 minutes. The membrane was washed with TBST (Tris: 20 mM Tris, 150 mM NaCl, 0.1% Tween-20), then incubated in blocking buffer

(5% low-fat milk [w/v] in TBST) for 60 minutes. The membrane was washed thrice for 5 minutes with TBST, then incubated with primary antibodies (1:1,000 α-HA antibodies or 1:5.000 α-tubulin antibodies in blocking buffer) for 60 minutes. The membrane was washed thrice for 10 minutes with TBST, then incubated with HRP-conjugated secondary antibodies (1:80,000 in blocking buffer) for 60 minutes. The membrane was washed thrice for 10 minutes with TBST, incubated with Pierce ECL Western blotting substrate (Thermo Scientific), then exposed and imaged using Syngene G:BOX gel imaging system and GeneSys software.

2.2.6 Statistical analysis

Data from RT-qPCR (2.2.4.3.2) was analysed using two-tailed Student's t-test.

CHAPTER 3. RESULTS

3.1 Construction of transfection vectors for genomic integration of LRV CP genes

3.1.1 Vector selection

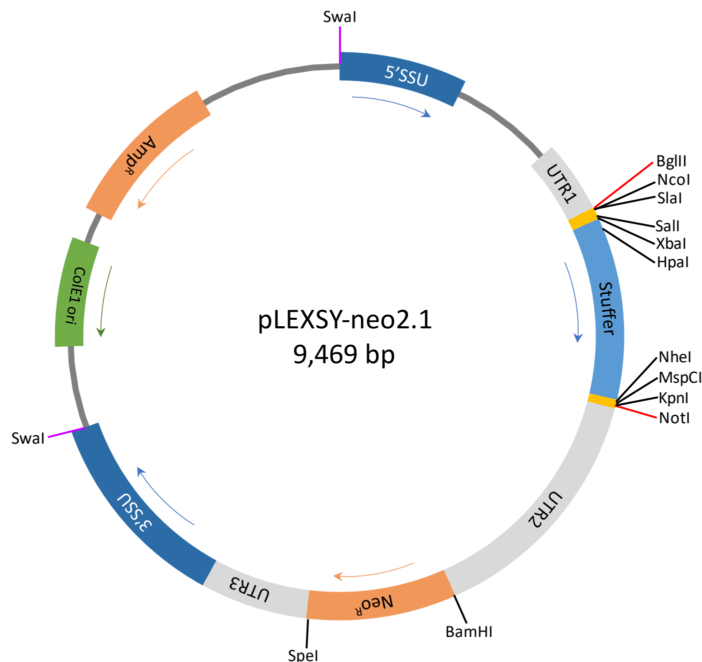


Figure 3.1. Plasmid map of pLEXSY-neo2.1. Restriction sites chosen for replacement of stuffer sequence with CP sequence indicated in red; SwaI restriction sites for vector linearisation prior to transfection indicated in purple. *Plasmid map constructed by author using Microsoft Excel and Microsoft PowerPoint.*

As the first aim of the project was generation of *L. guyanensis* and *L. major* parasite lines constitutively expressing five variants of LRV CPs from a genomic locus, transfection vectors were constructed from a pLEXSY integrative constitutive expression vector (Jena Bioscience). The pLEXSY-neo2.1 vector chosen for this study has the following features (Figure 3.1):

- Bacterial elements for propagation in *Escherichia coli*, which are removed upon vector linearisation with SwaI prior to transfection:
    - ColE1 ori: bacterial origin of replication

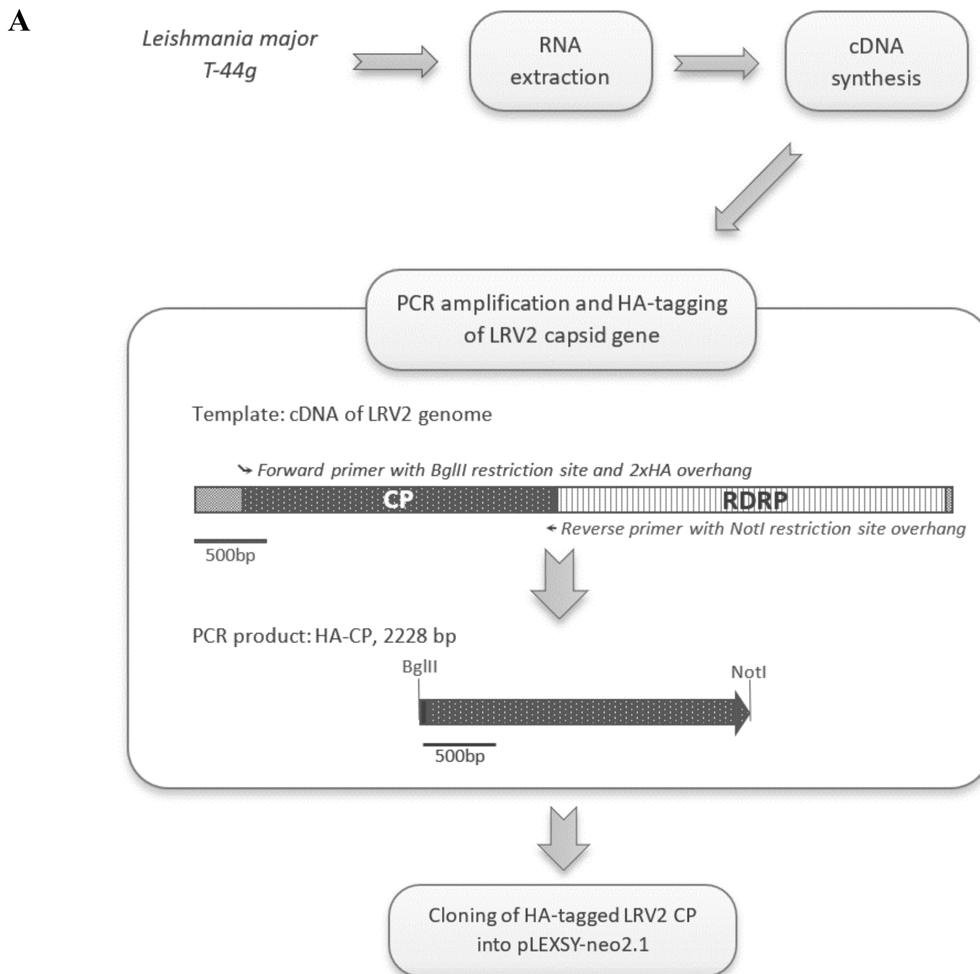
- Amp^R: beta-lactamase marker gene

- Features for chromosomal integration and expression of a target gene in *Leishmania*:
  - 5′ SSU: a ~600 bp fragment homologous to the 5′ end of the *Leishmania* 18S SSU rRNA locus
  - UTR1: the 5′ untranslated region of the adenine phosphoribosyltransferase gene of *Leishmania tarentolae* encoding the splice acceptor site for the target gene
  - Stuffer: a replaceable ~1 kb nonsense sequence, flanked by unique restriction sites
  - UTR2: a ~1.4 kb intergenic region from the calmodulin operon of *L. tarentolae* encoding the poly-A site for the target gene and the splice acceptor site for the marker gene
  - Neo^R: aminoglycoside phospho-transferase 3′(I), a marker gene which confers resistance to G418/neomycin
  - UTR3: the untranslated region of the dihydrofolate reductase-thymidylate synthase gene of *L. major* encoding the poly-A site for the marker gene
  - 3′ SSU: a ~1 kb fragment homologous to the 3′ end of the Leishmania 18S SSU rRNA locus

The sequences derived from the 5′ and 3′ parts of the *Leishmania* SSU are the sites at which homologous recombination takes place between the linearised pLEXSY vector and the genomic DNA of *Leishmania*. The parts of the vector flanked by these sequences get integrated into the chromosomal SSU locus. The UTRs enable efficient *trans*-splicing, polyadenylation, and translation of the exogenous CP gene and the antibiotic resistance marker gene.

3.1.2 Construction of LRV CP expression vectors derived from pLEXSY-neo2.1

A



B

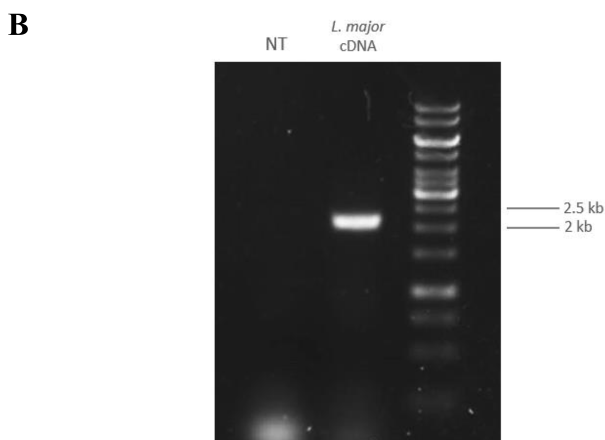


Figure 3.2 Cloning the LRV2 CP gene. A Strategy for cloning the LRV2 CP gene from LRV2 harboured by *Leishmania major*. Total RNA, which includes the dsRNA genome of LRV2, was isolated from a log phase culture of *L. major* T44g and converted to cDNA. The LRV2 CP gene was amplified from cDNA and tagged with HA. **B** Agarose gel electrophoresis of HA-tagged LRV2 CP PCR product. *CP, capsid protein; HA, hemagglutinin; NT, 'no template' control; RDRP, RNA-dependent RNA polymerase. Diagram in A constructed by author using Microsoft PowerPoint.*

Two LRV-infected *Leishmania* strains – *L. major* T44g (hereafter referred to as *L. major*) and *L. guyanensis* M4147 (hereafter referred to as *L. guyanensis*), harbouring LRV2 and LRV1-4, respectively – were available in Prof. Yurchenko's collection.

The LRV2 CP was cloned into pLEXSY-neo2.1 as follows (Figure 3.2). Total RNA was isolated from a log-phase culture of *L. major*, which included the dsRNA genome of its resident virus LRV2. RNA was converted into cDNA using random hexamer primers. The LRV2 CP sequence was amplified from total cDNA and tagged using a forward primer with an overhang consisting of a BglII restriction site and 2× hemagglutinin (HA) epitopes, and a reverse primer with a NotI restriction site overhang.

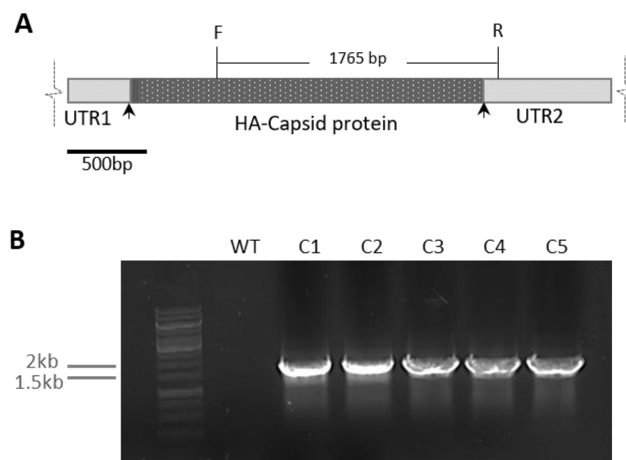


Figure 3.3 PCR verification of pLEXSY-neo2.1-LRV2_CP ligation products. A Map of pLEXSY-neo2.1-LRV2_CP fragment encompassing the vector-insert ligation junctions. Ligation junctions indicated by arrowheads. Primer binding sites for PCR verification of correct ligation product indicated by F and R. **B** Agarose gel electrophoresis of PCR products from 5 transformant colonies, showing amplicons of the expected size. WT, wild type. *Diagram in A constructed by author using Microsoft PowerPoint.*

The resulting PCR product (Figure 3.2B) was digested with BglII and NotI, then ligated with BglII- and NotI-digested pLEXSY-neo2.1 (see vector map, Figure 3.1). The ligation mix was used to transform *E.coli* XL Blue competent cells. Correct ligation products were verified via colony PCR using a forward primer binding to a region within the LRV2 CP sequence and a reverse primer binding to the pLEXSY backbone 93 bp downstream of the

cloning site (Figure 3.3). The resulting plasmid, pLEXSY-neo2.1-LRV2_CP, was sent for Sanger sequencing (Eurofins Genomics, Germany) to ensure that the LRV2 CP sequence was unaltered throughout the course of vector construction.

The CP gene of LRV1-4 (viral resident of *L. guyanensis*) had previously been cloned into pLEXSY-neo2.1 in a related CP-overexpression experiment by Alexandra Zakharova and Andreu Saura, following the same procedure as above. The resulting plasmid, pLEXSY-neo2.1-LRV1-4_CP, was also used as a vector for this project.

CPs of three additional strains of LRV (Table 3.1, white rows) were selected based on their varying phylogenetic distances (based on CP and RDRP protein sequences) from LRV1-4 (Kostygov, Grybchuk, et al., 2021). As the host strains *L. guyanensis* Lg2014, *L. braziliensis* LbrLEM2700, and *L. guyanensis* MHOM/SR/80/CUMC1 were not in Prof Yurchenko's trypanosomatid collection, HA-tagged CP genes from their resident LRVs were commercially synthesised (GeneCust, Boynes, France) and inserted into pLEXSY-neo2.1.

Table 3.1 LRV strains of CP genes used in this study. Strains are listed in increasing order of phylogenetic distance from LRV1-4. Highlighted in grey are strains available in Prof Yurchenko's lab.

LRV strain	NCBI accession number	Host	Shorthand in subsequent figures/text	Source of CP gene
LRV1-4	U01899	*L. guyanensis* M4147 (MHOM/BR/75/M4147)	LRV1-4	Cloned from host culture
LRV1-Lg2014	KY750611	*L. guyanensis* Lg2014	LRV1-2014	Commercially synthesised
LRV1-LbrLEM2700	KX808483	*L. braziliensis* LbrLEM2700	LRV1-2700	Commercially synthesised
LRV1-1	NC_002063	*L. guyanensis* CUMC1 (MHOM/SR/80/CUMC1)	LRV1-1	Commercially synthesised
LRV2-LmjUzb44tg	MZ926705	*L. major* T44g (MRHO/UZ/2003/IsvT44g)	LRV2	Cloned from host culture

3.2 Integration of five LRV CP gene variants into *L. guyanensis* and *L. major* genomic DNA

The pLEXSY-neo2.1 vector derivatives bearing HA-tagged CP genes from five strains of LRV were linearised with SwaI (see vector map, Figure 3.1). Each of the linearised vectors was used for transfection of *L. guyanensis* and *L. major* cultures via electroporation, generating 10 new parasite lines (Table 3.2).

Table 3.2 Newly constructed parasite lines expressing HA-tagged LRV CP from genomic SSU locus. Highlighted in grey are parasite lines constitutively expressing HA-tagged CPs of their resident LRV.

Parasite line	*Leishmania* species + strain	Resident LRV strain	LRV strain of CP expressed from SSU locus
Lguy/LRV1-4	*L. guyanensis* M4147	LRV1-4	LRV1-4
Lguy/LRV1-2014	*L. guyanensis* M4147	LRV1-4	LRV1-2014
Lguy/LRV1-2700	*L. guyanensis* M4147	LRV1-4	LRV1-2700
Lguy/LRV1-1	*L. guyanensis* M4147	LRV1-4	LRV1-1
Lguy/LRV2	*L. guyanensis* M4147	LRV1-4	LRV2
Lmaj/LRV1-4	*L. major* T44g	LRV2	LRV1-4
Lmaj/LRV1-2014	*L. major* T44g	LRV2	LRV1-2014
Lmaj/LRV1-2700	*L. major* T44g	LRV2	LRV1-2700
Lmaj/LRV1-1	*L. major* T44g	LRV2	LRV1-1
Lmaj/LRV2	*L. major* T44g	LRV2	LRV2

Antibiotic selection with three concentrations of G418 (50 µg/mL, 100 µg/mL, 200 ug/mL) was done in 24-well plates, with each well designated as one population of parasites. G418 concentrations were chosen in accordance with transfectant selection strategies previously used for *L. guyanensis* in our lab and for *L. major* in other labs (Beneke & Gluenz, 2019; Zakharova et al., 2022; Zhang et al., 2020). In nine out of ten setups, drug-resistant parasite populations became immediately apparent two weeks after transfection, at which point all parasites in the antibiotic-treated mock-transfected (control) setups had died. For the *L. major* setup transfected with pLEXSY-neo2.1-LRV2, drug-resistant parasites emerged one week after the death of the control parasites.

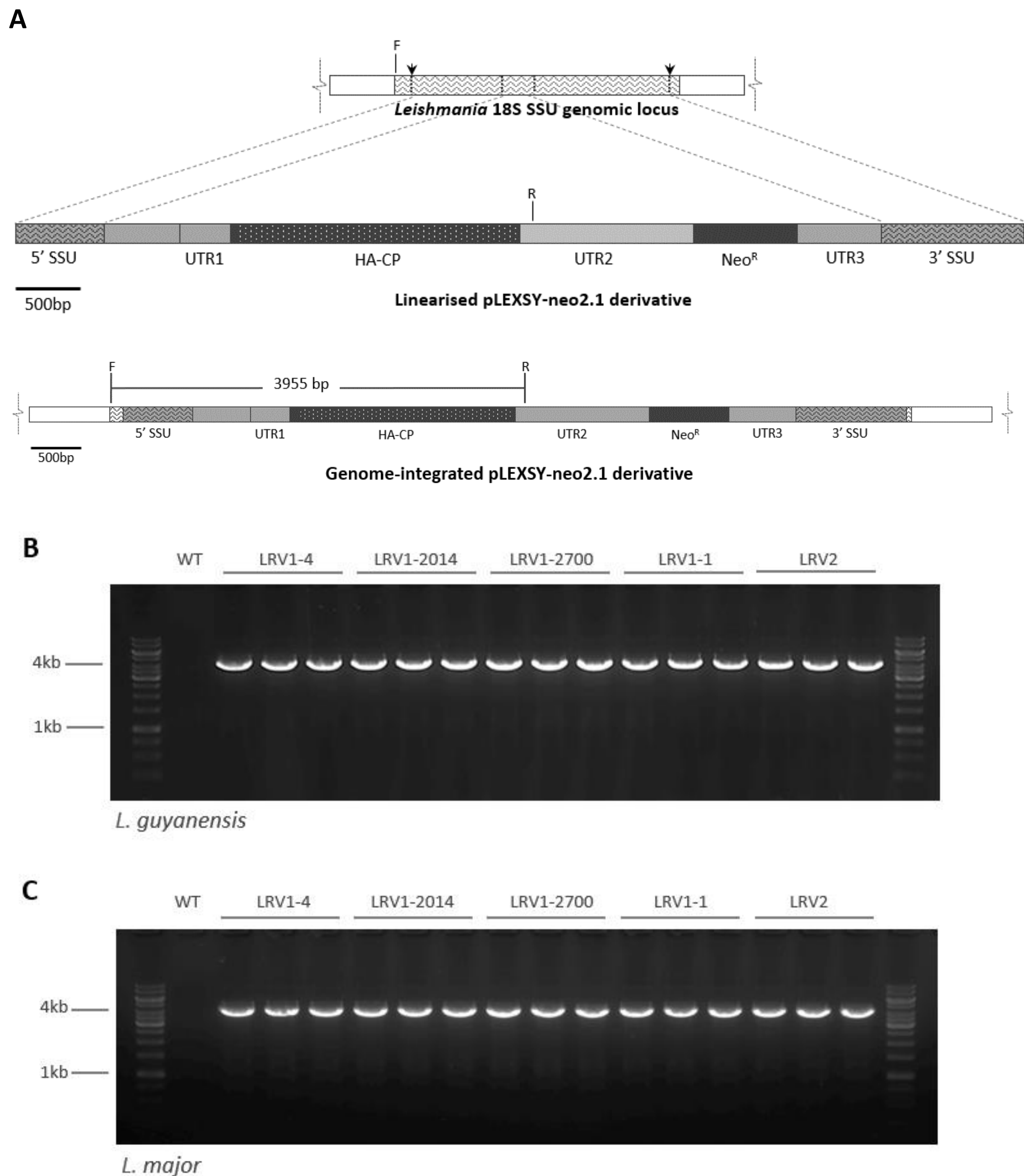


Figure 3.4. Integration of HA-CP harbouring pLEXSY-neo2.1 derivative into *Leishmania* genomic SSU locus. A SwaI-linearised HA-CP-bearing pLEXSY-neo2.1 derivative. Homologous recombination (HR) takes place between the *Leishmania* genomic SSU locus and the homology arms at the ends of the linearised vector (HR sites indicated by dotted lines). Arrowheads indicate the outer limits of the HR sites. F indicates the binding site of the forward primer upstream of the 5′ HR site; R indicates the binding site of the reverse primer within the pLEXSY vector backbone. **B** PCR showing genomic integration of pLEXSY-neo2.1-LRV1-4_CP, pLEXSY-neo2.1-LRV1_LRV1-2014_CP, pLEXSY-neo2.1-LRV1_LRV1-2700_CP, pLEXSY-neo2.1-LRV1-1_CP, and pLEXSY-neo2.1-LRV2_CP in *L. guyanensis*, 3 biological replicates per integration construct. **C** PCR demonstrating genomic integration of pLEXSY-neo2.1-LRV1-4_CP, pLEXSY-neo2.1-LRV1_2014_CP, pLEXSY-neo2.1-LRV1_LRV1-2700_CP, pLEXSY-neo2.1-LRV1-1_CP, and pLEXSY-neo2.1-LRV2_CP in *L. major* T44g, 3 biological replicates per integration construct. *Diagram in A constructed by author using Microsoft PowerPoint.*

Drug-resistant populations were passaged thrice in G418-supplemented medium, then thrice in the absence of antibiotics. PCRs were carried out on genomic DNA of three populations per parasite line to ascertain the correct genomic integration of pLEXSY-neo2.1 derivatives (Figure 3.4). Amplicons were sent for Sanger sequencing to verify the identities and sequence integrity of the genome-integrated LRV CPs.

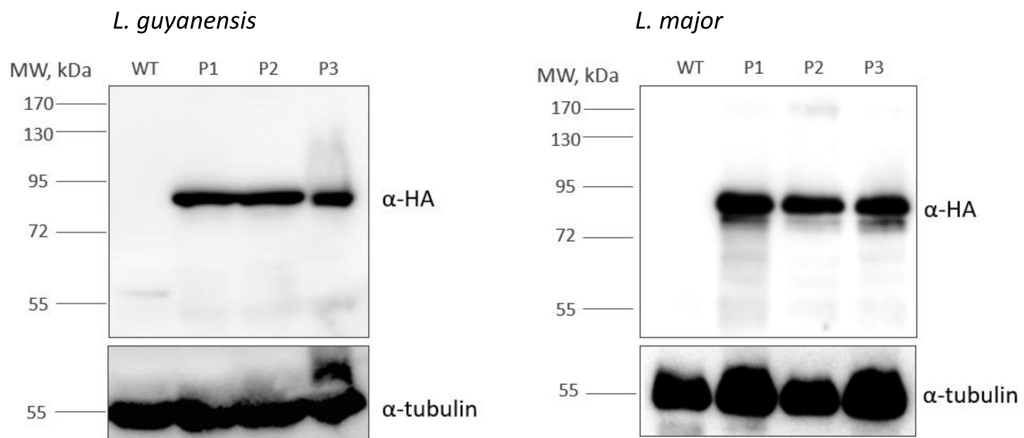
3.3 Confirmation of expression of genome-integrated LRV CPs

Western blot analyses were carried out to ascertain expression and correct folding of HA-tagged LRV CPs. Blots were carried out in batches, as the transfection vectors were not ready at the same time and transfection was carried out in batches.

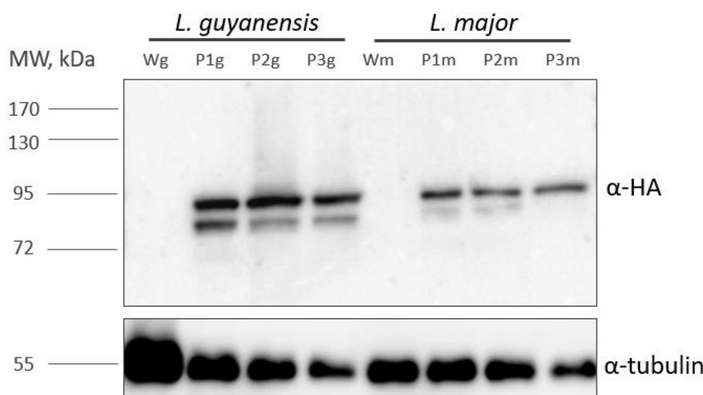
Western blots using α-HA antibodies on lysates of Lguy/LRV1-4 and Lmaj/LRV1-4 parasite lines detected bands of the expected size, 82kDa (Figure 3.5A). Bands of the correct size were also detected in lysates of Lguy/LRV2 and Lmaj/LRV2 (Figure 3.5B). However, a second band of lower molecular weight and lower intensity was also detected in these lysates.

When lysates from Lguy/LRV1-2700, Lguy/LRV1-2014, and Lguy/LRV1-1 parasite lines were subjected to Western blot analysis using α-HA antibodies, bands of the expected size were detected in the Lguy/LRV1-2700 and Lguy/LRV1-2014 lanes, but not in the Lguy/LRV1-1 lanes (Figure 3.5C). To investigate whether LRV1-1 CP was expressed at low levels by the Lguy/LRV1-1 parasite lines, the membrane was imaged twice more at higher exposure times. At the highest exposure time (5 minutes), faint bands were visible but their intensities were not much higher than background bands.

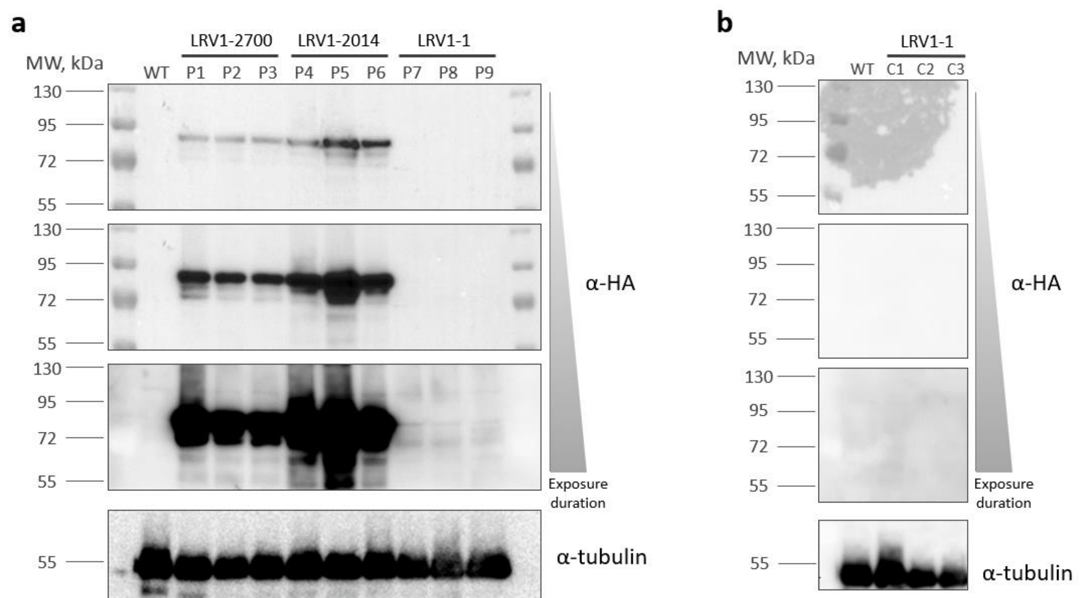
A LRV1-4 CP expression in *L. guyanensis* and *L. major*



B LRV2 CP expression in *L. guyanensis* and *L. major*



C. LRV1-2700 CP, LRV1-2014 CP, and LRV1-1 CP expression in *L. guyanensis*



D LRV1-1 CP, LRV1-2014 CP, and LRV1-2700 CP expression in *L. major*

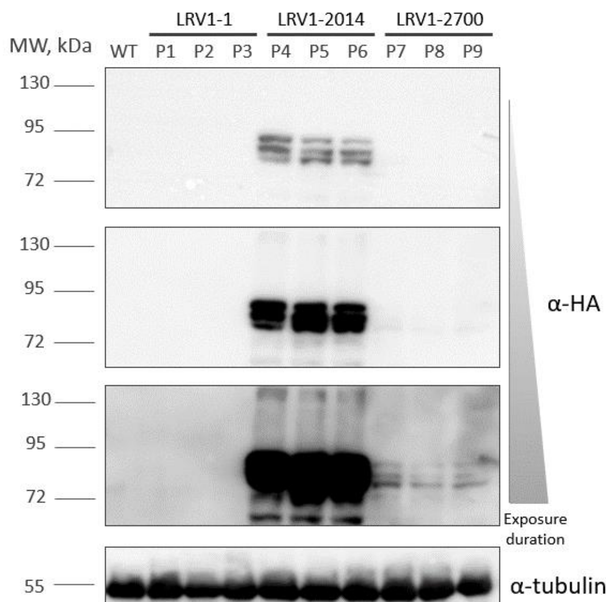


Figure 3.5 Western blot analysis of expression of 5 LRV CPs in *L. guyanensis* and *L. major*. Western blots with α-HA antibodies to detect HA-tagged CPs and α-tubulin antibodies (loading control). Each lane was loaded with lysates from 10^7 parasites. **A** Expression of HA-tagged LRV1-4 CP in wild type *L. guyanensis* (WT), 3 populations of Lguy/LRV1-4 (P1-P3), wild type *L. major* (WT, right), and 3 populations of Lmaj/LRV1-4 (P1-P3). **B** Expression of HA-tagged LRV2 CP in wild type *L. guyanensis* (Wg), 3 populations of Lguy/LRV2 (P1g, P2g, P3g), wild type *L. major* (Wm), and 3 populations of Lmaj/LRV2 (P1m, P2m, P3m). **C (a)** Expression of HA-tagged LRV CPs in wild type *L. guyanensis* (WT), 3 populations of Lguy/LRV1-2700 (P1-P3), 3 populations of Lguy/LRV1-2014 CP (P4-P6), and 3 populations of Lguy/LRV1-1 (P7-P9). Membrane probed with α-HA antibodies imaged at 3 exposure times. **(b)** Expression of HA-tagged LRV1-1 CP in wild type *L. guyanensis* (WT) and 3 clonal populations of Lguy/LRV1-1 (C1-C3). Membrane probed with α-HA antibodies imaged at 3 exposure times. **D** Expression of HA-tagged LRV CPs in wild type *L. major* (WT), 3 populations of Lmaj/LRV1-1 (P1-P3), 3 populations of Lmaj/LRV1-2014 (P4-P6), and 3 populations of Lmaj/LRV1-2700 CP (P7-P9). Membrane probed with α-HA antibodies imaged at 3 exposure times.

It was hypothesized that the low or absent expression of LRV1-1 CP may be attributed to the variable number of pLEXSY constructs integrated into SSU loci among non-clonal Lguy/LRV1-1 populations. To test this hypothesis, one population of Lguy/LRV1-1 was spread onto a G418-supplemented 1% agar M199 plate. Three clonal colonies were subsequently picked, cultured, and passaged thrice without antibiotics. The genome-integrated LRV1-1 CP genes in the clonal lines were verified via PCR, and the amplicons were sent for Sanger sequencing to confirm their sequence integrity (Figure 3.6). Western

blot analysis using α-HA antibodies was performed on lysates from the three clonal Lguy/LRV1-1 lines (Figure 3.5C(b)), but no LRV1-1 CP was detected at any of the exposure durations. The absence of LRV1-1 CP bands was not attributed to sample paucity or loading issues, as the levels of tubulin in the Lguy/LRV1-1 lanes were comparable to those of the wild type.

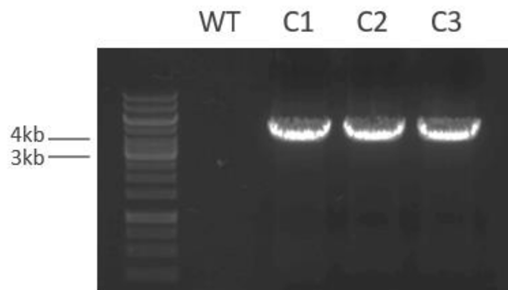


Figure 3.6 PCR verification of genomic integration of pLEXSY-neo2.1-LRV1-1_CP in three clonal lines. Primers as per Figure 3.4.

A Western blot with α-HA antibodies on lysates of Lmaj/LRV1-1, Lmaj/LRV1-2014, and Lmaj/LRV1-2700 detected bands of the expected size in the Lmaj/LRV1-2014 lanes at the lowest exposure, and in Lmaj/LRV1-2700 lanes at higher exposures (Figure 3.5C). Interestingly, each lane featured a triplet of bands of comparable intensity that migrated close to each other. As with *L. guyanensis*, no bands were detected in the Lmaj/LRV1-1 lanes.

3.4 Transcript levels of genome-integrated exogenous LRV CP genes

To ascertain whether the lack of LRV1-1 CP expression in both species and the overall variability in protein levels among different CP variants were due to transcriptional variability from having dissimilar number of CP genes integrated into SSU loci, RT-qPCR was carried out using strain-specific LRV CP primers (Figure 3.7). This was not done for Lguy/LRV1-4 and Lmaj/LRV2, as RT-qPCR using LRV1-4 CP primers in *L. guyanensis* or LRV2 CP primers in *L. major* would also bind to the +ssRNA transcripts and genomic

dsRNA of their respective LRVs and confound the quantification of CP transcripts from SSU loci. Transcript levels of LRV CPs measured by RT-qPCR were normalised against transcript levels of kinetoplastid membrane protein 11 (*kmp11*), a highly conserved housekeeping gene.

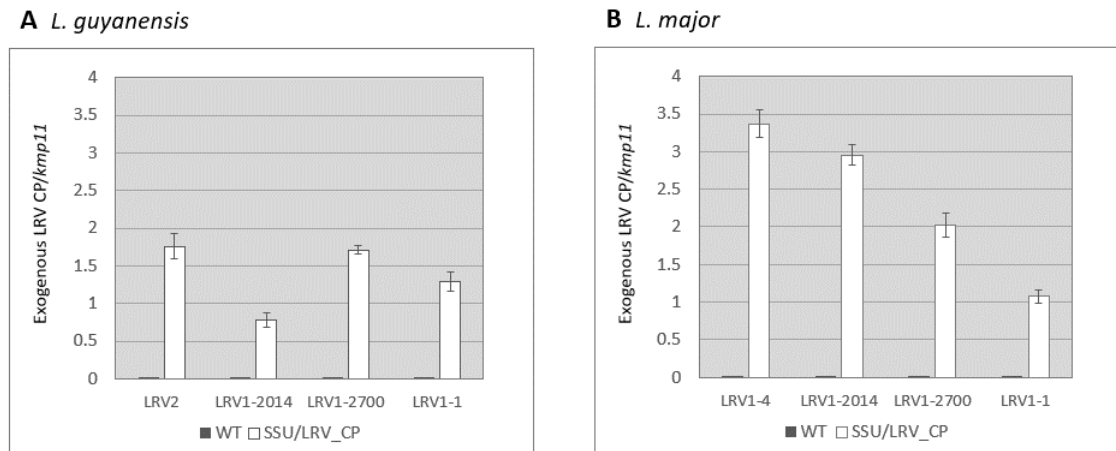


Figure 3.7 Transcript levels of genome-integrated, exogenous LRV CPs in *L. guyanensis* and *L. major*. Bar charts are clustered by CP variant (source strain indicated in X-axis). For each cluster, the dark grey bar shows the expression level (mean of 3 technical replicates) of the CP variant in wild-type parasites; the white bar shows the expression level (mean of 3 biological replicates × 3 technical replicates) of the CP variant in parasite lines harbouring that CP gene in their SSU. CP levels are normalised against *kmp11* levels. *Graphs constructed using Microsoft Excel.*

Among *L. guyanensis* parasite lines, there was transcriptional variability indicative of genomic copy number variability, however, no apparent correlation could be seen between transcript levels of exogenous LRV CP variants measured by RT-qPCR and protein levels detected by Western blot (Figure 3.7A). Lguy/LRV1-2014 parasite lines, which had the lowest levels of CP transcripts, had fairly robust CP expression in α-HA Western blots (Figure 3.5C). Lguy/LRV1-2700 lines had twice as much CP transcripts as Lguy/LRV1-2014 lines, despite having fainter CP bands than Lguy/LRV1-2014 lines in Western blot. The same is true for Lguy/LRV1-1 lines – they had significantly higher CP transcript levels than Lguy/LRV1-2014 lines despite having no (or barely) detectable CP bands in Western blot.

Among *L. major* lines, there appeared to be a correlation between LRV CP variant transcript level and protein expression. The *L. major* lines (excepting Lmaj/LRV1-4, which was on a separate blot) listed in order of α-HA Western blot band intensities are: Lmaj/LRV1-2014, Lmaj/LRV1-2700, Lmaj/LRV1-1. The same pattern can be seen in RT-qPCR transcript levels (Figure 3.7B).

Transcript levels of exogenous CPs in *L. major* were, in general, higher than in *L. guyanensis*, whereas protein levels were higher in *L. guyanensis.*

3.5 Effect of expression of LRV CPs on viral copy numbers of resident LRVs

To check whether overexpression of the 5 LRV CP variants in *L. guyanensis* and *L. major* affected the average copy number of their resident LRVs, RT-qPCR was carried out (Figure 3.8). For each strain of resident LRV, two pairs of primers were designed: one that amplifies a fragment within the 5′ UTR of the LRV genome, and another that amplifies a fragment within the coding sequence of RDRP (Figure 3.8 A and D). The 5′ UTR primers would detect only the genomic dsRNA of LRV1-4 and LRV2, while the RDRP primers would detect only endogenous RDRP transcripts or genomic dsRNA. None of the primer pairs would detect LRV1-4 or LRV2 CP transcripts from SSU loci, which makes them suitable for approximation of LRV copy numbers across all parasite lines.

In *L. guyanensis,* expression of all 5 LRV CP variants markedly diminished LRV1-4 copy number in both qPCR experiments. However, LRV1-1 CP – which was not (or barely) detectable in Western blot – failed to eliminate LRV1-4 completely (Figure 3.8 B and C).

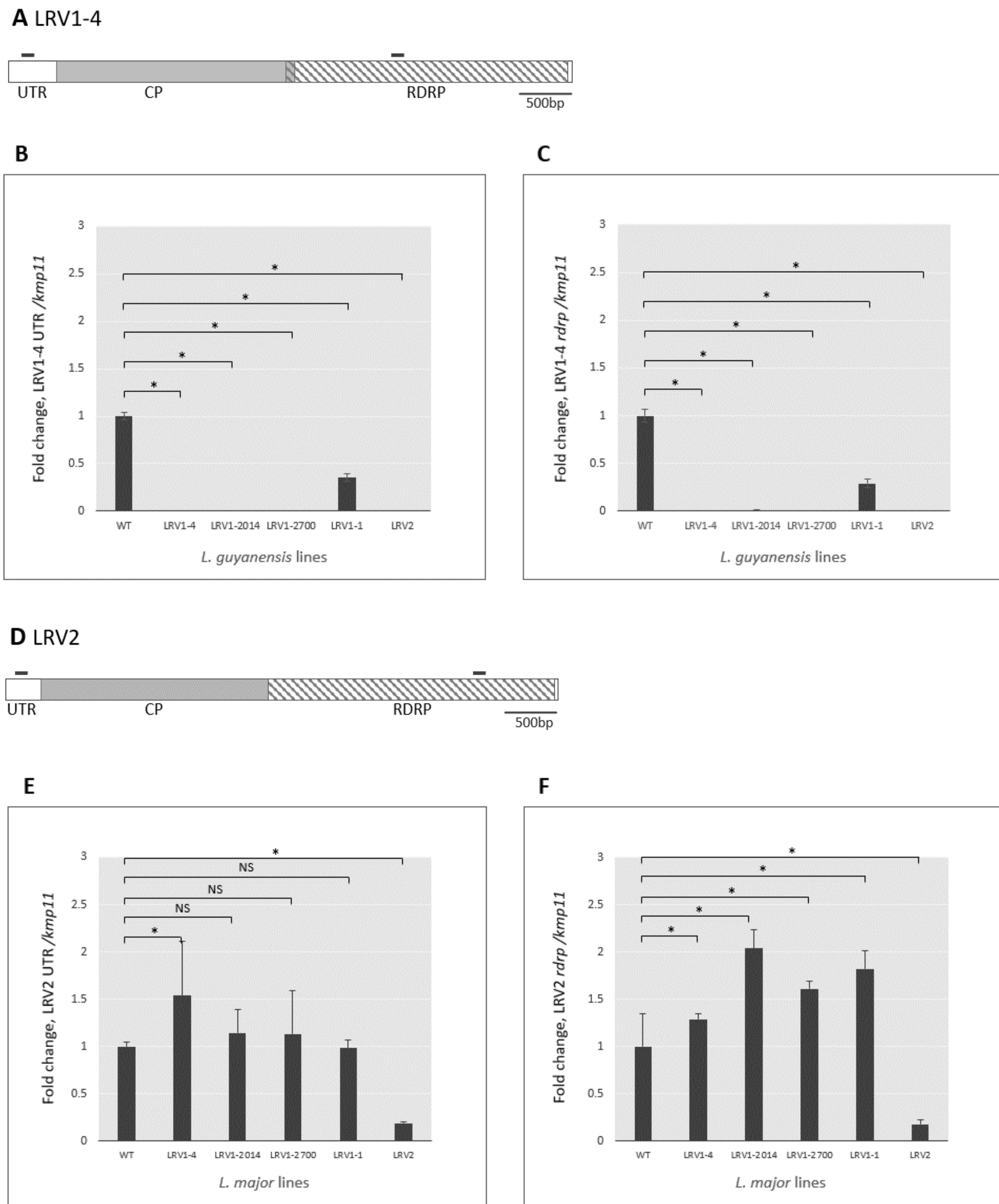


Figure 3.8 Effect of expression of LRV CPs on LRV1-4 copy number in *L. guyanensis* and LRV2 copy number in *L. major*. **A** Genome map of LRV1-4, showing locations (bars above map) of 5′UTR and RDRP fragments amplified by qPCR. **B** LRV1-4 copy number in *L. guyanensis* WT and 5× *L. guyanensis* lines expressing LRV CP variants from SSU, as measured by qPCR of LRV1-4 5′UTR fragment shown in A. **C** As in B, but using primers for LRV1-4 RDRP fragment in A. **D** Genome map of LRV2, showing locations of 5′UTR and RDRP fragments amplified by qPCR. **E** LRV2 copy number in *L. major* WT and *L. major* lines expressing various LRV CP variants from SSU locus, as measured by qPCR of LRV2 5′UTR fragment in D. **F** As in E, but using primers for LRV2 RDRP fragment in D. *Bars show mean measurements from 3 biological replicates × 3 technical repeats, normalised against* kmp11 *levels;* kmp11*-normalised WT levels set as 1. P-values indicated as follows: *, p-value < 0.01; NS, not significant. Diagrams constructed by author using Microsoft PowerPoint; graphs via Microsoft Excel.*

In *L. major*, expression of the 4 exogenous LRV CPs either did not affect (Figure 3.8 E, 5′UTR primers) or increased (Figure 3.8 F, RDRP primers) LRV2 copy number, depending on the primer pair used for RT-qPCR. Expression of the CP of its resident virus, LRV2, significantly diminished, but did not eliminate, LRV2 copy number.

CHAPTER 4. DISCUSSION

4.1 Integration of HA-tagged LRV CPs into *L. guyanensis* and *L. major* genomic DNA

Construction of the pLEXSY-neo2.1-LRV2_CP vector for genomic integration of HA-tagged LRV2 CP was a fairly straightforward process. Transfection of both *L. guyanensis* and *L. major* with five pLEXSY-derived vectors proceeded successfully, with all ten transfection setups yielding a number of G418-resistant populations. PCR and Sanger sequencing confirmed that the constructs integrated into the correct genomic locations. However, the emergence of antibiotic-resistant populations in *L. major* transfected with the vector bearing LRV2 CP was strikingly slower (by a full week) than in the *L. major* cultures transfected with the other four vectors. This indicates that, compared to the other four setups, the Lmaj/LRV2 setup had either a significantly lower number of successful transfectants per population or a significantly lower cell division rate of the transfectants. By contrast, no difference in timelines of drug-resistance emergence was observed among the five *L. guyanensis* cultures transfected with the five different CPs. Therefore, the lower transfection efficiency and/or reduced cell division rate of the Lmaj/LRV2 parasite lines were likely attributable to the HA-tagged LRV2 CP gene itself, as all five transfection vectors were otherwise identical.

4.2 The HA-tagged LRV CPs were expressed at dissimilar levels

CP variant-specific RT-qPCR demonstrated that all five genome-integrated CP genes were transcribed at a good rate.

Among the HA-tagged CPs, only LRV1-1 CP was not (or barely) detected in α-HA Western blots, despite robust LRV1-1 CP transcription in Lguy/LRV1-1 and Lmaj/LRV1-1 parasite lines. This suggests that either HA-tagged LRV1-1 CP was translated at a lower

rate than the other four CPs, or that this recombinant protein was markedly less stable than the other variants and was degraded at a higher rate. The first scenario is unlikely for the following reasons:

1. The same pLEXSY vector – bearing the same UTRs for post-transcriptional maturation and processing of an exogenous gene – was used for all HA-tagged CP expression constructs.
2. The transfection parameters for generating each parasite line were identical. Moreover, because the vectors bearing LRV1-1 CP, LRV1-2014 CP, and LRV1-2700 CP were commercially synthesised and became available at the same time, transfections with these three vectors were carried out simultaneously, using aliquots from the same *L. guyanensis* and *L. major* parasite suspensions. Therefore, the Lguy/LRV1-1 and Lmaj/LRV1-1 parasite lines were unlikely to have experienced unique translation-repressing stresses to which the other lines were not exposed.
3. The four CP variants that were detectable in α-HA Western blots had different band intensities/ protein levels in both *L. guyanensis* and *L. major*. Furthermore, rankings of relative expression (judged by α-HA band intensities relative to tubulin) of the CP variants were the same for both species: LRV1-4 CP was highest, followed by LRV2 CP, LRV1-2014 CP, and LRV1-2700 CP. Roughly similar rankings were observed in repeats of these Western blots[*]. Moreover, for every single parasite line, all three biological replicates had identical banding patterns and band intensities. These suggest that the dissimilar Western blot band intensities across different parasite lines were due

[*] A partial repeat of this Western blot (excluding the LRV1-1-expressing lines) can be found in the Supplementary Information files of the published paper (Appendix C).

to dissimilarities in variant-specific protein stability, rather than stochastic variations in translation rate.

4. The possibility that transcripts of HA-tagged LRV1-1 CP have specific secondary structures that were recognised by translation-repressive RBPs cannot be ruled out. This seems particularly plausible in the case of *L. major*, given that the host of LRV1-1 is of a different species and subgenus than *L. major*. It can be speculated that an RBP in *L. major*, which binds to the foreign LRV1-1 HA-CP mRNA, is absent from or binds a different target in *L. guyanensis* CUMC1, likely as an evolutionary adaptation of *L. guyanensis* CUMC1 to its LRV1-1 endosymbiont. However, this raises the question of why the same response was not triggered by transcripts of HA-tagged LRV1-4 CP, LRV1-2014 CP, and LRV1-2700 CP, all of which were equally foreign to *L. major*. A further argument against this explanation is that *L. guyanensis* M4147 appeared to process transcripts of HA-tagged LRV1-1 CP in the same way as *L. major*, despite its phylogenetic similarity to the host of LRV1-1 – the CUMC1 strain of the same species, *L. guyanensis*.

Overall, these observations suggest that all five recombinant HA-tagged CPs were successfully translated but had dissimilar turnover rates, with LRV1-1 CP having the highest turnover, hence its virtual absence in Western blots. Expression of the same HA-tagged LRV1-1 CP in *L. guyanensis* CUMC1 could clarify this. Unstable recombinant LRV1-1 CP would have high turnover even when expressed in its own host.

The α-HA Western blots showed either one band (Lguy/LRV1-4 and Lmaj/LRV1-4), one major band with one or two lower MW bands (Lguy/LRV2, Lmaj/LRV2, Lguy/LRV1-2014, Lguy/LRV1-2700), or multiple bands of comparable intensity migrating closely together (Lguy/LRV1-1, Lmaj/LRV1-2014, Lmaj/LRV-2700). The lower bands are

possibly degradation intermediates. Alternatively, the higher MW bands could be ubiquitinated (or otherwise post-translationally modified) derivatives of the CPs that earmark them for destruction. The latter theory could be tested via Western blotting with antibodies against ubiquitin.

4.3 Expression of all types of CPs led to elimination or copy number diminution of LRV1-4 in *L. guyanensis*

LRV1-1 CP was barely detected in the α-HA Western blot of lysates from Lguy/LRV1-1 lines (Figure 3.7). It was only visible at the longest exposure duration. Despite its scarcity, it brought about the reduction of endogenous LRV1-4 copy number to roughly a quarter of WT level. This supports the hypothesis that viral elimination by CP overexpression is protein-based and likely via competition: exogenous CPs polymerise with endogenous CPs with or without the inclusion of CP-RDRP. As the combined number of exogenous and endogenous CPs would be in excess of the number needed for proper virion formation, some empty VLPs are assembled which dilute the number of full virions passed on to *Leishmania* progeny with each round of division. The high turnover of LRV1-1 CP means it is never present in quantities that are high enough to completely eliminate LRV1-4.

The four other HA-tagged CPs which were detected in *L. guyanensis* Western blots – LRV1-4 CP, LRV1-2014 CP, LRV1-2700 CP, LRV2 CP – effectively eliminated endogenous LRV1-4 from *L. guyanensis*, adding further support for the theory that the mechanism behind viral elimination is contingent on protein abundance.

The ability of LRV2 CP to eliminate LRV1-4 from *L. guyanensis* was surprising, as the sequence identity between the protein sequences of LRV1-4 CP and LRV2 CP is low – roughly 40% (Table 4.1). There are two possible explanations for this outcome:

1. Despite the low protein sequence identity between LRV1-4 CP and LRV2 CP, they were structurally similar enough to polymerise with each other and form heterogeneous VLPs, and the key residues involved in polymerisation are identical. This bodes well for the CP-based vaccines being developed in mice (Castiglioni et al., 2017), as they would likely be effective against a broad range of LRV strains, not just the LRV1-4 strain from which they were derived. Future studies using immunofluorescence microscopy or coimmunoprecipitation (Chapter 5) could demonstrate whether LRV1-4 and LRV2 CPs coincide on the same capsid. Mutational studies could also be done on residues putatively involved in polymerisation.
2. CP-induced LRV elimination, while protein-based (given that LRV1-1 CP, which was scarcely detected in Western blot, could not eliminate LRV1-4), might not involve polymerisation between endogenous and exogenous CPs at all. Other mechanisms or interactions with heretofore unidentified factors may be at play.

Table 4.1 Identity between pairs of CPs. The value shown in each cell is the percent identity between the protein sequences of the two CP variants specified in the corresponding row and column headings.

	LRV1-4	**LRV1-2014**	**LRV1-2700**	**LRV1-1**
LRV1-2014	95.96			
LRV1-2700	89.81	90.84		
LRV1-1	90.70	92.18	90.84	
LRV2	39.81	40.74	40.59	40.97

4.4 Expression of LRV2 CP led to significant copy number reduction, but not elimination, of LRV2 from *L. major*, while other CPs had no effect

Expression of HA-tagged LRV2 CP significantly reduced endogenous LRV2 copy number to less than 20% of WT level. This is consistent with the results of the Widmer study

wherein episomal expression of LRV2-1 CP in *L. major* 5-ASKH led to a decrease in the copy number of LRV2-1, but not its elimination (Widmer, 1995). Nevertheless, it is noteworthy and surprising that overexpression of LRV2 CP could not fully eliminate LRV2 from its host *L. major* T44g in the present study, considering it was able to eliminate a phylogenetically distant LRV strain, LRV1-4, from a host that is phylogenetically distant from its own, *L. guyanensis*. Failure of episomally-expressed LRV2-1 CP to eliminate LRV2-1 from *L. major* ASKH in the 1995 study was speculated to be due to the variability of exogenous gene expression from an episome; hence, stabler and more robust expression of LRV2 CP from a genomic locus in this study was expected to eliminate endogenous LRV2 from *L. major* T44g. Retention of LRV2 in both studies suggests that *L. major* parasites have additional mechanisms or factors for preventing the loss of their resident LRV that are absent from *L. guyanensis.*

HA-tagged LRV1-1 CP was not detected in α-HA Western blots of lysates from Lmaj/LRV1-1 and, as expected, did not reduce LRV2 copy number. Failure of HA-tagged LRV1-2700 CP to eliminate or reduce LRV2 copy number might also be attributable to low protein level, as LRV1-2700 CP could only be detected in the α-HA Western blot at higher exposures. However, LRV1-4 CP and LRV1-2014 CP, which had robust levels in Western blots, also failed to reduce LRV2 copy number.

On the one hand, the ability of LRV2 CP to decrease LRV2 copy number in *L. major*, along with the inability of all 4 LRV1 CPs to alter LRV2 copy number, support the hypothesis that phylogenetically distant LRV1 CPs cannot polymerise with endogenous LRV2 CPs, and therefore, are unable to disrupt the CP:CP-RDRP ratio. Even the abundant LRV1-4 CPs or LRV1-2014 CPs in the cytoplasm can only assemble into homogeneous, empty VLPs and have no effect on LRV2.

On the other hand, it is perplexing that overexpression of LRV2 CP in *L. guyanensis* could eliminate LRV1-4 from *L. guyanensis*, whereas the reciprocal overexpression of LRV1-4 CP in *L. major* could not eliminate LRV2 from *L. major*. The former suggests that the two CP variants could heteropolymerise, while the latter suggests otherwise.

It is possible that the ability of LRV2 CP and LRV1-4 CP to form heterogeneous capsids is dependent on the activities of unidentified host factors. In *S. cerevisiae,* capsid assembly of the resident L-A virus (from the same family as LRV) requires acetylation of the viral coat protein, gag, by the host's MAK3 N-acetyltransferase (Tercero et al., 1993). Unassembled L-A gag proteins are reportedly unstable and have a high turnover rate (Tercero & Wickner, 1992). It is conceivable that a similar post-translational modification of CPs is necessary for LRV capsid assembly in *Leishmania* spp. It can be speculated that the host factor mediating this PTM in *L. guyanensis* is less specific than its counterpart in *L. major*, allowing it to modify and activate CPs of LRVs other than its own. By contrast, it can be speculated that *L. major*'s LRV2-selective host factor would not modify exogenous LRV1 CPs; consequently, these unmodified LRV1 CPs would be unable to assemble into homogeneous or heterogeneous VLPs. A co-immunoprecipitation experiment using antibodies against LRV CP, followed by mass spectrometric identification of the co-immunoprecipitated proteins, might identify the *Leishmania* host factors involved in modifying and regulating LRV CPs. Further experiments could then be carried out to test the function and specificity of these factors.

4.5 *L. major* and *L. guyanensis* have dissimilar relationships with their resident LRVs

L. major appears to have a more intricate interdependent relationship with its viral resident. Genetic manipulations that diminish LRV2 copy number have a detrimental effect on *L.*

major fitness, as evidenced by the later emergence of Lmaj/LRV2 parasite lines relative to other transfected *L. major* lines, which is indicative of a reduced cell division rate. The inability of LRV2 CP to completely eliminate LRV2 from its own host, *L. major*, in conjunction with its ability to eliminate LRV1-4 from a different host, *L. guyanensis*, suggests that *L. major* has additional and/or alternative mechanisms for LRV2 retention.

By contrast, elimination of LRV1-4 from *L. guyanensis* was achieved with markedly greater facility. The four CPs that had good levels of protein expression, including the phylogenetically distant LRV2 CP, eliminated LRV1-4 from *L. guyanensis*. Despite its high turnover, LRV1-1 CP also diminished LRV1-4 copy number significantly. The loss or copy number reduction of LRV1-4 had no discernible effect on the fitness of CP-expressing *L. guyanensis* parasite lines, which grew as well as wild type *L. guyanensis*.

Subsequent transcriptomic analyses conducted by a colleague (Appendix C) revealed that only two genes showed differential expression between wild-type (WT) and 2CMA-treated (LRV-negative) *L. guyanensis*, whereas 87 genes exhibited differential expression between WT and 2CMA-treated *L. major*. These findings provide additional evidence that *L. guyanensis* has undergone fewer adaptations to LRV presence, and that the loss of LRV is of lower consequence to this species.

Taking everything into account, the relationship between *L. major* and LRV2 appears to be more intimate than that between *L. guyanensis* and LRV1-4. Evidently, *L. major* has made more extensive genetic adaptations to accommodate, coexist with, and maintain its resident virus.

CHAPTER 5. SUMMARY AND FUTURE DIRECTIONS

5.1 Summary

Five transfection vectors were constructed to integrate and constitutively express genes from the genomic SSU locus of *Leishmania* parasites, bearing HA-tagged CP genes from LRV1-4, LRV1-2014, LRV1-2700, LRV1-1, and LRV2. *L. guyanensis* M4147 and *L. major T44g*, which harbour LRV1-4 and LRV2, respectively, were transfected with the five vectors. Successful CP-harbouring parasite lines emerged from all ten transfection setups, although the *L. major* parasite line with genome-integrated LRV2 CP emerged later than the other nine lines.

Western blot analyses with α-HA antibodies confirmed expression of LRV1-4 CP, LRV2 CP, LRV1-2014 CP, and LRV1-2700 CP in both *L. guyanensis* and *L. major*. Although banding patterns and intensities were identical between triplicates of each cell line, they varied widely between the eight types of cell lines. The fifth CP, LRV1-1, was barely detectable in *L. guyanensis*, and completely absent in *L. major*.

CP-variant-specific quantitative RT-PCR demonstrated robust expression of all five recombinant CPs from the SSU locus, suggesting that the dissimilar protein levels were not due to transcriptional insufficiency or failure. Notably, LRV1-1 CP was transcribed at a high level in both *L. guyanensis* and *L. major* despite its paucity or absence in Western blots. This suggests that the recombinant proteins have dissimilar stabilities and consequently dissimilar turnover rates.

Quantitative RT-PCR using primers for LRV1-4 genome sequences outside the CP gene demonstrated that, as hypothesised, expression of all four LRV1 CPs reduced LRV1-4 copy number in *L. guyanensis*. Only LRV1-1, which was barely detected in Western blots, failed to eliminate LRV1-4 completely, supporting the theory that LRV1-4 copy number

reduction is protein-based, possibly achieved by the dilution of CP-RDRP proteins by an excess of CPs. However, contrary to our hypothesis, overexpression of the robustly expressed but phylogenetically distant LRV2 CP also led to the elimination of LRV1-4. This calls for further investigation into heteropolymerisation between LRV1-4 CP and LRV2 CP. Evidence of the ability of LRV2 CP to form heterogeneous capsids with LRV1-4 CP would support our theory that CP-induced LRV1-4 elimination in *L. guyanensis* was due to disruption of the CP:CP-RDRP ratio; failure to detect heterogeneous LRV1-4 CP + LRV2 CP capsids would imply that an alternative mechanism for LRV2 CP-induced LRV1-4 elimination must have been at play.

Quantitative RT-PCR using primers targeting LRV2 genomic sequences outside the CP gene demonstrated that all four LRV1 CPs, even those with high expression levels in Western blots, had no effect on LRV2 copy number. Only LRV2 CP was able to diminish LRV2 copy number in *L. major*, and even this variant was unable to completely eliminate LRV2. While the inability of phylogenetically distant LRV1 CPs to reduce LRV2 copy number supports our initial hypothesis, their inability to form heterogeneous capsids with LRV2 CP could be due to other host-specific factors, given that while LRV2 CP was able to eliminate LRV1-4 from *L. guyanensis*, LRV1-4 CP was unable to bring about a reciprocal reduction in LRV2 copy number in *L. major*.

The reduced cell division rate of the Lmaj/LRV2 parasite line, the insusceptibility of LRV2 to the presence of any LRV1 CP variant, and the inability of LRV2 CP to completely eliminate LRV2 suggest that *L. major* has a more intimate relationship with its LRV2 resident than *L. guyanensis* has with LRV1-4. *L. major* appears to have evolved additional mechanisms for retention of LRV2, as reduction of LRV2 copy number negatively impacts this parasite.

5.2 Future Directions

5.2.1 Expressing HA-tagged LRV1-1 CP in *L. guyanensis* CUMC1

To verify whether the non-detection of HA-tagged LRV1-1 CP in α-HA Western blots was due to translational repression or instability of the recombinant protein, *L. guyanensis* CUMC1 could be transfected with the vector used in this study for expressing HA-tagged LRV1-1 CP. As *L. guyanensis* CUMC1 is the natural host of LRV1-1, it should not have mechanisms for translational repression of LRV1-1 CP transcripts. Absence or paucity of HA-tagged LRV1-1 CP in α-HA Western blot of lysates from *L. guyanensis* CUMC1 with genome-integrated pLEXSY-neo2.1-LRV1-1_CP would prove that the expression construct could not produce a stable recombinant protein. Conversely, the detection of a good level of LRV1-1 CP in α-HA Western blots would suggest that variant-specific translational repression took place in *L. guyanensis* M4147 and *L. major* Tg44.

5.2.2 Re-tagging of LRV1-1 CP

While RT-qPCR showed transcription at a comparable level to other variants, N-terminal HA-tagged LRV1-1 CP was not, or barely, detected in Western blots of lysates from Lguy/LRV1-1 and Lmaj/LRV1-1 parasite lines. This suggests that the tagged protein may have been misfolded or otherwise unstable. A few spacer amino acids between the HA tag and CP might help rectify the instability. Re-tagging LRV1-1 CP with HA at the C-terminal, rather than the N-terminal, might also produce a more stable protein. Alternatively, the N-terminal HA tag could be replaced with FLAG. If expressing a more stable tagged version of LRV1-1 CP led to elimination of LRV1-4 from *L. guyanensis* like the other CP subtypes, it would buttress the explanation put forward by this study that the elimination of LRV1-4 was due to exogenous CPs polymerising with native CPs and CP-RDRPs, leading to disruption of the normal CP:CP-RDRP ratio.

5.2.3 Immunofluorescence microscopy to detect formation of heterogeneous capsids

HA-tagged LRV1-4 CP could be co-expressed in *L. guyanensis* with FLAG-tagged versions of CPs from other LRV strains. The parasites could then be incubated with fluorochrome-linked anti-HA and anti-FLAG antibodies. Immunofluorescence microscopy would reveal whether fluorescence from HA-tagged CPs and FLAG-tagged CPs coincide on a common capsid, which would indicate heterogeneity of capsid composition. This would support the protein-based CP-RDRP dilution hypothesis. This would be especially illuminating in the case of exogenous LRV2 CP, which eliminated endogenous LRV1-4 despite low sequence identity with the latter's CP. A similar experiment could be carried out in *L. major* to illuminate why LRV1-4 CP overexpression did not affect LRV2 copy number.

5.2.4 Expression of CP-RDRP from the genomic SSU locus

Overexpression of CP alone eliminates LRV1-4 by disrupting the CP:CP-RDRP ratio. It would be interesting to see the effect of increasing the other component of that ratio: CP-RDRP. An HA-tagged CP-RDRP gene could be inserted into *L. guyanensis* SSU. The stop codon at the end of the CP coding sequence would be removed to ensure that the CP is not expressed by itself, and a base-pair would be removed to eliminate the need for ribosomal frameshifting.

There are two possible outcomes of CP-RDRP overexpression.

1. Due to its size, no more than 2 copies of the fusion protein can get incorporated into each capsid, so a higher number of CP-RDRP should not alter the composition of full capsids. An excess of CP-RDRP would not disrupt the CP:CP-RDRP ratio in fully assembled virions and should therefore not eliminate LRV1-4 from *L. guyanensis*.

2. On the other hand, CP-RDRP nucleates capsid assembly, so an excess of it might lead to numerous partially-formed open capsids that could not be completed due to insufficient quantities of CP. This might also lead to elimination of LRV1-4, as it would not be able to complete its replication cycle.

5.2.5 Incorporation of full LRV1-4 cDNA into the genomic SSU locus

Recently, the cDNA of the full genome of a totivirus-like virus was incorporated into the genomic DNA of its host, which jeopardised the latter's fitness (Wang et al., 2022). Incorporating the full LRV1-4 genome into *L. guyanensis* would theoretically lead to the production of CP and CP-RDRP of the correct ratio. This should not eliminate native LRV1-4, and may even increase the LRV1-4 copy number. It would be interesting to see if the increased viral copy number would affect the fitness or virulence of *L. guyanensis*, and what compensatory genetic changes *L. guyanensis* might resort to.

5.2.6 Testing the *in vivo* effect of LRV1-4 CP on the virulence of other LRV-bearing *Leishmania* strains

In a study in 2017 (Castiglioni et al.), inoculating mice with adjuvanted recombinant LRV1-4 CP prior to infection with *L. guyanensis* M4147 prevented the progression of cutaneous leishmaniasis to mucocutaneous leishmaniasis. If our CP-RDRP dilution theory is indeed the means by which CP overexpression eliminates LRV, then adjuvanted recombinant LRV1-4 CP would also confer protection to mice infected with other strains of LRV-positive *L. guyanensis* and *L. braziliensis*, provided they're phylogenetically close enough to *L. guyanensis* M4147. Injecting mice with this group's recombinant CPs and subsequently challenging them with a range of LRV-bearing *Leishmania* strains would give *in vivo* proof of the cross-specificity of CP-based viral elimination.

5.2.7 Using HA-CP expressing *Leishmania* lines to determine the effects of various conditions on EV-mediated LRV release

A project I'm currently involved in is experimental infection of LRV-negative trypanosomatid lines with LRV-containing EVs derived from *L. guyanensis* M4147 (presently at the early exploratory stage). As it is widely believed that EVs are released by *Leishmania* upon entry into mammalian blood to deploy virulence factors to mammalian cells, EV production in *Leishmania* EV studies is usually induced via a 4-hour "heat shock" at 37°C. However, one study found that the number of EVs released at 26°C (temperature of the sand fly gut) by *L. amazonensis* was 3-fold higher than at 37°C (Barbosa et al., 2018). While this was believed to be a peculiarity of that species, our current attempts suggest that *L. guyanensis* also releases more vesicles at 26°C than at 37°C.

The HA-CP-expressing *L. guyanensis* lines in this study could be used to determine the effects of temperature, pH, nutrient availability, nutrient type, and other variables on EV-mediated LRV release. Size and number of EVs are usually measured via nanoparticle tracking analysis and Western blotting of EV marker proteins. Additional Western blot with α-HA antibodies could reveal whether release of HA-CP VLPs quantitatively tallies with release of EVs in general, or if certain conditions lead to enrichment of LRVs in EVs.

5.2.8 Identification of *Leishmania* host factors involved in LRV maintenance

The ability of LRV2 CP to eradicate LRV1-4 from *L. guyanensis* and the inability of LRV1-4 CP to alter LRV2 copy number in *L. major* suggest that interactions between exogenous and endogenous CPs might be subject to regulation by one or more unidentified host factors. Apart from this, other interactions between *Leishmania* and LRV – e.g. the hijacking of *Leishmania* enzymes by LRV for its replication, the monitoring of LRV

presence and copy number by the host, the targeting of excess virions for destruction – are likely to be CP-mediated as well. This is because the capsid is the only component of LRV that is accessible to the host for most of the viral replication cycle, with the dsRNA genome of LRV and the two copies of RDRP remaining sequestered in the capsid.

An ongoing project I'm involved in began with a co-immunoprecipitation (co-IP) experiment using the LRV1-4 CP as bait. The proteins pulled down by LRV1-4 CP were identified by LC-MS/MS. I'm currently developing a reverse genetics strategy for knocking out genes pulled down in the co-IP experiment using the CRISPR/Cas9 system. I will subsequently assess the effects of these knockouts on LRV1-4 copy number in *L. guyanensis*. Knocking out an *L. guyanensis* gene involved in keeping LRV1-4 copy number low would lead to an increase in LRV1-4 copy number, whereas knocking out a gene involved in LRV1-4 maintenance would lead to a decrease in LRV1-4 copy number. One such gene might be an enzyme that post-translational modifies LRV1-4 CP, rendering it competent for assembly.

REFERENCES

Abtahi, M., Eslami, G., Cavallero, S., Vakili, M., Hosseini, S. S., Ahmadian, S., Boozhmehrani, M. J., & Khamesipour, A. (2020). Relationship of *Leishmania* RNA virus (LRV) and treatment failure in clinical isolates of *Leishmania major*. *BMC Research Notes*, *13*(1), 126. https://doi.org/10.1186/s13104-020-04973-y

Aitmanaitė, L., Konovalovas, A., Medvedevas, P., Servienė, E., & Serva, S. (2021). Specificity determination in *Saccharomyces cerevisiae* killer kirus systems. *Microorganisms*, *9*(2), 236. https://doi.org/10.3390/microorganisms9020236

Aronson, N., Herwaldt, B. L., Libman, M., Pearson, R., Lopez-Velez, R., Weina, P., Carvalho, E., Ephros, M., Jeronimo, S., & Magill, A. (2017). Diagnosis and treatment of leishmaniasis: Clinical practice guidelines by the Infectious Diseases Society of America (IDSA) and the American Society of Tropical Medicine and Hygiene (ASTMH). *The American Journal of Tropical Medicine and Hygiene*, *96*(1), 24–45. https://doi.org/10.4269/ajtmh.16-84256

Atayde, V. D., da Silva Lira Filho, A., Chaparro, V., Zimmermann, A., Martel, C., Jaramillo, M., & Olivier, M. (2019). Exploitation of the *Leishmania* exosomal pathway by *Leishmania* RNA virus 1. *Nature Microbiology*, *4*(4), 714–723. https://doi.org/10.1038/s41564-018-0352-y

Barbosa, F. M. C., Dupin, T. V., Toledo, M. dos S., Reis, N. F. dos C., Ribeiro, K., Cronemberger-Andrade, A., Rugani, J. N., De Lorenzo, B. H. P., Novaes e Brito, R. R., Soares, R. P., Torrecilhas, A. C., & Xander, P. (2018). Extracellular vesicles released by *Leishmania (Leishmania) amazonensis* promote disease progression and induce the production of different cytokines in macrophages and B-1 cells. *Frontiers in Microbiology*, *9*. https://doi.org/10.3389/fmicb.2018.03056

Bates, P. A. (2007). Transmission of *Leishmania* metacyclic promastigotes by phlebotomine sand flies. *International Journal for Parasitology*, *37*(10), 1097–1106. https://doi.org/10.1016/j.ijpara.2007.04.003

Beneke, T., & Gluenz, E. (2019). LeishGEdit: A method for rapid gene knockout and tagging using CRISPR-Cas9. In J. Clos (Ed.), *Leishmania*: *Methods and Protocols* (pp. 189–210). Humana Press. https://doi.org/10.1007/978-1-4939-9210-2_9

Beneke, T., Madden, R., Makin, L., Valli, J., Sunter, J., & Gluenz, E. (2017). A CRISPR Cas9 high-throughput genome editing toolkit for kinetoplastids. *Royal Society Open Science*, *4*(5), 170095. https://doi.org/10.1098/rsos.170095

Bourreau, E., Ginouves, M., Prévot, G., Hartley, M.-A., Gangneux, J.-P., Robert-Gangneux, F., Dufour, J., Sainte-Marie, D., Bertolotti, A., Pratlong, F., Martin, R., Schütz, F., Couppié, P., Fasel, N., & Ronet, C. (2016). Presence of *Leishmania* RNA virus 1 in *Leishmania guyanensis* increases the risk of first-line treatment failure and symptomatic relapse. *Journal of Infectious Diseases*, *213*(1), 105–111. https://doi.org/10.1093/infdis/jiv355

Bringaud, F., Müller, M., Cerqueira, G. C., Smith, M., Rochette, A., El-Sayed, N. M. A., Papadopoulou, B., & Ghedin, E. (2007). Members of a large retroposon family are

determinants of post-transcriptional gene expression in *Leishmania*. *PLoS Pathogens*, *3*(9), e136. https://doi.org/10.1371/journal.ppat.0030136

Cantanhêde, L. M., da Silva Júnior, C. F., Ito, M. M., Felipin, K. P., Nicolete, R., Salcedo, J. M. V., Porrozzi, R., Cupolillo, E., & Ferreira, R. de G. M. (2015). Further evidence of an association between the presence of *Leishmania* RNA virus 1 and the mucosal manifestations in tegumentary leishmaniasis patients. *PLOS Neglected Tropical Diseases*, *9*(9), e0004079. https://doi.org/10.1371/journal.pntd.0004079

Cantanhêde, L. M., Fernandes, F. G., Ferreira, G. E. M., Porrozzi, R., Ferreira, R. de G. M., & Cupolillo, E. (2018). New insights into the genetic diversity of *Leishmania* RNA virus 1 and its species-specific relationship with *Leishmania* parasites. *PLOS ONE*, *13*(6), e0198727. https://doi.org/10.1371/journal.pone.0198727

Cantanhêde, L. M., Mata-Somarribas, C., Chourabi, K., Pereira da Silva, G., Dias das Chagas, B., de Oliveira R. Pereira, L., Côrtes Boité, M., & Cupolillo, E. (2021). The Maze Pathway of Coevolution: A critical review over the *Leishmania* and its endosymbiotic history. *Genes*, *12*(5), 657. https://doi.org/10.3390/genes12050657

Carrion, R., Ro, Y.-T., & Patterson, J. L. (2008). Leishmaniaviruses. In B.W.J. Mahy & M.H.V. Van Regenmortel (Eds.), *Encyclopedia of Virology* 3rd Edition (pp. 220–224). Elsevier. https://doi.org/10.1016/B978-012374410-4.00566-5

Castiglioni, P., Hartley, M.-A., Rossi, M., Prevel, F., Desponds, C., Utzschneider, D. T., Eren, R.-O., Zangger, H., Brunner, L., Collin, N., Zehn, D., Kuhlmann, F. M., Beverley, S. M., Fasel, N., & Ronet, C. (2017). Exacerbated leishmaniasis caused by a viral endosymbiont can be prevented by immunization with its viral capsid. *PLOS Neglected Tropical Diseases*, *11*(1), e0005240. https://doi.org/10.1371/journal.pntd.0005240

Cohen-Freue, G., Holzer, T. R., Forney, J. D., & McMaster, W. R. (2007). Global gene expression in *Leishmania*. *International Journal for Parasitology*, *37*(10), 1077–1086. https://doi.org/10.1016/j.ijpara.2007.04.011

Curotto de Lafaille, M. A., Laban, A., & Wirth, D. F. (1992). Gene expression in *Leishmania*: analysis of essential 5′ DNA sequences. *Proceedings of the National Academy of Sciences*, *89*(7), 2703–2707. https://doi.org/10.1073/pnas.89.7.2703

da Silva Lira Filho, A., Fajardo, E. F., Chang, K. P., Clément, P., & Olivier, M. (2022). *Leishmania* exosomes/extracellular vesicles containing GP63 are essential for enhanced cutaneous leishmaniasis development upon co-inoculation of *Leishmania amazonensis* and its exosomes. *Frontiers in Cellular and Infection Microbiology*, *11*. https://doi.org/10.3389/fcimb.2021.709258

Dantas-Torres, F. (2007). The role of dogs as reservoirs of *Leishmania* parasites, with emphasis on *Leishmania (Leishmania) infantum* and *Leishmania (Viannia) braziliensis*. *Veterinary Parasitology*, *149*(3–4), 139–146. https://doi.org/10.1016/j.vetpar.2007.07.007

d'Avila-Levy, C. M., Boucinha, C., Kostygov, A., Santos, H. L. C., Morelli, K. A., Grybchuk-Ieremenko, A., Duval, L., Votýpka, J., Yurchenko, V., Grellier, P., & Lukeš, J. (2015). Exploring the environmental diversity of kinetoplastid flagellates in the high-throughput DNA sequencing era. *Memórias Do Instituto Oswaldo Cruz*, *110*(8), 956–965. https://doi.org/10.1590/0074-02760150253

de Carvalho, R. V. H., Lima-Junior, D. S., da Silva, M. V. G., Dilucca, M., Rodrigues, T. S., Horta, C. V., Silva, A. L. N., da Silva, P. F., Frantz, F. G., Lorenzon, L. B., Souza, M. M., Almeida, F., Cantanhêde, L. M., Ferreira, R. de G. M., Cruz, A. K., & Zamboni, D. S. (2019). *Leishmania* RNA virus exacerbates leishmaniasis by subverting innate immunity via TLR3-mediated NLRP3 inflammasome inhibition. *Nature Communications*, *10*(1), 5273. https://doi.org/10.1038/s41467-019-13356-2

Douanne, N., Dong, G., Amin, A., Bernardo, L., Blanchette, M., Langlais, D., Olivier, M., & Fernandez-Prada, C. (2022). *Leishmania* parasites exchange drug-resistance genes through extracellular vesicles. *Cell Reports*, *40*(3), 111121. https://doi.org/10.1016/j.celrep.2022.111121

Du, R., Hotez, P. J., Al-Salem, W. S., & Acosta-Serrano, A. (2016). Old World cutaneous leishmaniasis and refugee crises in the Middle East and North Africa. *PLOS Neglected Tropical Diseases*, *10*(5), e0004545. https://doi.org/10.1371/journal.pntd.0004545

Espada, C. R., Ferreira, B. A., Ortiz, P. A., Uliana, S. R. B., & Coelho, A. C. (2021). Full nucleotide sequencing of ribosomal DNA internal transcribed spacer of *Leishmania* species causing cutaneous leishmaniasis in Brazil and its potential for species typing. *Acta Tropica*, *223*, 106093. https://doi.org/10.1016/j.actatropica.2021.106093

Esteban, R., Fujimura, T., & Wickner, R. B. (1988). Site-specific binding of viral plus single-stranded RNA to replicase-containing open virus-like particles of yeast. *Proceedings of the National Academy of Sciences*, *85*(12), 4411–4415. https://doi.org/10.1073/pnas.85.12.4411

Fernández-Prada, C., Douanne, N., Minguez-Menendez, A., Pena, J., Tunes, L. G., Pires, D. E. V., & Monte-Neto, R. L. (2019). Repurposed molecules: A new hope in tackling neglected infectious diseases. In K. Roy (Ed.), *In Silico Drug Design* (pp. 119–160). Elsevier. https://doi.org/10.1016/B978-0-12-816125-8.00005-5

Ferreira, T. R., Dowle, A. A., Parry, E., Alves-Ferreira, E. V. C., Hogg, K., Kolokousi, F., Larson, T. R., Plevin, M. J., Cruz, A. K., & Walrad, P. B. (2020). PRMT7 regulates RNA-binding capacity and protein stability in *Leishmania* parasites. *Nucleic Acids Research*, *48*(10), 5511–5526. https://doi.org/10.1093/nar/gkaa211

Fujimura, T., Ribas, J. C., Makhov, A. M., & Wickner, R. B. (1992). Pol of gag–pol fusion protein required for encapsidation of viral RNA of yeast L-A virus. *Nature*, *359*(6397), 746–749. https://doi.org/10.1038/359746a0

González, C., Wang, O., Strutz, S. E., González-Salazar, C., Sánchez-Cordero, V., & Sarkar, S. (2010). Climate change and risk of leishmaniasis in North America: Predictions from ecological niche models of vector and reservoir species. *PLoS*

Neglected Tropical Diseases, *4*(1), e585. https://doi.org/10.1371/journal.pntd.0000585

Gossage, S. M., Rogers, M. E., & Bates, P. A. (2003). Two separate growth phases during the development of *Leishmania* in sand flies: implications for understanding the life cycle. *International Journal for Parasitology*, *33*(10), 1027–1034. https://doi.org/10.1016/S0020-7519(03)00142-5

Gradoni, L. (2018). A brief introduction to leishmaniasis epidemiology. In F. Bruschi and L. Gradoni (Eds.), *The Leishmaniases: Old Neglected Tropical Diseases* (pp. 1–13). Springer International Publishing. https://doi.org/10.1007/978-3-319-72386-0_1

Grünebast, J., & Clos, J. (2020). *Leishmania*: Responding to environmental signals and challenges without regulated transcription. *Computational and Structural Biotechnology Journal*, *18*, 4016–4023. https://doi.org/10.1016/j.csbj.2020.11.058

Handman, E., & Bullen, D. V. R. (2002). Interaction of *Leishmania* with the host macrophage. *Trends in Parasitology*, *18*(8), 332–334. https://doi.org/10.1016/S1471-4922(02)02352-8

Hartley, M.-A., Ronet, C., Zangger, H., Beverley, S. M., & Fasel, N. (2012). *Leishmania* RNA virus: when the host pays the toll. *Frontiers in Cellular and Infection Microbiology*, *2*. https://doi.org/10.3389/fcimb.2012.00099

Hillman, B. I., & Cohen, A. B. (2021). Totiviruses (Totiviridae). In B.W.J. Mahy & M.H.V. Van Regenmortel (Eds.), *Encyclopedia of Virology* 3rd Edition (pp. 648–657). Elsevier. https://doi.org/10.1016/B978-0-12-809633-8.21347-2

Ives, A., Ronet, C., Prevel, F., Ruzzante, G., Fuertes-Marraco, S., Schutz, F., Zangger, H., Revaz-Breton, M., Lye, L.-F., Hickerson, S. M., Beverley, S. M., Acha-Orbea, H., Launois, P., Fasel, N., & Masina, S. (2011). *Leishmania* RNA virus controls the severity of mucocutaneous leishmaniasis. *Science*, *331*(6018), 775–778. https://doi.org/10.1126/science.1199326

Karamysheva, Z. N., Gutierrez Guarnizo, S. A., & Karamyshev, A. L. (2020). Regulation of translation in the protozoan parasite *Leishmania*. *International Journal of Molecular Sciences*, *21*(8), 2981. https://doi.org/10.3390/ijms21082981

Kaye, P., & Scott, P. (2011). Leishmaniasis: complexity at the host–pathogen interface. *Nature Reviews Microbiology*, *9*(8), 604–615. https://doi.org/10.1038/nrmicro2608

Koch, L. K., Kochmann, J., Klimpel, S., & Cunze, S. (2017). Modeling the climatic suitability of leishmaniasis vector species in Europe. *Scientific Reports*, *7*(1), 13325. https://doi.org/10.1038/s41598-017-13822-1

Kostygov, A. Y., Grybchuk, D., Kleschenko, Y., Chistyakov, D. S., Lukashev, A. N., Gerasimov, E. S., & Yurchenko, V. (2021). Analyses of *Leishmania*-LRV co-phylogenetic patterns and evolutionary variability of viral proteins. *Viruses*, *13*(11), 2305. https://doi.org/10.3390/v13112305

Kostygov, A. Y., Karnkowska, A., Votýpka, J., Tashyreva, D., Maciszewski, K., Yurchenko, V., & Lukeš, J. (2021). Euglenozoa: taxonomy, diversity and ecology, symbioses and viruses. *Open Biology*, *11*(3). https://doi.org/10.1098/rsob.200407

Kuhlmann, F. M., Robinson, J. I., Bluemling, G. R., Ronet, C., Fasel, N., & Beverley, S. M. (2017). Antiviral screening identifies adenosine analogs targeting the endogenous dsRNA *Leishmania* RNA virus 1 (LRV1) pathogenicity factor. *Proceedings of the National Academy of Sciences*, *114*(5). https://doi.org/10.1073/pnas.1619114114

Lopinski, J. D., Dinman, J. D., & Bruenn, J. A. (2000). Kinetics of ribosomal pausing during programmed −1 translational frameshifting. *Molecular and Cellular Biology*, *20*(4), 1095–1103. https://doi.org/10.1128/MCB.20.4.1095-1103.2000

Lukeš, J., Mauricio, I. L., Schönian, G., Dujardin, J.-C., Soteriadou, K., Dedet, J.-P., Kuhls, K., Tintaya, K. W. Q., Jirků, M., Chocholová, E., Haralambous, C., Pratlong, F., Oborník, M., Horák, A., Ayala, F. J., & Miles, M. A. (2007). Evolutionary and geographical history of the *Leishmania donovani* complex with a revision of current taxonomy. *Proceedings of the National Academy of Sciences*, *104*(22), 9375–9380. https://doi.org/10.1073/pnas.0703678104

Lukša, J., Ravoitytė, B., Konovalovas, A., Aitmanaitė, L., Butenko, A., Yurchenko, V., Serva, S., & Servienė, E. (2017). Different metabolic pathways are involved in response of *Saccharomyces cerevisiae* to L-A and M viruses. *Toxins*, *9*(8), 233. https://doi.org/10.3390/toxins9080233

Lye, L.-F., Owens, K., Shi, H., Murta, S. M. F., Vieira, A. C., Turco, S. J., Tschudi, C., Ullu, E., & Beverley, S. M. (2010). Retention and loss of RNA interference pathways in trypanosomatid protozoans. *PLoS Pathogens*, *6*(10), e1001161. https://doi.org/10.1371/journal.ppat.1001161

Machado, G. U., Prates, F. V., & Machado, P. R. L. (2019). Disseminated leishmaniasis: clinical, pathogenic, and therapeutic aspects. *Anais Brasileiros de Dermatologia*, *94*(1), 9–16. https://doi.org/10.1590/abd1806-4841.20198775

Malvolti, S., Malhame, M., Mantel, C. F., Le Rutte, E. A., & Kaye, P. M. (2021). Human leishmaniasis vaccines: Use cases, target population and potential global demand. *PLOS Neglected Tropical Diseases*, *15*(9), e0009742. https://doi.org/10.1371/journal.pntd.0009742

Martínez-Calvillo, S., Sunkin, S. M., Yan, S., Fox, M., Stuart, K., & Myler, P. J. (2001). Genomic organization and functional characterization of the *Leishmania major* Friedlin ribosomal RNA gene locus. *Molecular and Biochemical Parasitology*, *116*(2), 147–157. https://doi.org/10.1016/S0166-6851(01)00310-3

Miles, S. A., Conrad, S. M., Alves, R. G., Jeronimo, S. M. B., & Mosser, D. M. (2005). A role for IgG immune complexes during infection with the intracellular pathogen *Leishmania*. *Journal of Experimental Medicine*, *201*(5), 747–754. https://doi.org/10.1084/jem.20041470

Moafi, M., Rezvan, H., Sherkat, R., & Taleban, R. (2019). *Leishmania* vaccines entered in clinical trials: A review of literature. *International Journal of Preventive Medicine*, *10*(1), 95. https://doi.org/10.4103/ijpvm.IJPVM_116_18

Molyneux, D. H. (1974). Virus-like particles in *Leishmania* parasites. *Nature*, *249*(5457), 588–589. https://doi.org/10.1038/249588a0

Nandan, D., Thomas, S. A., Nguyen, A., Moon, K.-M., Foster, L. J., & Reiner, N. E. (2017). Comprehensive identification of mRNA-binding proteins of *Leishmania donovani* by interactome capture. *PLOS ONE*, *12*(1), e0170068. https://doi.org/10.1371/journal.pone.0170068

Nocua, P. A., Ramirez, C. A., Requena, J. M., & Puerta, C. J. (2017). *Leishmania braziliensis* SCD6 and RBP42 proteins, two factors with RNA binding capacity. *Parasites & Vectors*, *10*(1), 610. https://doi.org/10.1186/s13071-017-2557-y

Olías-Molero, A. I., de la Fuente, C., Cuquerella, M., Torrado, J. J., & Alunda, J. M. (2021). Antileishmanial drug discovery and development: time to reset the model? *Microorganisms*, *9*(12), 2500. https://doi.org/10.3390/microorganisms9122500

Olivier, M., & Zamboni, D. S. (2020). *Leishmania Viannia guyanensis*, LRV1 virus and extracellular vesicles: a dangerous trio influencing the faith of immune response during muco-cutaneous leishmaniasis. *Current Opinion in Immunology*, *66*, 108–113. https://doi.org/10.1016/j.coi.2020.08.004

Ponte-Sucre, A., Gamarro, F., Dujardin, J.-C., Barrett, M. P., López-Vélez, R., García-Hernández, R., Pountain, A. W., Mwenechanya, R., & Papadopoulou, B. (2017). Drug resistance and treatment failure in leishmaniasis: A 21st century challenge. *PLOS Neglected Tropical Diseases*, *11*(12), e0006052. https://doi.org/10.1371/journal.pntd.0006052

Procházková, M., Füzik, T., Grybchuk, D., Falginella, F. L., Podešvová, L., Yurchenko, V., Vácha, R., & Plevka, P. (2021). Capsid structure of *Leishmania* RNA virus 1. *Journal of Virology*, *95*(3). https://doi.org/10.1128/JVI.01957-20

Rath, C. T., Vivarini, Á. C., dos Santos, J. V., Medina, J. M., Saliba, A. M., Mottram, J. C., Lima, A. P. C. A., Calegari-Silva, T. C., Pereira, R. M., & Lopes, U. G. (2022). Toll-Like Receptor 3 (TLR3) is engaged in the intracellular survival of the protozoan parasite *Leishmania (Leishmania) amazonensis*. *Infection and Immunity*, *90*(9). https://doi.org/10.1128/iai.00324-22

Ro, Y. T., Scheffter, S. M., & Patterson, J. L. (1997). Hygromycin B resistance mediates elimination of *Leishmania* virus from persistently infected parasites. *Journal of Virology*, *71*(12), 8991–8998. https://doi.org/10.1128/jvi.71.12.8991-8998.1997

Roberts, S. C. (2011). The genetic toolbox for *Leishmania* parasites. *Bioengineered Bugs*, *2*(6), 320–326. https://doi.org/10.4161/bbug.2.6.18205

Robinson, J. I., & Beverley, S. M. (2018). Concentration of 2′C-methyladenosine triphosphate by *Leishmania guyanensis* enables specific inhibition of *Leishmania* RNA virus 1 via its RNA polymerase. *Journal of Biological Chemistry*, *293*(17), 6460–6469. https://doi.org/10.1074/jbc.RA117.001515

Ronet, C., Beverley, S. M., & Fasel, N. (2011). Muco-cutaneous leishmaniasis in the New World. *Virulence*, *2*(6), 547–552. https://doi.org/10.4161/viru.2.6.17839

Scott, P., & Novais, F. O. (2016). Cutaneous leishmaniasis: immune responses in protection and pathogenesis. *Nature Reviews Immunology*, *16*(9), 581–592. https://doi.org/10.1038/nri.2016.72

Shapira, M., Zilka, A., Garlapati, S., Dahan, E., Dahan, I., & Yavesky, V. (2001). Post transcriptional control of gene expression in Leishmania. *Medical Microbiology and Immunology*, *190*(1–2), 23–26. https://doi.org/10.1007/s004300100073

Shrivastava, R., Drory-Retwitzer, M., & Shapira, M. (2019). Nutritional stress targets LeishIF4E-3 to storage granules that contain RNA and ribosome components in *Leishmania*. *PLOS Neglected Tropical Diseases*, *13*(3), e0007237. https://doi.org/10.1371/journal.pntd.0007237

Silverman, J. M., & Reiner, N. E. (2012). Leishmania exosomes deliver preemptive strikes to create an environment permissive for early infection. *Frontiers in Cellular and Infection Microbiology*, *1*. https://doi.org/10.3389/fcimb.2011.00026

Stevens, A., Muratore, K., Cui, Y., Johnson, P. J., & Zhou, Z. H. (2021). Atomic structure of the *Trichomonas vaginalis* double-stranded RNA virus 2. *MBio*, *12*(2). https://doi.org/10.1128/mBio.02924-20

Steverding, D. (2017). The history of leishmaniasis. *Parasites & Vectors*, *10*(1), 82. https://doi.org/10.1186/s13071-017-2028-5

Tarr, P. I., Aline, R. F., Smiley, B. L., Scholler, J., Keithly, J., & Stuart, K. (1988). LR1: a candidate RNA virus of *Leishmania*. *Proceedings of the National Academy of Sciences*, *85*(24), 9572–9575. https://doi.org/10.1073/pnas.85.24.9572

Tercero, J. C., Dinman, J. D., & Wickner, R. B. (1993). Yeast MAK3 N-acetyltransferase recognizes the N-terminal four amino acids of the major coat protein (gag) of the L-A double-stranded RNA virus. *Journal of Bacteriology*, *175*(10), 3192–3194. https://doi.org/10.1128/jb.175.10.3192-3194.1993

Tercero, J. C., & Wickner, R. B. (1992). MAK3 encodes an N-acetyltransferase whose modification of the L-A gag NH2 terminus is necessary for virus particle assembly. *Journal of Biological Chemistry*, *267*(28), 20277–20281. https://doi.org/10.1016/S0021-9258(19)88697-0

Torres-Guerrero, E., Quintanilla-Cedillo, M. R., Ruiz-Esmenjaud, J., & Arenas, R. (2017). Leishmaniasis: a review. *F1000Research*, *6*, 750. https://doi.org/10.12688/f1000research.11120.1

Ubeda, J.-M., Légaré, D., Raymond, F., Ouameur, A., Boisvert, S., Rigault, P., Corbeil, J., Tremblay, M. J., Olivier, M., Papadopoulou, B., & Ouellette, M. (2008). Modulation of gene expression in drug resistant *Leishmania* is associated with gene amplification, gene deletion and chromosome aneuploidy. *Genome Biology*, *9*(7), R115. https://doi.org/10.1186/gb-2008-9-7-r115

Ubeda, J.-M., Raymond, F., Mukherjee, A., Plourde, M., Gingras, H., Roy, G., Lapointe, A., Leprohon, P., Papadopoulou, B., Corbeil, J., & Ouellette, M. (2014). Genome-wide stochastic adaptive DNA amplification at direct and inverted DNA repeats in the Parasite *Leishmania*. *PLoS Biology*, *12*(5), e1001868. https://doi.org/10.1371/journal.pbio.1001868

Velez, R., & Gállego, M. (2020). Commercially approved vaccines for canine leishmaniosis: a review of available data on their safety and efficacy. *Tropical Medicine & International Health*, *25*(5), 540–557. https://doi.org/10.1111/tmi.13382

Wang, H., de Matos Filipe, D., & Okamoto, K. (2022). A full-length infectious cDNA clone of a dsRNA totivirus-like virus. *Virology*, *576*, 127–133. https://doi.org/10.1016/j.virol.2022.09.010

Wickner, R. B., Ghabrial, S., Nibert, M., Patterson, J., & Wang, C. (2012). Totiviridae. In A.M.Q. King, M.J. Adams, E.B. Carstens, & E.J. Lefkowitz (Eds.) *Virus Taxonomy* (pp. 639–650). Elsevier. https://doi.org/10.1016/B978-0-12-384684-6.00052-5

Widmer, G. (1995). Suppression of *Leishmania* RNA virus replication by capsid protein overexpression. *Journal of Virology*, *69*(7), 4122–4126. https://doi.org/10.1128/jvi.69.7.4122-4126.1995

Widmer, G., & Dooley, S. (1995). Phylogenetic analysis of *Leishmania* RNA virus and *Leishmania* suggests ancient virus-parasite association. *Nucleic Acids Research*, *23*(12), 2300–2304. https://doi.org/10.1093/nar/23.12.2300

World Health Organization. (2022). *Leishmaniasis Fact Sheet*. https://www.who.int/news-room/fact-sheets/detail/leishmaniasis. Accessed 28 March 2023.

Yanik, M., Gurel, M. S., Simsek, Z., & Kati, M. (2004). The psychological impact of cutaneous leishmaniasis. *Clinical and Experimental Dermatology*, *29*(5), 464–467. https://doi.org/10.1111/j.1365-2230.2004.01605.x

Yao, W., Muqtadir, K., & Bruenn, J. A. (1995). Packaging in a yeast double-stranded RNA virus. *Journal of Virology*, *69*(3), 1917–1919. https://doi.org/10.1128/jvi.69.3.1917-1919.1995

Zakharova, A., Albanaz, A. T. S., Opperdoes, F. R., Škodová-Sveráková, I., Zagirova, D., Saura, A., Chmelová, L., Gerasimov, E. S., Leštinová, T., Bečvář, T., Sádlová, J., Volf, P., Lukeš, J., Horváth, A., Butenko, A., & Yurchenko, V. (2022). *Leishmania guyanensis* M4147 as a new LRV1-bearing model parasite: Phosphatidate phosphatase 2-like protein controls cell cycle progression and intracellular lipid content. *PLOS Neglected Tropical Diseases*, *16*(6), e0010510. https://doi.org/10.1371/journal.pntd.0010510

Zangger, H., Ronet, C., Desponds, C., Kuhlmann, F. M., Robinson, J., Hartley, M.-A., Prevel, F., Castiglioni, P., Pratlong, F., Bastien, P., Müller, N., Parmentier, L., Saravia, N. G., Beverley, S. M., & Fasel, N. (2013). Detection of *Leishmania* RNA

virus in *Leishmania* parasites. *PLoS Neglected Tropical Diseases*, *7*(1), e2006. https://doi.org/10.1371/journal.pntd.0002006

Zhang, W.-W., Karmakar, S., Gannavaram, S., Dey, R., Lypaczewski, P., Ismail, N., Siddiqui, A., Simonyan, V., Oliveira, F., Coutinho-Abreu, I. V., DeSouza-Vieira, T., Meneses, C., Oristian, J., Serafim, T. D., Musa, A., Nakamura, R., Saljoughian, N., Volpedo, G., Satoskar, M., Nakhasi, H. L. (2020). A second generation leishmanization vaccine with a markerless attenuated *Leishmania major* strain using CRISPR gene editing. *Nature Communications*, *11*(1), 3461. https://doi.org/10.1038/s41467-020-17154-z

Zijlstra, E., Musa, A., Khalil, E., El Hassan, I., & El-Hassan, A. (2003). Post-kala-azar dermal leishmaniasis. *The Lancet Infectious Diseases*, *3*(2), 87–98. https://doi.org/10.1016/S1473-3099(03)00517-6

Zirpel, H., & Clos, J. (2019). Gene replacement by homologous recombination. In J. Clos (Ed.), *Leishmania: Methods and Protocols* (pp. 169–188). Humana Press. https://doi.org/10.1007/978-1-4939-9210-2_8

APPENDIX A. PRIMERS

Primer pairs used in this study. Restriction sites underlined. Hemagglutinin sequences rendered in bold.

Primer name	Sequence	Target	Purpose
BglII_HA_LRV2_CP_F	GCCACCAGATCTATG**TACCCATACGATGTTCCAGATTACGCT**GCT**TACCCATACGATGTTCCAGATTACGCT**ACTTCACCGATTCCACAAATGAA	LRV2 CP 5' end	Making HA-tagged LRV2 CP, 2228 bp
NotI_LRV2_CP_R	AGGAGGGCGGCCGCTTACACATTTTCGCCTGGTGG	LRV2 CP 3' end	
LRV2_CP_qPCR_F	ACCTACCGAACGCACTCAA	LRV2 CP, internal	Checking junction between LRV2 CP and pLEXSY backbone, 1765 bp
pLEXSY_insert_R	CATCTATAGAGAAGTACACGTAAAAG	pLEXSY backbone, downstream of cloning site	
SSU_int_F	CATATGCTTGTTTCAAGGAC	*Leishmania* SSU, upstream of pLEXSY 5'SSU HR site	Checking junction between *Leishmania* genomic DNA and pLEXSY backbone, 3955 bp
pLEXSY_insert_R	CATCTATAGAGAAGTACACGTAAAAG	pLEXSY backbone, downstream of cloning site	
LRV1-4_UTR_qPCR_F	CTGACTGGACGGGGGGTAAT	5'UTR of LRV1-4 genome	Measuring LRV1-4 copy number in *L. guyanensis*
LRV1-4_UTR_qPCR_R	CAAAACACTCCCTTACGC	5'UTR of LRV1-4 genome	
LRV1-4_RDRP_qPCR_F	GTCTGTTTCGTACCCGCCG	LRV1-4 RDRP gene	Measuring LRV1-4 copy number in *L. guyanensis*
LRV1-4_RDRP_qPCR_R	AAGCTCAGGATGTGCATGTTCCA	LRV1-4 RDRP gene	
LRV2_UTR_qPCR_F	CGACCCAGTGTTCAATAGTTGC	5'UTR of LRV2 genome	Measuring LRV2 copy number in *L. major*
LRV2_UTR_qPCR_R	TTCCGGCATCTACAGGTGTTAG	5'UTR of LRV2 genome	
LRV2_RDRP_qPCR_F	CTTGGACCGACGAGAAGCA	LRV2 RDRP gene	Measuring LRV2 copy number in *L. major*
LRV2_RDRP_qPCR_R	TCGCTTCGCAACGCTTAATG	LRV2 RDRP gene	
LRV1-4_CP_qPCR_F	GTCGCAAGCCTATGTTTTGC	LRV1-4 CP transcript	Measuring LRV1-4 CP transcript level
LRV1-4_CP_qPCR_R	AATCGCGCATCATACACACG	LRV1-4 CP transcript	
LRV2_CP_qPCR_F	ACCTACCGAACGCACTCAA	LRV2 CP transcript	Measuring LRV2 CP transcript level
LRV2_CP_qPCR_R	GGGCTGGTACGTTGTATGC	LRV2 CP transcript	
LRV1-Lg2014_CP_qPCR_F	TGAGCAATGCCGACAGTTTG	Lg2014 CP transcript	Measuring Lg2014 CP transcript level
LRV1-Lg2014_CP_qPCR_R	TACTTGCGTCACATGCACAC	Lg2014 CP transcript	
LRV1-LBr2700_CP_qPCR_F	TCAGATTGCGCCCATTCAAC	Lbr2700 CP transcript	Measuring Lbr2700 CP transcript level
LRV1-LBr2700_CP_qPCR_R	TTCGCAGCATTACAGCATCG	Lbr2700 CP transcript	
LRV1-1_CP_qPCR_F	ATGAATGTGATGGCGCCATG	LRV1-1 CP transcript	Measuring LRV1-1 CP transcript level
LRV1-1_CP_qPCR_R	AAGAATTGGTGGCAGCAAGC	LRV1-1 CP transcript	
kmp11_F	GCCTGGATGAGGAGTTCAACA	*kmp11* transcript	Measuring *kmp11* transcript level for qPCR normalisation
kmp11_R	GTGCTCCTTCATCTCGGG	*kmp11* transcript	

APPENDIX B. PLASMIDS

Plasmid name	Features	Eukaryotic gene/s	Source
pLEXSY-neo2.1	*Leishmania* SSU integration vector; pBluescript II KS(-) origin of replication Ori, Amp^R	Neo^R	Jena Bioscience
pLEXSY-neo2.1-LRV1-4_CP	*Leishmania* SSU integration vector; pBluescript II KS(-) origin of replication Ori, Amp^R	Neo^R, HA-tagged LRV1-4 capsid protein	Alexandra Zakharova, Ostrava University
pLEXSY-neo2.1-LRV1-1_CP	*Leishmania* SSU integration vector; pBluescript II KS(-) origin of replication Ori, Amp^R	Neo^R, HA-tagged LRV1-1 capsid protein	Synthesised by GeneCust for this study
pLEXSY-neo2.1-LRV1-2014_CP	*Leishmania* SSU integration vector; pBluescript II KS(-) origin of replication Ori, Amp^R	Neo^R, HA- tagged LRV1-Lg2014 capsid protein	Synthesised by GeneCust for this study
pLEXSY-neo2.1-LRV1-2700_CP	*Leishmania* SSU integration vector; pBluescript II KS(-) origin of replication Ori, Amp^R	Neo^R, HA- tagged LRV1-LBrLEM2700 capsid protein	Synthesised by GeneCust for this study
pLEXSY-neo2.1-LRV2_CP	*Leishmania* SSU integration vector; pBluescript II KS(-) origin of replication Ori, Amp^R	Neo^R, HA-tagged LRV2 capsid protein	This study

Sequences of plasmid constructed in this study (CP sequence rendered in bold)**:**

pLEXSY-neo2.1-LRV1-1_CP

```
   1 aaattggata acttggcgaa acgccaagct aatacatgaa ccaaccgggt gttctccact
  61 ccagacggtg ggcaaccatc gtcgtgagac gcccagcgaa tgaatgacag taaaaccaat
 121 gccttcactg gcagtaacac ccagcagtgt tgactcaatt cattccgtgc gaaagccggc
 181 ttgttccggc gtcttttgac gaacaactgc cctatcagct ggtgatggcc gtgtagtgga
 241 ctgccatgca tggcgttgac gggagcgggg gattagggtt cgattccgga gagggagcct
 301 gagaaatagc taccacttct acggagggca gcaggcgcgc aaattgccca atgtcaaaac
 361 aaaacgatga ggcagcgaaa agaaatagag ttgtcagtcc atttggattg tcatttcaat
 421 gggggatatt taaacccatc caatatcgag taacaattgg aggacaagtc tggtgccagc
 481 acccgcggta attccagctc caaaagcgta tattaatgct gttgctgtta aagggttcgt
 541 agttgaactg tgggctgtgc aggtttgttc ctggtcgtcc cgtccatgtc ggatttggtg
 601 acccaggccc ttgcagcccg tgaacattca aagaaacaag aaacacggga gtggttcctt
 661 tcctgattta cgcatgtcat gcatgcgtcg aggatctggg cccgaggtct gcgattgacg
 721 taggagttgc aagggggagg gggtatgaac ggggtgggta gagctttttt aggtggaagt
 781 agtgagaggg tgggcttgag agaatttgag gtgtgttcgt gatgtgtgga tcttatcggg
 841 ggctcggttg agtttttggt ttggtgattt tgtaggggtg aagggcggtg gaagggttgc
 901 tgtgtgcgtg ccattccgtg aagcgtcatt cctatgtccg acactgaaat gcgtccgcga
 961 ttgtgtggat tgtgtggaag tgtctgggcg acgactgaat tgagaagggg agtcgaaacg
1021 gtgcgtggat gccgtgtttg tgctacacac aagcaaaggc aacaacgtaa gcgcatcgag
1081 ccagacgaac caacccaata ttcacacaac tgtacgtggc agcgcgctgt atcgttggaa
1141 agacgatgcg tgtgcgcgtg ccaccaccat gtacaagttt ccagcctaac tgcttgtata
1201 ccttgtcttt tgcgttgtgc ctgcgcgtgc tgtgtacctc tgctcgtgtg tccagatcca
1261 gtcgcagcct gaccgcatca cacatcaagg cgttacagcc tctccgtctt tgctgcaccc
1321 aagtgtattc gtgtgaccgc gttgagggat ctaatggatt cggaactttg gtcgtttggt
```

1381 ctctcatgct gcccaatccc atgcgctttt ctccgtgcct ctctgtctcc ccttccgttc
1441 tatttggaca acgatgtgta ctgtaatgcg tgcgccatct gacaacgaat agatcagcag
1501 cattcgcaca cttgcacata cccagtgaag cttttgtgtc tgtcgtattg acaacaccga
1561 ctgcaacaag gtgtagatag aagttggcct tctcgctcgc tcgcacgctc ttcacgctcc
1621 tgctttcctt gctgtgcctt gccaccagat ct**atgtaccc atacgatgtt ccagattacg**
1681 ctgcttaccc atacgatgtt ccagattacg cttctgatac accaaactct gataaagtcg
1741 cgtgtggaag gaagcccatg ttctgcgaga ttataaaact cgcgaacagg aagcggttga
1801 ttttggatac aactggtgag cgagtgtacg acgcccgtat aaattactgc tcaacggcag
1861 atgcgaccgt tcaagccgat tgccatatat actggcgctt gaagctgcga cgcactgatg
1921 ccgtttttga agaatatact ggtcagggct acgcgctcga cactgcttca tacccccaac
1981 agtatgctga tatcatcagg gggtattaca gtaagcatgt gtcaagtgca ctagctgcca
2041 acacccaaca ctttgtgaac gttttggcta tgttacgtca tgctggcgct tgtattgcgc
2101 attattgcat gacagggaaa attgattttg acatcttaag caaaaagaaa cataagaata
2161 aagaagttgt gacactaagc agtgctgaca gtttgagttt tttgcctcac tcggctctat
2221 atctgccttc agcgttacgt gcaagcgatc cagaaatatt taatgttttg tacctgcttg
2281 gttgtgcttg tgatgcaagc atagcaatgg acaacattgc gaacacgagt ggggcagcta
2341 aatatgcgat gcctcattat aacccactgc aactatctca cgctctacat gtaaccatat
2401 tctatatgct gggcttgatg gacagttgcg ggtatgggga agatgcagtc ctggctttaa
2461 cggccggcct gcattctgtc acaacggtga tagcacacag tgatgaagga gctataactc
2521 gtgatgcttt aagagagctc tcatacacac agccttatgg cacaatgccg gtgccgattg
2581 caggttactt tcagcacata aatgtgctgt ttacaacaca gcctgcttgg gatcaatttg
2641 cagggatttg ggattatgtc atcttggcaa ctgctgcatt ggttcatctt tctgatcctg
2701 gaatgacagt gaatgacgtc acatatccga ctacactcac tacaaaagtt gcatcagtag
2761 atggccgtaa cagtgacata gcagcacaga tgatgcacag tgccacacgg ttctgtgata
2821 tatacattga gaacttatgt ctcttttggg gtgtggtcgc gaacccagac gggaatgcga
2881 gccaagcttt actgcacgca ttcaatgtgg ttgcttgtgg cacagagcca aacaggcatc
2941 ttgaaatgaa tgtgatggcg ccatggtatt gggttgagag ttcagcacta ttcagcgact
3001 actcccaatt tcgatcacca atctcctcag cgggatatgg gccgcaaagc atctacggtg
3061 cacgtttggt gcttgctgcc accaattctt tggaattcac gggggaagca ggcgattact
3121 cagcataccg ttttgagtgg acgacaatgc ggcacaaccc actatttaac attctcaata
3181 aacgcgtcgg tgatggcctt gctaatgtgg atttcaggct gcgaccgttt aatgagtggt
3241 tgctagaagg ccaaccgagc cgtcgaagct gtgatgctgt tggccacggc acacctacgg
3301 ctacatgtaa ccacaagact ccgaaccgta acacagtgga cgagtttatt tggggcagca
3361 cttcctgtga cctctttcat ccagctgaac taacttcata caccgccgtg tgtgttaggt
3421 tcaggaatta tttgtctgaa gccgacggtg atgtacgcgt actaaacaca cccactaggg
3481 aggttattga ggggaatgtt gttacaaggt gtgacggcat acgctgcttg gacagcaaca
3541 agcgcataca acatgttcct gaggtcgcca gacgttattg catgatggca aggtatttgc
3601 aagcgcgtac ctttggagcg ttaactatag gtgacgacat aatacgtgga tttgataagg
3661 tcgataaggt tgtgaaaatg cataagagca ataatagatt aggtcaaatg cccttaatag
3721 atgtaactgg attgtgtcag ccaatgattg aaacgagtac tgctcgtgct tcgatgacta
3781 ctcgcattga tccaaacaga cttgctgcag ccacagctcg tatagaacta ccactagcac
3841 ctcgttgcac gtcctccctg attccgtcat ccgatgcagt gcccgaacca gaaccacatg
3901 ttggtgaacc gggcaatggt agtggctgcg cttgagcggc cgccctcctc ctcctttctt
3961 gttcctttca cgtcgccttc tcggttgtag ctggcagacg acgagtctta cttttacgtg
4021 tacttctcta tagatgatgt atgatctctc tgcatgcgtg ttcgtgcatg tgtccgtgtg
4081 ttgtgtacgc gtgcgtctcg cctcagctct ccgcgtgaaa gggtttgact gcccatgatg
4141 cgtgtgtata tccacgcgca ggcacgcaca cacacacaca cacacacaca cacaggcaca
4201 cacaggcaca caaacgcatc tcaggccgag ccgcatacgt ctctcgcacg gtctcgttta
4261 tttgatcatg tagttgagta ttaaattggg aagacaaaaa cataatagca cgaagagtcg
4321 ggcacgaaaa gcccgatctc tctctctctc tctcgtgcgc gcaggggcgt gtgggtgcac
4381 gacgacgagg acggaggggg gaagggaggc ataaacggat cgatccccca tcgcacatgc
4441 gggtacgagc attatctgct gtgtctggtc ctttatcata tcgctacccg cccgcccccc
4501 gccccctcc cccccccccc cccgcggctc ctgtcactgt cgttcctgcg actccccaca
4561 cacccgcgca acgctcatgt cacgaagaga aaagtagaag gcgtggcgat gcgttgcgcg
4621 gccccctgtc cgctggcatg cagacgtcga gccgtggaac ggatggtagc gagagacaac
4681 gcggcgatgc aggaaaggag atttactgca ggacgtctac acacgcgcgc acaccactgc
4741 gaatggcacc gtgctgagga aaattggggg ggagggcgca cgggggcggg gcaggaaacg
4801 gttggatcag cagcactctc aacttgtgtc cgtattgcga aggagggaga ggagcctgcc
4861 cgtattatgg gcgctgaaac gtcggaacca cctatgaatc ccacttctgc gtcgcggagg
4921 cgtactgctt ccgaccacct ccgcagctca cgcgcgtcga gctccctccc tttccccttc
4981 tttccgttgt gtgtgtgtgt gtgtgtgtgt cttgtcacga tgcatggcca tctctcctcc
5041 aaccctccac gccttcctct cccccgccca tcaaacgcgc tacggccaca accatttctt
5101 caagtatcac atccaccacc actcttaccc acctaccctc ttcgactctg acacagccac
5161 tgacgccctc tccgcctctc tcgctcgctg cacattgact ctccacagcc cccctccct
5221 cgcacagcca gccccaccca cccacccacc cacccaccgc tctacacacc tccacagcag
5281 ccggatccaa tatgggatcg gccattgaac aagatggatt gcacgcaggt tctccggccg
5341 cttgggtgga gaggctattc ggctatgact gggcacaaca gacaatcggc tgctctgatg
5401 ccgccgtgtt ccggctgtca gcgcaggggc gcccggttct tttgtcaag accgacctgt

```
5461 ccggtgccct gaatgaactg caggacgagg cagcgcggct atcgtggctg gccacgacgg
5521 gcgttccttg cgcagctgtg ctcgacgttg tcactgaagc gggaagggac tggctgctat
5581 tgggcgaagt gccggggcag gatctcctgt catctcacct tgctcctgcc gagaaagtat
5641 ccatcatggc tgatgcaatg cggcggctgc atacgcttga tccggctacc tgcccattcg
5701 accaccaagc gaaacatcgc atcgagcgag cacgtactcg gatggaagcc ggtcttgtcg
5761 atcaggatga tctggacgaa gagcatcagg ggctcgcgcc agccgaactg ttcgccaggc
5821 tcaaggcgcg catgcccgac ggcgaggatc tcgtcgtgac gcatggcgat gcctgcttgc
5881 cgaatatcat ggtggaaaat ggccgctttt ctggattcat cgactgtggc cggctgggtg
5941 tggcggaccg ctatcaggac atagcgttgg ctacccgtga tattgctgaa gagcttggcg
6001 gcgaatgggc tgaccgcttc ctcgtgcttt acggtatcgc cgctcccgat tcgcagcgca
6061 tcgccttcta tcgccttctt gacgagttct tctgaactag ttctagttct aggggccgaa
6121 ttcagatcct cgtgtgagcg ttcgcggaat cggtcgctcg tgtttatgcc cgtcttggtg
6181 ttgtgctcgc aaggcggtgc agcaggatac cgtcgccctc ctctctcctt gcttctctgt
6241 tcttcaattc gcgatctcac agaggccggc tgtgcacgcc cttcctcacc ctccttttcc
6301 cacctctcgg ccaccggtcg gctccgttcc gtctgccgtc gagaagggac gggcatgtgc
6361 agctcctccc tttctctcgc gcgcgcatct tctcttgctt gtggcactca cgctcatgcg
6421 tcaaggcggc cccacgccga gcccctgcgc tcccttccct cttgcgcatc cgtagcgcgg
6481 atgccgtcga tgcgcaaggc cggcatgaag gagcgcgtgc cctcaagagg gcacactatc
6541 atgccctacg tgggccacgc agcgatgagg ccggcttcgg cggagatgcg tcacgcacgt
6601 gccagatgat gcgctacgcc ttccttgact tgcgcccccc tctcttcctc cgtctctcac
6661 tctctctctc tcacacacac acacacacac acacacacac aaagctccgg ttcgtccggc
6721 cgtaacgcct tttcaactca cggcctctag gaatgaagga gggtagttcg gggggagaacg
6781 tactggggcg tcagaggtga aattcttaga ccgcaccaag acgaactaca gcgaaggcat
6841 tcttcaagga taccttcctc aatcaagaac caaagtgtgg agatcgaaga tgattagaga
6901 ccattgtagt ccacactgca aacgatgaca cccatgaatt ggggatctta tgggccggcc
6961 tgcggcaggg tttaccctgt gtcagcaccg cgcccgcttt taccaactta cgtatctttt
7021 ctattcggcc tttaccggcc acccacggga atatcctcag cacgttttct gttttttcac
7081 gcgaaagctt tgaggttaca gtctcagggg ggagtacgtt cgcaagagtg aaacttaaag
7141 aaattgacgg aatggcacca caagacgtgg agcgtgcggt ttaatttgac tcaacacggg
7201 gaactttacc agatccggac aggatgagga ttgacagatt gagtgttctt tctcgattcc
7261 ctgaatggtg gtgcatggcc gcttttggtc ggtggagtga tttgtttggt tgattccgtc
7321 aacggacgag atccaagctg cccagtagaa ttcagaattg cccatagaat agcaaactca
7381 tcggcgggtt ttacccaacg gtgggccgca ttcggtcgaa ttcttctctg cgggattcct
7441 ttgtaattgc acaaggtgaa attttgggca acagcaggtc tgtgatgctc ctcaatgttc
7501 tgggcgacac gcgcactaca atgtcagtga gaacaagaaa aacgactttt gtcgaaccta
7561 cttgatcaaa agagtgggga aaccccggaa tcacatagac ccacttggga ccgaggattg
7621 caattattgg tcgcgcaacg aggaatgtct cgtaggcgca gctcatcaaa ctgtgccgat
7681 tacgtccctg ccatttgtac acaccgcccg tcgttgtttc cgatgatggt gcaatacagg
7741 tgatcggaca ggcggtgttt tatccgcccg aaagttcacc gatatttaaa tccagctttt
7801 gttcccttta gtgagggtta attgcgcgct tggcgtaatc atggtcatag ctgtttcctg
7861 tgtgaaattg ttatccgctc acaattccac acaacatacg agccggaagc ataaagtgta
7921 aagcctgggg tgcctaatga gtgagctaac tcacattaat tgcgttgcgc tcactgcccg
7981 ctttccagtc gggaaacctg tcgtgccagc tgcattaatg aatcggccaa cgcgcgggga
8041 gaggcggttt gcgtattggg cgctcttccg cttcctcgct cactgactcg ctgcgctcgg
8101 tcgttcggct gcggcgagcg gtatcagctc actcaaaggc ggtaatacgg ttatccacag
8161 aatcagggga taacgcagga aagaacatgt gagcaaaagg ccagcaaaag gccaggaacc
8221 gtaaaaaggc cgcgttgctg gcgtttttcc ataggctccg cccccctgac gagcatcaca
8281 aaaatcgacg ctcaagtcag aggtggcgaa acccgacagg actataaaga taccaggcgt
8341 ttccccctgg aagctccctc gtgcgctctc ctgttccgac cctgccgctt accggatacc
8401 tgtccgcctt tctcccttcg ggaagcgtgg cgctttctca tagctcacgc tgtaggtatc
8461 tcagttcggt gtaggtcgtt cgctccaagc tgggctgtgt gcacgaaccc cccgttcagc
8521 ccgaccgctg cgccttatcc ggtaactatc gtcttgagtc caacccggta agacacgact
8581 tatcgccact ggcagcagcc actggtaaca ggattagcag agcgaggtat gtaggcggtg
8641 ctacagagtt cttgaagtgg tggcctaact acggctacac tagaagaaca gtatttggta
8701 tctgcgctct gctgaagcca gttaccttcg gaaaaagagt tggtagctct tgatccggca
8761 aacaaaccac cgctggtagc ggtggttttt ttgtttgcaa gcagcagatt acgcgcagaa
8821 aaaaaggatc tcaagaagat cctttgatct tttctacggg gtctgacgct cagtggaacg
8881 aaaactcacg ttaagggatt ttggtcatga gattatcaaa aaggatcttc acctagatcc
8941 ttttaaatta aaaatgaagt tttaaatcaa tctaaagtat atatgagtaa acttggtctg
9001 acagttacca atgcttaatc agtgaggcac ctatctcagc gatctgtcta tttcgttcat
9061 ccatagttgc ctgactcccc gtcgtgtaga taactacgat acgggagggc ttaccatctg
9121 gccccagtgc tgcaatgata ccgcgagacc cacgctcacc ggctccagat ttatcagcaa
9181 taaaccagcc agccggaagg gccgagcgca gaagtggtcc tgcaacttta tccgcctcca
9241 tccagtctat taattgttgc cgggaagcta gagtaagtag ttcgccagtt aatagtttgc
9301 gcaacgttgt tgccattgct acaggcatcg tggtgtcacg ctcgtcgttt ggtatggctt
9361 cattcagctc cggttcccaa cgatcaaggc gagttacatg atcccccatg ttgtgcaaaa
9421 aagcggttag ctccttcggt cctccgatcg ttgtcagaag taagttggcc gcagtgttat
9481 cactcatggt tatggcagca ctgcataatt ctcttactgt catgccatcc gtaagatgct
```

9541 tttctgtgac tggtgagtac tcaaccaagt cattctgaga atagtgtatg cggcgaccga
9601 gttgctcttg cccggcgtca atacgggata ataccgcgcc acatagcaga actttaaaag
9661 tgctcatcat tggaaaacgt tcttcggggc gaaaactctc aaggatctta ccgctgttga
9721 gatccagttc gatgtaaccc actcgtgcac ccaactgatc ttcagcatct tttactttca
9781 ccagcgtttc tgggtgagca aaaacaggaa ggcaaaatgc cgcaaaaaag ggaataaggg
9841 cgacacggaa atgttgaata ctcatactct tcctttttca atattattga agcatttatc
9901 agggttattg tctcatgagc ggatacatat ttgaatgtat ttagaaaaat aaacaaatag
9961 gggttccgcg cacatttccc cgaaaagtgc cacctgacgc gccctgtagc ggcgcattaa
10021 gcgcggcggg tgtggtggtt acgcgcagcg tgaccgctac acttgccagc gccctagcgc
10081 ccgctccttt cgctttcttc ccttcctttc tcgccacgtt cgccggcttt ccccgtcaag
10141 ctctaaatcg ggggcgccct ttagggttcc gatttagtgc tttacggcac ctcgacccca
10201 aaaaacttga ttagggtgat ggttcacgta gtgggccatc gccctgatag acggtttttc
10261 gccctttgac gttggagtcc acgttcttta atagtggact cttgttccaa actggaacaa
10321 cactcaaccc tatctcggtc tattcttttg atttataagg gattttgccg atttcggcct
10381 attggttaaa aaatgagct gatttaacaa aaatttaacg cgaattttaa caaaatatta
10441 acgcttacaa tttccattcg ccattcaggc tgcgcaactg ttgggaaggg cgatcggtgc
10501 gggcctcttc gctattacgc cagctggcga aagggggatg tgctgcaagg cgattaagtt
10561 gggtaacgcc agggttttcc cagtcacgac gttgtaaaac gacggccagt gagcgcgcgt
10621 aatacgactc actatagggc gaattggatt t
//

pLEXSY-neo2.1-LRV1-2014_CP

1 aaattggata acttggcgaa acgccaagct aatacatgaa ccaaccgggt gttctccact
61 ccagacggtg ggcaaccatc gtcgtgagac gcccagcgaa tgaatgacag taaaaccaat
121 gccttcactg gcagtaacac ccagcagtgt tgactcaatt cattccgtgc gaaagccggc
181 ttgttccggc gtcttttgac gaacaactgc cctatcagct ggtgatggcc gtgtagtgga
241 ctgccatgca tggcgttgac gggagcgggg gattagggtt cgattccgga gagggagcct
301 gagaaatagc taccacttct acggagggca gcaggcgcgc aaattgccca atgtcaaaac
361 aaaacgatga ggcagcgaaa agaaatagag ttgtcagtcc atttggattg tcatttcaat
421 gggggatatt taaacccatc caatatcgag taacaattgg aggacaagtc tggtgccagc
481 acccgcggta attccagctc caaaagcgta tattaatgct gttgctgtta aagggttcgt
541 agttgaactg tgggctgtgc aggtttgttc ctggtcgtcc cgtccatgtc ggatttggtg
601 acccaggccc ttgcagcccg tgaacattca aagaaacaag aaacacggga gtggttcctt
661 tcctgattta cgcatgtcat gcatgcgtcg aggatctggg cccgaggtct gcgattgacg
721 taggagttgc aagggggagg gggtatgaac ggggtgggta gagctttttt aggtggaagt
781 agtgagaggg tgggcttgag agaatttgag gtgtgttcgt gatgtgtgga tcttatcggg
841 ggctcggttg agtttttggt ttggtgattt tgtaggggtg aagggcggtg gaagggttgc
901 tgtgtgcgtg ccattccgtg aagcgtcatt cctatgtccg acactgaaat gcgtccgcga
961 ttgtgtggat tgtgtggaag tgtctgggcg acgactgaat tgagaagggg agtcgaaacg
1021 gtgcgtggat gccgtgtttg tgctacacac aagcaaaggc aacaacgtaa gcgcatcgag
1081 ccagacgaac caacccaata ttcacacaac tgtacgtggc agcgcgctgt atcgttggaa
1141 agacgatgcg tgtgcgcgtg ccaccaccat gtacaagttt ccagcctaac tgcttgtata
1201 ccttgtcttt tgcgttgtgc ctgcgcgtgc tgtgtacctc tgctcgtgtg tccagatcca
1261 gtcgcagcct gaccgcatca cacatcaagg cgttacagcc tctccgtctt tgctgcaccc
1321 aagtgtattc gtgtgaccgc gttgagggat ctaatggatt cggaactttg gtcgtttggt
1381 ctctcatgct gcccaatccc atgcgctttt ctccgtgcct ctctgtctcc ccttccgttc
1441 tatttggaca acgatgtgta ctgtaatgcg tgcgccatct gacaacgaat agatcagcag
1501 cattcgcaca cttgcacata cccagtgaag cttttgtgtc tgtcgtattg acaacaccga
1561 ctgcaacaag gtgtagatag aagttggcct tctcgctcgc tcgcacgctc ttcacgctcc
1621 tgctttcctt gctgtgcctt gccaccagat ct**atgtaccc atacgatgtt ccagattacg**
1681 ctgcttaccc atacgatgtt ccagattacg ctgccgatat accaaactcc gataaaatcg
1741 cttgcggtcg caagcctatg ttttgcgaga ttataaaact tgcaaacagg aagaggttga
1801 tctttgacac gaccagtgaa cgagtgtatg acgcacgatt gaactactgc tctactgctg
1861 attcaacggt gcaggctgac tgccatattt actggcggct caagctgcga cgtgctgatg
1921 ctgtttttga agagtatact ggtcagggct actcacttga tacagctgcc taccctcagc
1981 aatatacaga cattatcaga gggtattaca gcaaacacgt gtcaagctct ctcgctgcaa
2041 acactcaaca ctttgtgaac gttttggcta tgctgcgcca tgctggtgca tgtattgcac
2101 attattgcat gactgggaag attgattttg atattttaag taagaaaaaa tacaaaaata
2161 aggaggtggt gaccttgagc aatgccgaca gtttgagctt cttgccccac tccgccctat
2221 atctgccttc tcctttacgt gccagcgatc ctgaaatttt caatatgtta tatctacttg
2281 ggtgtgcatg tgacgcaagt attgcgatgg ataacatctc taatacaagc ggggctgcta
2341 aatacgctat gcctcactac aacccactgc agctatccca tgccttgcat gttacgatat
2401 tttacatgct aagcctgatg gacagttgtg ggtacggcga tgacgctgtt ctggcactta
2461 cctcgggctt acattctgta actacggtta tagcacatag cgatgagggt ggtataacac
2521 gcgacgccct acgtgagctg tcatatacac agccctatgg tactatgcct gtcccgattg
2581 ctggttactt ccaacatatc aatgtgctgt tcaccacaca gcctgcatgg gaccaatttg

2641 ctggaatttg ggattacgtc attttggcga cagctgcact agtacatcta tcagaccctg
2701 gaatgaccgt caacgatgtt acttacccca caactttaac aactaaagtg gcaacggttg
2761 atgggcggaa cagtgatttg gcggcccaga tgatgcacag tgcaactcga ttttgtgaca
2821 tatttattga gaatttaagc acgttctggg gcgtcgttgc caatccagat ggaaatgcaa
2881 gccaagcttt actccacgcc tttaacatcg ttgcttgtgc cgcagaacct aatagacatc
2941 tggaaatgaa tgtgatggct ccttggtatt ggatagaaag ttcagctctg ttctgtgact
3001 acacacagtt ccgatcaccg ctatcttcag caggatacgg tcctcaatgt atttatggtg
3061 cgaggctggt acttgcggct acgaactcac tagagttcac aggtgaagca ggtgattatt
3121 ctgcctaccg ctttgagtgg acgacaatgc gccacaaccc gttattcaac atccttaaca
3181 aacgggttgg cgacggtctg gccaatgtag atttcagatt acgccctttc aacgaatggt
3241 tgttagaagg tcaaccaagc cgccgaagct gtaactctgc tggtcatggc acgcctactg
3301 ccacgtgttc tcacaaagct ccaaaccgtg acacgttaga tgaatatatt tggggcagca
3361 cgtcatgtga cctatttcat ccagcagagc tgacttcgta cacggcagtg tgtgtcagat
3421 tccgtaacta catatcgggt gctgacggcg atgtcagaat actgaacacg ccgacgaggg
3481 aggtgattga aggaaatgtg gttacaagat gtgacggaat aaggtgtttg gacagcaaca
3541 aacgaataca gcatgttcct gaggtagcta gacgctactg tatgatggcg cggtacctcg
3601 cacaagcccg cacttttggt gcattaacta taggggatga tattatacgt ggctttgata
3661 aagttgaaaa agttgtcaaa atgcataaga gtaacaatag gttggatcaa atgcccttaa
3721 ttgacgtaac tgggttgtgt cagccaatga ttgaaacaag tacggtgcgt gcctccacaa
3781 cgacgcgtat cgatcccaac aaactcgcag ccgcgacggc tcgtgttgag ttaccactag
3841 caccacgttg tacatcttcc ctaatcccct catctgatgc agtgcccgaa ccagagcctc
3901 aagttggcga gtcgggtgac caaggcagct gcgcttaagc ggccgccctc ctcctccttt
3961 cttgttcctt tcacgtcgcc ttctcggttg tagctggcag acgacgagtc ttacttttac
4021 gtgtacttct ctatagatga tgtatgatct ctctgcatgc gtgttcgtgc atgtgtccgt
4081 gtgttgtgta cgcgtgcgtc tcgcctcagc tctccgcgtg aaagggtttg actgcccatg
4141 atgcgtgtgt atatccacgc gcaggcacgc acacacacac acacacacac acacacaggc
4201 acacacaggc acacaaacgc atctcaggcc gagccgcata cgtctctcgc acggtctcgt
4261 ttatttgatc atgtagttga gtattaaatt gggaagacaa aaacataata gcacgaagag
4321 tcgggcacga aaagcccgat ctctctctct ctctctcgtg cgcgcagggg cgtgtgggtg
4381 cacgacgacg aggacggagg ggggaaggga ggcataaacg gatcgatccc ccatcgcaca
4441 tgcgggtacg agcattatct gctgtgtctg gtcctttatc atatcgctac ccgcccgccc
4501 cccgcccccc tccccccccc ccccccgcgg ctcctgtcac tgtcgttcct gcgactcccc
4561 acacacccgc gcaacgctca tgtcacgaag agaaaagtag aaggcgtggc gatgcgttgc
4621 gcggccccct gtccgctggc atgcagacgt cgagccgtgg aacggatggt agcgagagac
4681 aacgcggcga tgcaggaaag gagatttact gcaggacgtc tacacacgcg cgcacaccac
4741 tgcgaatggc accgtgctga ggaaaattgg gggggagggc gcacgggggc ggggcaggaa
4801 acggttggat cagcagcact ctcaacttgt gtccgtattg cgaaggaggg agaggagcct
4861 gcccgtatta tgggcgctga aacgtcggaa ccacctatga atcccacttc tgcgtcgcgg
4921 aggcgtactg cttccgacca cctccgcagc tcacgcgcgt cgagctccct ccctttcccc
4981 ttctttccgt tgtgtgtgtg tgtgtgtgtg tgtcttgtca cgatgcatgg ccatctctcc
5041 tccaaccctc cacgccttcc tctcccccgc ccatcaaacg cgctacggcc acaaccattt
5101 cttcaagtat cacatccacc accactctta cccacctacc ctcttcgact ctgacacagc
5161 cactgacgcc ctctccgcct ctctcgctcg ctgcacattg actctccaca gcccccccctc
5221 cctcgcacag ccagccccac ccacccaccc acccacccac cgctctacac acctccacag
5281 cagccggatc caatatggga tcggccattg aacaagatgg attgcacgca ggttctccgg
5341 ccgcttgggt ggagaggcta ttcggctatg actgggcaca acagacaatc ggctgctctg
5401 atgccgccgt gttccggctg tcagcgcagg ggcgcccggt tctttttgtc aagaccgacc
5461 tgtccggtgc cctgaatgaa ctgcaggacg aggcagcgcg gctatcgtgg ctggccacga
5521 cgggcgttcc ttgcgcagct gtgctcgacg ttgtcactga agcgggaagg gactggctgc
5581 tattgggcga agtgccgggg caggatctcc tgtcatctca ccttgctcct gccgagaaag
5641 tatccatcat ggctgatgca atgcggcggc tgcatacgct tgatccggct acctgcccat
5701 tcgaccacca agcgaaacat cgcatcgagc gagcacgtac tcggatggaa gccggtcttg
5761 tcgatcagga tgatctggac gaagagcatc aggggctcgc gccagccgaa ctgttcgcca
5821 ggctcaaggc gcgcatgccc gacggcgagg atctcgtcgt gacgcatggc gatgcctgct
5881 tgccgaatat catggtggaa aatggccgct tttctggatt catcgactgt ggccggctgg
5941 gtgtggcgga ccgctatcag gacatagcgt tggctacccg tgatattgct gaagagcttg
6001 gcggcgaatg ggctgaccgc ttcctcgtgc tttacggtat cgccgctccc gattcgcagc
6061 gcatcgcctt ctatcgcctt cttgacgagt tcttctgaac tagttctagt tctaggggcc
6121 gaattcagat cctcgtgtga gcgttcgcgg aatcggtcgc tcgtgtttat gcccgtcttg
6181 gtgttgtgct cgcaaggcgg tgcagcagga taccgtcgcc ctcctctctc cttgcttctc
6241 tgttcttcaa ttcgcgatct cacagaggcc ggctgtgcac gcccttcctc accctccttt
6301 tcccacctct cggccaccgg tcggctccgt tccgtctgcc gtcgagaagg gacgggcatg
6361 tgcagctcct ccctttctct cgcgcgcgca tcttctcttg cttgtggcac tcacgctcat
6421 gcgtcaaggc ggccccacgc cgagcccctg cgctcccttc cctcttgcgc atccgtagcg
6481 cggatgccgt cgatgcgcaa ggccggcatg aaggagcgcg tgccctcaag agggcacact
6541 atcatgccct acgtgggcca cgcagcgatg aggccggctt cggcggagat gcgtcacgca
6601 cgtgccagat gatgcgctac gccttccttg acttgcgccc ccctctcttc ctccgtctct
6661 cactctctct ctctcacaca cacacacaca cacacacaca cacaaagctc cggttcgtcc

```
     6721 ggccgtaacg ccttttcaac tcacggcctc taggaatgaa ggagggtagt tcgggggaga
     6781 acgtactggg gcgtcagagg tgaaattctt agaccgcacc aagacgaact acagcgaagg
     6841 cattcttcaa ggataccttc ctcaatcaag aaccaaagtg tggagatcga agatgattag
     6901 agaccattgt agtccacact gcaaacgatg acacccatga attggggatc ttatgggccg
     6961 gcctgcggca gggtttaccc tgtgtcagca ccgcgcccgc ttttaccaac ttacgtatct
     7021 tttctattcg gcctttaccg gccacccacg ggaatatcct cagcacgttt tctgtttttt
     7081 cacgcgaaag ctttgaggtt acagtctcag gggggagtac gttcgcaaga gtgaaactta
     7141 aagaaattga cggaatggca ccacaagacg tggagcgtgc ggtttaattt gactcaacac
     7201 ggggaacttt accagatccg gacaggatga ggattgacag attgagtgtt ctttctcgat
     7261 tccctgaatg gtggtgcatg gccgcttttg gtcggtggag tgatttgttt ggttgattcc
     7321 gtcaacggac gagatccaag ctgcccagta gaattcagaa ttgcccatag aatagcaaac
     7381 tcatcggcgg gttttaccca acggtgggcc gcattcggtc gaattcttct ctgcgggatt
     7441 cctttgtaat tgcacaaggt gaaattttgg gcaacagcag gtctgtgatg ctcctcaatg
     7501 ttctgggcga cacgcgcact acaatgtcag tgagaacaag aaaaacgact tttgtcgaac
     7561 ctacttgatc aaaagagtgg ggaaaccccg gaatcacata gacccacttg ggaccgagga
     7621 ttgcaattat tggtcgcgca acgaggaatg tctcgtaggc gcagctcatc aaactgtgcc
     7681 gattacgtcc ctgccatttg tacacaccgc ccgtcgttgt ttccgatgat ggtgcaatac
     7741 aggtgatcgg acaggcggtg ttttatccgc ccgaaagttc accgatattt aaatccagct
     7801 tttgttccct ttagtgaggg ttaattgcgc gcttggcgta atcatggtca tagctgtttc
     7861 ctgtgtgaaa ttgttatccg ctcacaattc cacacaacat acgagccgga agcataaagt
     7921 gtaaagcctg gggtgcctaa tgagtgagct aactcacatt aattgcgttg cgctcactgc
     7981 ccgctttcca gtcgggaaac ctgtcgtgcc agctgcatta atgaatcggc caacgcgcgg
     8041 ggagaggcgg tttgcgtatt gggcgctctt ccgcttcctc gctcactgac tcgctgcgct
     8101 cggtcgttcg gctgcggcga gcggtatcag ctcactcaaa ggcggtaata cggttatcca
     8161 cagaatcagg ggataacgca ggaaagaaca tgtgagcaaa aggccagcaa aaggccagga
     8221 accgtaaaaa ggccgcgttg ctggcgtttt tccataggct ccgcccccct gacgagcatc
     8281 acaaaaatcg acgctcaagt cagaggtggc gaaacccgac aggactataa agataccagg
     8341 cgtttccccc tggaagctcc ctcgtgcgct ctcctgttcc gaccctgccg cttaccggat
     8401 acctgtccgc ctttctccct tcgggaagcg tggcgctttc tcatagctca cgctgtaggt
     8461 atctcagttc ggtgtaggtc gttcgctcca agctgggctg tgtgcacgaa ccccccgttc
     8521 agcccgaccg ctgcgcctta tccggtaact atcgtcttga gtccaacccg gtaagacacg
     8581 acttatcgcc actggcagca gccactggta acaggattag cagagcgagg tatgtaggcg
     8641 gtgctacaga gttcttgaag tggtggccta actacggcta cactagaaga acagtatttg
     8701 gtatctgcgc tctgctgaag ccagttacct tcggaaaaag agttggtagc tcttgatccg
     8761 gcaaacaaac caccgctggt agcggtggtt tttttgtttg caagcagcag attacgcgca
     8821 gaaaaaaagg atctcaagaa gatcctttga tcttttctac ggggtctgac gctcagtgga
     8881 acgaaaactc acgttaaggg attttggtca tgagattatc aaaaaggatc ttcacctaga
     8941 tccttttaaa ttaaaaatga agttttaaat caatctaaag tatatatgag taaacttggt
     9001 ctgacagtta ccaatgctta atcagtgagg cacctatctc agcgatctgt ctatttcgtt
     9061 catccatagt tgcctgactc cccgtcgtgt agataactac gatacgggag ggcttaccat
     9121 ctggccccag tgctgcaatg ataccgcgag acccacgctc accggctcca gatttatcag
     9181 caataaacca gccagccgga agggccgagc gcagaagtgg tcctgcaact ttatccgcct
     9241 ccatccagtc tattaattgt tgccgggaag ctagagtaag tagttcgcca gttaatagtt
     9301 tgcgcaacgt tgttgccatt gctacaggca tcgtggtgtc acgctcgtcg tttggtatgg
     9361 cttcattcag ctccggttcc caacgatcaa ggcgagttac atgatccccc atgttgtgca
     9421 aaaaagcggt tagctccttc ggtcctccga tcgttgtcag aagtaagttg gccgcagtgt
     9481 tatcactcat ggttatggca gcactgcata attctcttac tgtcatgcca tccgtaagat
     9541 gcttttctgt gactggtgag tactcaacca agtcattctg agaatagtgt atgcggcgac
     9601 cgagttgctc ttgcccggcg tcaatacggg ataataccgc gccacatagc agaactttaa
     9661 aagtgctcat cattggaaaa cgttcttcgg ggcgaaaact ctcaaggatc ttaccgctgt
     9721 tgagatccag ttcgatgtaa cccactcgtg cacccaactg atcttcagca tcttttactt
     9781 tcaccagcgt ttctgggtga gcaaaaacag gaaggcaaaa tgccgcaaaa aagggaataa
     9841 gggcgacacg gaaatgttga atactcatac tcttcctttt tcaatattat tgaagcattt
     9901 atcagggtta ttgtctcatg agcggataca tatttgaatg tatttagaaa aataaacaaa
     9961 taggggttcc gcgcacattt ccccgaaaag tgccacctga cgcgccctgt agcggcgcat
    10021 taagcgcggc gggtgtggtg gttacgcgca gcgtgaccgc tacacttgcc agcgccctag
    10081 cgcccgctcc tttcgctttc ttcccttcct ttctcgccac gttcgccggc tttccccgtc
    10141 aagctctaaa tcgggggcgc cctttagggt tccgatttag tgctttacgg cacctcgacc
    10201 ccaaaaaact tgattagggt gatggttcac gtagtgggcc atcgccctga tagacggttt
    10261 ttcgcccttt gacgttggag tccacgttct ttaatagtgg actcttgttc caaactggaa
    10321 caacactcaa ccctatctcg gtctattctt ttgatttata agggattttg ccgatttcgg
    10381 cctattggtt aaaaaaatga gctgatttaa caaaaattta acgcgaattt taacaaaata
    10441 ttaacgctta caatttccat tcgccattca ggctgcgcaa ctgttgggaa gggcgatcgg
    10501 tgcgggcctc ttcgctatta cgccagctgg cgaaaggggg atgtgctgca aggcgattaa
    10561 gttgggtaac gccagggttt tcccagtcac gacgttgtaa aacgacggcc agtgagcgcg
    10621 cgtaatacga ctcactatag ggcgaattgg attt
//
```

pLEXSY-neo2.1-LRV1-2700_CP

1 aaattggata acttggcgaa acgccaagct aatacatgaa ccaaccgggt gttctccact
61 ccagacggtg ggcaaccatc gtcgtgagac gcccagcgaa tgaatgacag taaaaccaat
121 gccttcactg gcagtaacac ccagcagtgt tgactcaatt cattccgtgc gaaagccggc
181 ttgttccggc gtcttttgac gaacaactgc cctatcagct ggtgatggcc gtgtagtgga
241 ctgccatgca tggcgttgac gggagcgggg gattagggtt cgattccgga gagggagcct
301 gagaaatagc taccacttct acggagggca gcaggcgcgc aaattgccca atgtcaaaac
361 aaaacgatga ggcagcgaaa agaaatagag ttgtcagtcc atttggattg tcatttcaat
421 gggggatatt taaacccatc caatatcgag taacaattgg aggacaagtc tggtgccagc
481 acccgcggta attccagctc caaaagcgta tattaatgct gttgctgtta aagggttcgt
541 agttgaactg tgggctgtgc aggtttgttc ctggtcgtcc cgtccatgtc ggatttggtg
601 acccaggccc ttgcagcccg tgaacattca aagaaacaag aaacacggga gtggttcctt
661 tcctgattta cgcatgtcat gcatgcgtcg aggatctggg cccgaggtct gcgattgacg
721 taggagttgc aagggggagg gggtatgaac ggggtgggta gagctttttt aggtggaagt
781 agtgagaggg tgggcttgag agaatttgag gtgtgttcgt gatgtgtgga tcttatcggg
841 ggctcggttg agtttttggt ttggtgattt tgtaggggtg aagggcggtg gaagggttgc
901 tgtgtgcgtg ccattccgtg aagcgtcatt cctatgtccg acactgaaat gcgtccgcga
961 ttgtgtggat tgtgtggaag tgtctgggcg acgactgaat tgagaagggg agtcgaaacg
1021 gtgcgtggat gccgtgtttg tgctacacac aagcaaaggc aacaacgtaa gcgcatcgag
1081 ccagacgaac caacccaata ttcacacaac tgtacgtggc agcgcgctgt atcgttggaa
1141 agacgatgcg tgtgcgcgtg ccaccaccat gtacaagttt ccagcctaac tgcttgtata
1201 ccttgtcttt tgcgttgtgc ctgcgcgtgc tgtgtacctc tgctcgtgtg tccagatcca
1261 gtcgcagcct gaccgcatca cacatcaagg cgttacagcc tctccgtctt tgctgcaccc
1321 aagtgtattc gtgtgaccgc gttgagggat ctaatggatt cggaactttg gtcgtttggt
1381 ctctcatgct gcccaatccc atgcgctttt ctccgtgcct ctctgtctcc ccttccgttc
1441 tatttggaca acgatgtgta ctgtaatgcg tgcgccatct gacaacgaat agatcagcag
1501 cattcgcaca cttgcacata cccagtgaag cttttgtgtc tgtcgtattg acaacaccga
1561 ctgcaacaag gtgtagatag aagttggcct tctcgctcgc tcgcacgctc ttcacgctcc
1621 tgctttcctt gctgtgcctt gccaccagat ct**atgtaccc atacgatgtt ccagattacg
1681 ctgcttaccc atacgatgtt ccagattacg cttctgatgt gccaaactct gataaggtgg
1741 catgtgggcg gaaacccatg ttctgtgaga ttataaaact cgcgaacagg aagaggttgg
1801 tccttgacac aacaagcgag agagtgtatg atgcgcgagt gaattattgt tcaacagctg
1861 acgcaaccgt gcaggctgat tgccacatat attggcggct gaaactgcgg cgggttgatg
1921 ctgtctttga ggaatacact ggccaaggat attcgctcga cacggctgct tacccccaac
1981 agtatgcaga tattataagg ggatactaca gcaaacatgt ttcgagtgct ctagctgcta
2041 acactcaaca cttcgtgaac gttttggcta tgctacggca tgcaggcgcc tgtattgcgc
2101 attactgcat gacaggtaaa attgatttcg acatcttaag taaaaagaaa cacagaaaca
2161 aggaagtggt gactttgagc aatgctgata gtttaagttt cttgcctcat tctgcactgt
2221 atttaccttc tgcgttacgt gttagtgacc cagagatatt taacacacta tatttgcttg
2281 gttgtgcatg tgatgcgagc attgcaatgg acaacatttc aaacacgagt ggggctgcca
2341 agtacgctat gccacactat aacccactac aactatcaca tgcactgcat gtcactatat
2401 tctacatgtt gggcttgatg gacagttgtg gatatgggga cgacgcaatc ttggcactga
2461 ctgctggtct acactcggtc acaaccgtga tagcccacag tgatgaaggt ggaataacgc
2521 gggacgccct acgagaacta tcatatacac aaccctatgg gactatgccc gtaccaattg
2581 cgggttattt tcaacacata aatgtgctct tcacatcgca gccagcttgg gatcaatttg
2641 ctggaatctg ggactatgtc atcttagcaa ctgcagcact agttcacctg tccgacccag
2701 gtatgactgt gaacgatgtc acttacccga ctactctaac gacgaaggta gccacaattg
2761 atggtcgtaa cagtgatata gccgcacaga tgatgcatag cgccacaagg ttttgtgata
2821 tatatgttga aaacttatgt gctttctggg gtgtagtagc taatcctgat ggaaatgcta
2881 gccaagcact actacacgca ttcaatgttg tcgcatgtgg aatcgaaccc aacagacacc
2941 ttgaaatgaa cgtgatggcg ccatggtatt ggatcgagag ttctgctttg ttcagtgact
3001 attcatcgtt ccgttcacct atctcgtcag ccggttatgg gccacaatgc gtgtatggtg
3061 ccagacttgt gtttgctgct acaaattcac ttgaatttac tggagaagca ggagattatt
3121 ctgcctaccg ttttgagtgg acaacaatgc gacataaccc actattcacc atcttaaaca
3181 agagagtggg tgacggtttg gccaatgtcg atttcagatt gcgcccattc aacgaatggc
3241 tcctcgaagg ccaaccgagc cgtcgatgct gtaatgctgc gaaccatggc acaccggcag
3301 ctgcttgttg ccacaaaact ccaaatcgtg actcacttga cgagtacatc tggggcagca
3361 cttcatgcga cctatttcat cctgctgagt tgacttcata cacagctgtc tgtgtgagat
3421 tccgaaattt tatgtcgacg gctgatggcg atgtgcaaat attgaacacg cctacgagag
3481 aggtcattga gggcaatgtc atcacaaggt gtgacgggat taggtgtctt gatagcaaca
3541 aacgtattca gcatgtccct gaggtcgcta gacgttactg tgttatggcc aggtacctag
3601 cacaagcacg tacgtttggg gctctcacta ttggtgatga tattatacgt ggttttgata
3661 aagttgataa aatcccgaag atgtataaga gtaataatag attaggtcaa atgcccttaa
3721 ttgatgtgac tggactgtgc caacccatga tagaaaccag tactgtgcga gcttctacca
3781 ccactcgcat tgacccgaac aaactggctg cagctgcgac acgcactgaa ctgccactgg
3841 ccccacgttc tacatcctct ctgataccat cctctgacgc tgttcccgat gcaaggctac
3901 aggctgatga gtcgggcggt ggagacagct ctgcctag**gc ggccgccctc ctcctccttt

3961 cttgttcctt tcacgtcgcc ttctcggttg tagctggcag acgacgagtc ttacttttac
4021 gtgtacttct ctatagatga tgtatgatct ctctgcatgc gtgttcgtgc atgtgtccgt
4081 gtgttgtgta cgcgtgcgtc tcgcctcagc tctccgcgtg aaagggtttg actgcccatg
4141 atgcgtgtgt atatccacgc gcaggcacgc acacacacac acacacacac acacacaggc
4201 acacacaggc acacaaacgc atctcaggcc gagccgcata cgtctctcgc acggtctcgt
4261 ttatttgatc atgtagttga gtattaaatt gggaagacaa aaacataata gcacgaagag
4321 tcgggcacga aaagcccgat ctctctctct ctctctcgtg cgcgcagggg cgtgtgggtg
4381 cacgacgacg aggacggagg ggggaaggga ggcataaacg gatcgatccc ccatcgcaca
4441 tgcgggtacg agcattatct gctgtgtctg gtcctttatc atatcgctac ccgcccgccc
4501 cccgcccccc tccccccccc ccccccgcgg ctcctgtcac tgtcgttcct gcgactcccc
4561 acacacccgc gcaacgctca tgtcacgaag agaaaagtag aaggcgtggc gatgcgttgc
4621 gcggccccct gtccgctggc atgcagacgt cgagccgtgg aacggatggt agcgagagac
4681 aacgcggcga tgcaggaaag gagatttact gcaggacgtc tacacacgcg cgcacaccac
4741 tgcgaatggc accgtgctga ggaaaattgg gggggagggc gcacgggggc ggggcaggaa
4801 acggttggat cagcagcact ctcaacttgt gtccgtattg cgaaggaggg agaggagcct
4861 gcccgtatta tgggcgctga aacgtcggaa ccacctatga atcccacttc tgcgtcgcgg
4921 aggcgtactg cttccgacca cctccgcagc tcacgcgcgt cgagctccct ccctttcccc
4981 ttctttccgt tgtgtgtgtg tgtgtgtgtg tgtcttgtca cgatgcatgg ccatctctcc
5041 tccaaccctc cacgccttcc tctcccccgc ccatcaaacg cgctacggcc acaaccattt
5101 cttcaagtat cacatccacc accactctta cccacctacc ctcttcgact ctgacacagc
5161 cactgacgcc ctctccgcct ctctcgctcg ctgcacattg actctccaca gcccccccctc
5221 cctcgcacag ccagccccac ccacccaccc acccacccac cgctctacac acctccacag
5281 cagccggatc caatatggga tcggccattg aacaagatgg attgcacgca ggttctccgg
5341 ccgcttgggt ggagaggcta ttcggctatg actgggcaca acagacaatc ggctgctctg
5401 atgccgccgt gttccggctg tcagcgcagg ggcgcccggt tctttttgtc aagaccgacc
5461 tgtccggtgc cctgaatgaa ctgcaggacg aggcagcgcg gctatcgtgg ctggccacga
5521 cgggcgttcc ttgcgcagct gtgctcgacg ttgtcactga agcgggaagg gactggctgc
5581 tattgggcga agtgccgggg caggatctcc tgtcatctca ccttgctcct gccgagaaag
5641 tatccatcat ggctgatgca atgcggcggc tgcatacgct tgatccggct acctgcccat
5701 tcgaccacca agcgaaacat cgcatcgagc gagcacgtac tcggatggaa gccggtcttg
5761 tcgatcagga tgatctggac gaagagcatc aggggctcgc gccagccgaa ctgttcgcca
5821 ggctcaaggc gcgcatgccc gacggcgagg atctcgtcgt gacgcatggc gatgcctgct
5881 tgccgaatat catggtggaa aatggccgct tttctggatt catcgactgt ggccggctgg
5941 gtgtggcgga ccgctatcag gacatagcgt tggctacccg tgatattgct gaagagcttg
6001 gcggcgaatg ggctgaccgc ttcctcgtgc tttacggtat cgccgctccc gattcgcagc
6061 gcatcgcctt ctatcgcctt cttgacgagt tcttctgaac tagttctagt tctaggggcc
6121 gaattcagat cctcgtgtga gcgttcgcgg aatcggtcgc tcgtgtttat gcccgtcttg
6181 gtgttgtgct cgcaaggcgg tgcagcagga taccgtcgcc ctcctctctc cttgcttctc
6241 tgttcttcaa ttcgcgatct cacagaggcc ggctgtgcac gcccttcctc accctccttt
6301 tcccacctct cggccaccgg tcggctccgt tccgtctgcc gtcgagaagg gacgggcatg
6361 tgcagctcct ccctttctct cgcgcgcgca tcttctcttg cttgtggcac tcacgctcat
6421 gcgtcaaggc ggccccacgc cgagcccctg cgctcccttc cctcttgcgc atccgtagcg
6481 cggatgccgt cgatgcgcaa ggccggcatg aaggagcgcg tgccctcaag agggcacact
6541 atcatgccct acgtgggcca cgcagcgatg aggccggctt cggcggagat gcgtcacgca
6601 cgtgccagat gatgcgctac gccttccttg acttgcgccc ccctctcttc ctccgtctct
6661 cactctctct ctctcacaca cacacacaca cacacacaca cacaaagctc cggttcgtcc
6721 ggccgtaacg ccttttcaac tcacggcctc taggaatgaa ggagggtagt tcgggggaga
6781 acgtactggg gcgtcagagg tgaaattctt agaccgcacc aagacgaact acagcgaagg
6841 cattcttcaa ggataccttc ctcaatcaag aaccaaagtg tggagatcga agatgattag
6901 agaccattgt agtccacact gcaaacgatg acacccatga attggggatc ttatgggccg
6961 gcctgcggca gggtttaccc tgtgtcagca ccgcgcccgc ttttaccaac ttacgtatct
7021 tttctattcg gcctttaccg gccacccacg ggaatatcct cagcacgttt tctgtttttt
7081 cacgcgaaag ctttgaggtt acagtctcag gggggagtac gttcgcaaga gtgaaactta
7141 aagaaattga cggaatggca ccacaagacg tggagcgtgc ggtttaattt gactcaacac
7201 ggggaacttt accagatccg gacaggatga ggattgacag attgagtgtt ctttctcgat
7261 tccctgaatg gtggtgcatg gccgcttttg gtcggtggag tgatttgttt ggttgattcc
7321 gtcaacggac gagatccaag ctgcccagta gaattcagaa ttgcccatag aatagcaaac
7381 tcatcggcgg gttttaccca acggtgggcc gcattcggtc gaattcttct ctgcgggatt
7441 cctttgtaat tgcacaaggt gaaattttgg gcaacagcag gtctgtgatg ctcctcaatg
7501 ttctgggcga cacgcgcact acaatgtcag tgagaacaag aaaaacgact tttgtcgaac
7561 ctacttgatc aaaagagtgg ggaaaccccg gaatcacata gacccacttg ggaccgagga
7621 ttgcaattat tggtcgcgca acgaggaatg tctcgtaggc gcagctcatc aaactgtgcc
7681 gattacgtcc ctgccatttg tacacaccgc ccgtcgttgt ttccgatgat ggtgcaatac
7741 aggtgatcgg acaggcggtg ttttatccgc ccgaaagttc accgatattt aaatccagct
7801 tttgttccct ttagtgaggg ttaattgcgc gcttggcgta atcatggtca tagctgtttc
7861 ctgtgtgaaa ttgttatccg ctcacaattc cacacaacat acgagccgga agcataaagt
7921 gtaaagcctg gggtgcctaa tgagtgagct aactcacatt aattgcgttg cgctcactgc
7981 ccgctttcca gtcgggaaac ctgtcgtgcc agctgcatta atgaatcggc caacgcgcgg

```
     8041 ggagaggcgg tttgcgtatt gggcgctctt ccgcttcctc gctcactgac tcgctgcgct
     8101 cggtcgttcg gctgcggcga gcggtatcag ctcactcaaa ggcggtaata cggttatcca
     8161 cagaatcagg ggataacgca ggaaagaaca tgtgagcaaa aggccagcaa aaggccagga
     8221 accgtaaaaa ggccgcgttg ctggcgtttt tccataggct ccgcccccct gacgagcatc
     8281 acaaaaatcg acgctcaagt cagaggtggc gaaacccgac aggactataa agataccagg
     8341 cgtttccccc tggaagctcc ctcgtgcgct ctcctgttcc gaccctgccg cttaccggat
     8401 acctgtccgc ctttctccct tcgggaagcg tggcgctttc tcatagctca cgctgtaggt
     8461 atctcagttc ggtgtaggtc gttcgctcca agctgggctg tgtgcacgaa ccccccgttc
     8521 agcccgaccg ctgcgcctta tccggtaact atcgtcttga gtccaacccg gtaagacacg
     8581 acttatcgcc actggcagca gccactggta acaggattag cagagcgagg tatgtaggcg
     8641 gtgctacaga gttcttgaag tggtggccta actacggcta cactagaaga acagtatttg
     8701 gtatctgcgc tctgctgaag ccagttacct tcggaaaaag agttggtagc tcttgatccg
     8761 gcaaacaaac caccgctggt agcggtggtt tttttgtttg caagcagcag attacgcgca
     8821 gaaaaaaagg atctcaagaa gatcctttga tcttttctac ggggtctgac gctcagtgga
     8881 acgaaaactc acgttaaggg attttggtca tgagattatc aaaaaggatc ttcacctaga
     8941 tccttttaaa ttaaaaatga agttttaaat caatctaaag tatatatgag taaacttggt
     9001 ctgacagtta ccaatgctta atcagtgagg cacctatctc agcgatctgt ctatttcgtt
     9061 catccatagt tgcctgactc cccgtcgtgt agataactac gatacgggag ggcttaccat
     9121 ctggccccag tgctgcaatg ataccgcgag acccacgctc accggctcca gatttatcag
     9181 caataaacca gccagccgga agggccgagc gcagaagtgg tcctgcaact ttatccgcct
     9241 ccatccagtc tattaattgt tgccgggaag ctagagtaag tagttcgcca gttaatagtt
     9301 tgcgcaacgt tgttgccatt gctacaggca tcgtggtgtc acgctcgtcg tttggtatgg
     9361 cttcattcag ctccggttcc caacgatcaa ggcgagttac atgatccccc atgttgtgca
     9421 aaaaagcggt tagctccttc ggtcctccga tcgttgtcag aagtaagttg gccgcagtgt
     9481 tatcactcat ggttatggca gcactgcata attctcttac tgtcatgcca tccgtaagat
     9541 gcttttctgt gactggtgag tactcaacca agtcattctg agaatagtgt atgcggcgac
     9601 cgagttgctc ttgcccggcg tcaatacggg ataataccgc gccacatagc agaactttaa
     9661 aagtgctcat cattggaaaa cgttcttcgg ggcgaaaact ctcaaggatc ttaccgctgt
     9721 tgagatccag ttcgatgtaa cccactcgtg cacccaactg atcttcagca tcttttactt
     9781 tcaccagcgt ttctgggtga gcaaaaacag gaaggcaaaa tgccgcaaaa aagggaataa
     9841 gggcgacacg gaaatgttga atactcatac tcttcctttt tcaatattat tgaagcattt
     9901 atcagggtta ttgtctcatg agcggataca tatttgaatg tatttagaaa aataaacaaa
     9961 taggggttcc gcgcacattt ccccgaaaag tgccacctga cgcgccctgt agcggcgcat
    10021 taagcgcggc gggtgtggtg gttacgcgca gcgtgaccgc tacacttgcc agcgccctag
    10081 cgcccgctcc tttcgctttc ttcccttcct ttctcgccac gttcgccggc tttccccgtc
    10141 aagctctaaa tcgggggcgc cctttagggt tccgatttag tgctttacgg cacctcgacc
    10201 ccaaaaaact tgattagggt gatggttcac gtagtgggcc atcgccctga tagacggttt
    10261 ttcgcccttt gacgttggag tccacgttct ttaatagtgg actcttgttc caaactggaa
    10321 caacactcaa ccctatctcg gtctattctt ttgatttata agggattttg ccgatttcgg
    10381 cctattggtt aaaaaatga gctgatttaa caaaaattta acgcgaattt taacaaaata
    10441 ttaacgctta caatttccat tcgccattca ggctgcgcaa ctgttgggaa gggcgatcgg
    10501 tgcgggcctc ttcgctatta cgccagctgg cgaaaggggg atgtgctgca aggcgattaa
    10561 gttgggtaac gccagggttt tcccagtcac gacgttgtaa aacgacggcc agtgagcgcg
    10621 cgtaatacga ctcactatag ggcgaattgg attt
//
```

pLEXSY-neo2.1-LRV2_CP

```
        1 AAATTGGATA ACTTGGCGAA ACGCCAAGCT AATACATGAA CCAACCGGGT GTTCTCCACT
       61 CCAGACGGTG GGCAACCATC GTCGTGAGAC GCCCAGCGAA TGAATGACAG TAAAACCAAT
      121 GCCTTCACTG GCAGTAACAC CCAGCAGTGT TGACTCAATT CATTCCGTGC GAAAGCCGGC
      181 TTGTTCCGGC GTCTTTTGAC GAACAACTGC CCTATCAGCT GGTGATGGCC GTGTAGTGGA
      241 CTGCCATGCA TGGCGTTGAC GGGAGCGGGG GATTAGGGTT CGATTCCGGA GAGGGAGCCT
      301 GAGAAATAGC TACCACTTCT ACGGAGGGCA GCAGGCGCGC AAATTGCCCA ATGTCAAAAC
      361 AAAACGATGA GGCAGCGAAA AGAAATAGAG TTGTCAGTCC ATTTGGATTG TCATTTCAAT
      421 GGGGGATATT TAAACCCATC CAATATCGAG TAACAATTGG AGGACAAGTC TGGTGCCAGC
      481 ACCCGCGGTA ATTCCAGCTC CAAAAGCGTA TATTAATGCT GTTGCTGTTA AAGGGTTCGT
      541 AGTTGAACTG TGGGCTGTGC AGGTTTGTTC CTGGTCGTCC CGTCCATGTC GGATTTGGTG
      601 ACCCAGGCCC TTGCAGCCCG TGAACATTCA AAGAAACAAG AAACACGGGA GTGGTTCCTT
      661 TCCTGATTTA CGCATGTCAT GCATGCGTCG AGGATCTGGG CCCGAGGTCT GCGATTGACG
      721 TAGGAGTTGC AAGGGGGAGG GGGTATGAAC GGGGTGGGTA GAGCTTTTTT AGGTGGAAGT
      781 AGTGAGAGGG TGGGCTTGAG AGAATTTGAG GTGTGTTCGT GATGTGTGGA TCTTATCGGG
      841 GGCTCGGTTG AGTTTTTGGT TTGGTGATTT TGTAGGGGTG AAGGGCGGTG GAAGGGTTGC
      901 TGTGTGCGTG CCATTCCGTG AAGCGTCATT CCTATGTCCG ACACTGAAAT GCGTCCGCGA
      961 TTGTGTGGAT TGTGTGGAAG TGTCTGGGCG ACGACTGAAT TGAGAAGGGG AGTCGAAACG
     1021 GTGCGTGGAT GCCGTGTTTG TGCTACACAC AAGCAAAGGC AACAACGTAA GCGCATCGAG
     1081 CCAGACGAAC CAACCCAATA TTCACACAAC TGTACGTGGC AGCGCGCTGT ATCGTTGGAA
```

1141 AGACGATGCG TGTGCGCGTG CCACCACCAT GTACAAGTTT CCAGCCTAAC TGCTTGTATA
1201 CCTTGTCTTT TGCGTTGTGC CTGCGCGTGC TGTGTACCTC TGCTCGTGTG TCCAGATCCA
1261 GTCGCAGCCT GACCGCATCA CACATCAAGG CGTTACAGCC TCTCCGTCTT TGCTGCACCC
1321 AAGTGTATTC GTGTGACCGC GTTGAGGGAT CTAATGGATT CGGAACTTTG GTCGTTTGGT
1381 CTCTCATGCT GCCCAATCCC ATGCGCTTTT CTCCGTGCCT CTCTGTCTCC CCTTCCGTTC
1441 TATTTGGACA ACGATGTGTA CTGTAATGCG TGCGCCATCT GACAACGAAT AGATCAGCAG
1501 CATTCGCACA CTTGCACATA CCCAGTGAAG CTTTTGTGTC TGTCGTATTG ACAACACCGA
1561 CTGCAACAAG GTGTAGATAG AAGTTGGCCT TCTCGCTCGC TCGCACGCTC TTCACGCTCC
1621 TGCTTTCCTT GCTGTGCCTT GCCACCAGAT CT**ATGTACCC ATACGATGTT CCAGATTACG**
1681 CTGCTTACCC ATACGATGTT CCAGATTACG CTACTTCACC GATTCCACAA ATGAATACGT
1741 CTGTATTCAA TCTAACTCGA GGTCGACGTG TCGTATCAGG AGATACGTCT GAAAGAGTTT
1801 ACACAGCGGT CCTCAACTAT TGCTCGACGG TTCGTCAGTC TTACCAAGCT ACGACGAATG
1861 TTCATTGGCG TCTAAAACAG AGAGGAAAAA AGGCGTCACA TGTAAGCCGA GCACATGCTG
1921 GTATTTCTGC ATCATTTGAT TATGCTGATA ACTTGAAACA GGAGATACTA AAATACCACT
1981 CTCGACACGT GTGTACAGAG ATAGTTGTAG GCCGAGATTA CAGGTTCGAT TTTCGTAGCT
2041 TATTACATCA TGCGGGGGCT TGTATGGCTC ACTACTGTTT GACTGGCGTT TTGGACTTGA
2101 ACGTGCTCAG CAGTGGTAAG TCAGGTGGTA CACGAGTCAC TGTGTTATCC AACGCAGATG
2161 GGTTGTCGAT CTTGAGCGCA GACGTCTTAT ACCTACCGAA CGCACTCAAG ATGATGGATA
2221 TAAATATATT TAACTGTATA TATCTACTTG CTGCTGCATG TGAGTGTAAC ATCGTTTTTG
2281 ATGATGTTAT TTTAACTGGG GGCACTTCAG CATATAACGT ACCAGCCCTA AACTCGGCAC
2341 AGTGGAGTAA GCACCTACAC AACACAATTA TATATCTGTT GCATTTACAA AGCAACGGAG
2401 ATGGCGGTTG TGATGGAGCA TTGGCACTTA CGGCAGGTTT ACATACTGTT GTAACAGTCG
2461 TTGCGCACAG TGATGAAGGA GGATTGATGC GTGATGTTTT GCGAGCGCTA CACTATGCAC
2521 CGGCGAAAGG TATTGTCGTG GCTGACCCAA ACCGTACTGT CAACCATTCG CTACCATTAC
2581 TGATGTATGA ACCAACGTGG GATTCTGTTT GTGGTATGTG GGATTACGTG GCTGTGGCAA
2641 CAGCAGGTCT GGTACACTTG TCCGACCCAT GTGAGCTAAT AGCTCATGAT ATCTACCCAA
2701 CAATAATCAC TGCTCCCGAT GAAAACAACG TACGTGCAAT AGAAGCTCGT GTTGAGGCTG
2761 CACTACCTGC TTTCGCTTCA ATATACATCC ATAATCTGGC AATGTTCATG GGTGTAGGTG
2821 GTGACGATTG TGGACGTGCG GTGGATACGT TGGTTGAAGC TGGTGTTTAT ATCTGCTCAC
2881 TTACAGGCAA TGATGCCAGT CGTCATATGC TTAAGGGAAC TATGGCACCA TGGTTCTGGG
2941 TTGAGTCTAC AGGCTTGTTT CGGGACCTCA CGTGCTTCAA CATCCCTGGT TTGAGCGCTG
3001 GCTATGGTCC TCAAGCGATA TATGGTACAG CTGTTACCCA GCCAGCTATG CAAGCTTGCG
3061 AATATACAGG TACTCGAGGA TCATACGATT ATTACAATGT CGGGTGGACA TCATTACGCA
3121 AGCACCCGCT TCTTGTCCTT ACAAATAACA GAACTGGTGA CGGCATTGCA CATATGGAGG
3181 TGGGGCGTGA TGCTGCCATA CCGTGGGTCT TGCCGGGACA ACCACAAAGA CGTGAGTGTA
3241 CTGCACAAGG TCACGGGACC CAGACTGCAA CGTGTGGACA TACAAGACAC AATCGAGGGA
3301 CATTGGATGA ATATATCTGG GGTAGATTCT CCACGGGCCT ATTCCATCCA GCCGAGCTGA
3361 CAACTTTTAC TAACGTGGAG TTCCGCGTTA AGTGTTGGAC GGAAGACAAC GACGGTAACG
3421 TACTCGAAAC AGGAGCACCA GTCAAGGATA TTGTGGAGGG GGGGGTTACT GTGAGTGTGA
3481 ACGCAATACT ACTTACAAAT TCCACCAATC ACATACGTAC GGTACCAGCG GTTCGACGAA
3541 GTTACCAAGC TGGCGCGAAG TATCTCGAAG AAGCACGAAG CCGTGGTACG ATGTCTACAC
3601 TTAATAGAAT AATAGGAGGA CAACTTTTTA GGGATCTCCA ACCTATAAGC AAGCGAGGGG
3661 TGCCGCCACA GCCCGAGCCT GTATGTGCAA GCATGCCTGT TGAAACAACG AAAGAGGTTA
3721 GTCTGTCAAG GATCGGACCA ATTCGTATCA ACCAGAACGT TTTTAGACCA CGACCACAAC
3781 CTACAGAAGA AGAGATGCAT GAAGAACCAA CGCCAGTAGC GCTCGAAGCA GAGCCACCAG
3841 GCGAAAATGT GTAAGCGGCC GCCCTCCTCC TCCTTTCTTG TTCCTTTCAC GTCGCCTTCT
3901 CGGTTGTAGC TGGCAGACGA CGAGTCTTAC TTTTACGTGT ACTTCTCTAT AGATGATGTA
3961 TGATCTCTCT GCATGCGTGT TCGTGCATGT GTCCGTGTGT TGTGTACGCG TGCGTCTCGC
4021 CTCAGCTCTC CGCGTGAAAG GGTTTGACTG CCCATGATGC GTGTGTATAT CCACGCGCAG
4081 GCACGCACAC ACACACACAC ACACACACAC ACAGGCACAC ACAGGCACAC AAACGCATCT
4141 CAGGCCGAGC CGCATACGTC TCTCGCACGG TCTCGTTTAT TTGATCATGT AGTTGAGTAT
4201 TAAATTGGGA AGACAAAAAC ATAATAGCAC GAAGAGTCGG GCACGAAAAG CCCGATCTCT
4261 CTCTCTCTCT CTCGTGCGCG CAGGGGCGTG TGGGTGCACG ACGACGAGGA CGGAGGGGGG
4321 AAGGGAGGCA TAAACGGATC GATCCCCCAT CGCACATGCG GGTACGAGCA TTATCTGCTG
4381 TGTCTGGTCC TTTATCATAT CGCTACCCGC CCGCCCCCCG CCCCCCTCCC CCCCCCCCCC
4441 CCGCGGCTCC TGTCACTGTC GTTCCTGCGA CTCCCCACAC ACCCGCGCAA CGCTCATGTC
4501 ACGAAGAGAA AAGTAGAAGG CGTGGCGATG CGTTGCGCGG CCCCCTGTCC GCTGGCATGC
4561 AGACGTCGAG CCGTGGAACG GATGGTAGCG AGAGACAACG CGGCGATGCA GGAAAGGAGA
4621 TTTACTGCAG GACGTCTACA CACGCGCGCA CACCACTGCG AATGGCACCG TGCTGAGGAA
4681 AATTGGGGGG GAGGGCGCAC GGGGGCGGGG CAGGAAACGG TTGGATCAGC AGCACTCTCA
4741 ACTTGTGTCC GTATTGCGAA GGAGGGAGAG GAGCCTGCCC GTATTATGGG CGCTGAAACG
4801 TCGGAACCAC CTATGAATCC CACTTCTGCG TCGCGGAGGC GTACTGCTTC CGACCACCTC
4861 CGCAGCTCAC GCGCGTCGAG CTCCCTCCCT TTCCCCTTCT TTCCGTTGTG TGTGTGTGTG
4921 TGTGTGTGTC TTGTCACGAT GCATGGCCAT CTCTCCTCCA ACCCTCCACG CCTTCCTCTC
4981 CCCCGCCCAT CAAACGCGCT ACGGCCACAA CCATTTCTTC AAGTATCACA TCCACCACCA
5041 CTCTTACCCA CCTACCCTCT TCGACTCTGA CACAGCCACT GACGCCCTCT CCGCCTCTCT
5101 CGCTCGCTGC ACATTGACTC TCCACAGCCC CCCCTCCCTC GCACAGCCAG CCCCACCCAC
5161 CCACCCACCC ACCCACCGCT CTACACACCT CCACAGCAGC CGGATCCAAT ATGGGATCGG

```
5221 CCATTGAACA AGATGGATTG CACGCAGGTT CTCCGGCCGC TTGGGTGGAG AGGCTATTCG
5281 GCTATGACTG GGCACAACAG ACAATCGGCT GCTCTGATGC CGCCGTGTTC CGGCTGTCAG
5341 CGCAGGGGCG CCCGGTTCTT TTTGTCAAGA CCGACCTGTC CGGTGCCCTG AATGAACTGC
5401 AGGACGAGGC AGCGCGGCTA TCGTGGCTGG CCACGACGGG CGTTCCTTGC GCAGCTGTGC
5461 TCGACGTTGT CACTGAAGCG GGAAGGGACT GGCTGCTATT GGGCGAAGTG CCGGGGCAGG
5521 ATCTCCTGTC ATCTCACCTT GCTCCTGCCG AGAAAGTATC CATCATGGCT GATGCAATGC
5581 GGCGGCTGCA TACGCTTGAT CCGGCTACCT GCCCATTCGA CCACCAAGCG AAACATCGCA
5641 TCGAGCGAGC ACGTACTCGG ATGGAAGCCG GTCTTGTCGA TCAGGATGAT CTGGACGAAG
5701 AGCATCAGGG GCTCGCGCCA GCCGAACTGT TCGCCAGGCT CAAGGCGCGC ATGCCCGACG
5761 GCGAGGATCT CGTCGTGACG CATGGCGATG CCTGCTTGCC GAATATCATG GTGGAAAATG
5821 GCCGCTTTTC TGGATTCATC GACTGTGGCC GGCTGGGTGT GGCGGACCGC TATCAGGACA
5881 TAGCGTTGGC TACCCGTGAT ATTGCTGAAG AGCTTGGCGG CGAATGGGCT GACCGCTTCC
5941 TCGTGCTTTA CGGTATCGCC GCTCCCGATT CGCAGCGCAT CGCCTTCTAT CGCCTTCTTG
6001 ACGAGTTCTT CTGAACTAGT TCTAGTTCTA GGGGCCGAAT TCAGATCCTC GTGTGAGCGT
6061 TCGCGGAATC GGTCGCTCGT GTTTATGCCC GTCTTGGTGT TGTGCTCGCA AGGCGGTGCA
6121 GCAGGATACC GTCGCCCTCC TCTCTCCTTG CTTCTCTGTT CTTCAATTCG CGATCTCACA
6181 GAGGCCGGCT GTGCACGCCC TTCCTCACCC TCCTTTTCCC ACCTCTCGGC CACCGGTCGG
6241 CTCCGTTCCG TCTGCCGTCG AGAAGGGACG GGCATGTGCA GCTCCTCCCT TTCTCTCGCG
6301 CGCGCATCTT CTCTTGCTTG TGGCACTCAC GCTCATGCGT CAAGGCGGCC CCACGCCGAG
6361 CCCCTGCGCT CCCTTCCCTC TTGCGCATCC GTAGCGCGGA TGCCGTCGAT GCGCAAGGCC
6421 GGCATGAAGG AGCGCGTGCC CTCAAGAGGG CACACTATCA TGCCCTACGT GGGCCACGCA
6481 GCGATGAGGC CGGCTTCGGC GGAGATGCGT CACGCACGTG CCAGATGATG CGCTACGCCT
6541 TCCTTGACTT GCGCCCCCCT CTCTTCCTCC GTCTCTCACT CTCTCTCTCT CACACACACA
6601 CACACACACA CACACACACA AAGCTCCGGT TCGTCCGGCC GTAACGCCTT TTCAACTCAC
6661 GGCCTCTAGG AATGAAGGAG GGTAGTTCGG GGGAGAACGT ACTGGGGCGT CAGAGGTGAA
6721 ATTCTTAGAC CGCACCAAGA CGAACTACAG CGAAGGCATT CTTCAAGGAT ACCTTCCTCA
6781 ATCAAGAACC AAAGTGTGGA GATCGAAGAT GATTAGAGAC CATTGTAGTC CACACTGCAA
6841 ACGATGACAC CCATGAATTG GGGATCTTAT GGGCCGGCCT GCGGCAGGGT TTACCCTGTG
6901 TCAGCACCGC GCCCGCTTTT ACCAACTTAC GTATCTTTTC TATTCGGCCT TTACCGGCCA
6961 CCCACGGGAA TATCCTCAGC ACGTTTTCTG TTTTTTCACG CGAAAGCTTT GAGGTTACAG
7021 TCTCAGGGGG GAGTACGTTC GCAAGAGTGA AACTTAAAGA AATTGACGGA ATGGCACCAC
7081 AAGACGTGGA GCGTGCGGTT TAATTTGACT CAACACGGGG AACTTTACCA GATCCGGACA
7141 GGATGAGGAT TGACAGATTG AGTGTTCTTT CTCGATTCCC TGAATGGTGG TGCATGGCCG
7201 CTTTTGGTCG GTGGAGTGAT TTGTTTGGTT GATTCCGTCA ACGGACGAGA TCCAAGCTGC
7261 CCAGTAGAAT TCAGAATTGC CCATAGAATA GCAAACTCAT CGGCGGGTTT TACCCAACGG
7321 TGGGCCGCAT TCGGTCGAAT TCTTCTCTGC GGGATTCCTT TGTAATTGCA CAAGGTGAAA
7381 TTTTGGGCAA CAGCAGGTCT GTGATGCTCC TCAATGTTCT GGGCGACACG CGCACTACAA
7441 TGTCAGTGAG AACAAGAAAA ACGACTTTTG TCGAACCTAC TTGATCAAAA GAGTGGGGAA
7501 ACCCCGGAAT CACATAGACC CACTTGGGAC CGAGGATTGC AATTATTGGT CGCGCAACGA
7561 GGAATGTCTC GTAGGCGCAG CTCATCAAAC TGTGCCGATT ACGTCCCTGC CATTTGTACA
7621 CACCGCCCGT CGTTGTTTCC GATGATGGTG CAATACAGGT GATCGGACAG GCGGTGTTTT
7681 ATCCGCCCGA AAGTTCACCG ATATTTAAAT CCAGCTTTTG TTCCCTTTAG TGAGGGTTAA
7741 TTGCGCGCTT GGCGTAATCA TGGTCATAGC TGTTTCCTGT GTGAAATTGT TATCCGCTCA
7801 CAATTCCACA CAACATACGA GCCGGAAGCA TAAAGTGTAA AGCCTGGGGT GCCTAATGAG
7861 TGAGCTAACT CACATTAATT GCGTTGCGCT CACTGCCCGC TTTCCAGTCG GGAAACCTGT
7921 CGTGCCAGCT GCATTAATGA ATCGGCCAAC GCGCGGGGAG AGGCGGTTTG CGTATTGGGC
7981 GCTCTTCCGC TTCCTCGCTC ACTGACTCGC TGCGCTCGGT CGTTCGGCTG CGGCGAGCGG
8041 TATCAGCTCA CTCAAAGGCG GTAATACGGT TATCCACAGA ATCAGGGGAT AACGCAGGAA
8101 AGAACATGTG AGCAAAAGGC CAGCAAAAGG CCAGGAACCG TAAAAAGGCC GCGTTGCTGG
8161 CGTTTTTCCA TAGGCTCCGC CCCCCTGACG AGCATCACAA AAATCGACGC TCAAGTCAGA
8221 GGTGGCGAAA CCCGACAGGA CTATAAAGAT ACCAGGCGTT TCCCCCTGGA AGCTCCCTCG
8281 TGCGCTCTCC TGTTCCGACC CTGCCGCTTA CCGGATACCT GTCCGCCTTT CTCCCTTCGG
8341 GAAGCGTGGC GCTTTCTCAT AGCTCACGCT GTAGGTATCT CAGTTCGGTG TAGGTCGTTC
8401 GCTCCAAGCT GGGCTGTGTG CACGAACCCC CCGTTCAGCC CGACCGCTGC GCCTTATCCG
8461 GTAACTATCG TCTTGAGTCC AACCCGGTAA GACACGACTT ATCGCCACTG GCAGCAGCCA
8521 CTGGTAACAG GATTAGCAGA GCGAGGTATG TAGGCGGTGC TACAGAGTTC TTGAAGTGGT
8581 GGCCTAACTA CGGCTACACT AGAAGAACAG TATTTGGTAT CTGCGCTCTG CTGAAGCCAG
8641 TTACCTTCGG AAAAAGAGTT GGTAGCTCTT GATCCGGCAA ACAAACCACC GCTGGTAGCG
8701 GTGGTTTTTT TGTTTGCAAG CAGCAGATTA CGCGCAGAAA AAAAGGATCT CAAGAAGATC
8761 CTTTGATCTT TTCTACGGGG TCTGACGCTC AGTGGAACGA AAACTCACGT TAAGGGATTT
8821 TGGTCATGAG ATTATCAAAA AGGATCTTCA CCTAGATCCT TTTAAATTAA AAATGAAGTT
8881 TTAAATCAAT CTAAAGTATA TATGAGTAAA CTTGGTCTGA CAGTTACCAA TGCTTAATCA
8941 GTGAGGCACC TATCTCAGCG ATCTGTCTAT TTCGTTCATC CATAGTTGCC TGACTCCCCG
9001 TCGTGTAGAT AACTACGATA CGGGAGGGCT TACCATCTGG CCCCAGTGCT GCAATGATAC
9061 CGCGAGACCC ACGCTCACCG GCTCCAGATT TATCAGCAAT AAACCAGCCA GCCGGAAGGG
9121 CCGAGCGCAG AAGTGGTCCT GCAACTTTAT CCGCCTCCAT CCAGTCTATT AATTGTTGCC
9181 GGGAAGCTAG AGTAAGTAGT TCGCCAGTTA ATAGTTTGCG CAACGTTGTT GCCATTGCTA
9241 CAGGCATCGT GGTGTCACGC TCGTCGTTTG GTATGGCTTC ATTCAGCTCC GGTTCCCAAC
```

```
     9301 GATCAAGGCG AGTTACATGA TCCCCCATGT TGTGCAAAAA AGCGGTTAGC TCCTTCGGTC
     9361 CTCCGATCGT TGTCAGAAGT AAGTTGGCCG CAGTGTTATC ACTCATGGTT ATGGCAGCAC
     9421 TGCATAATTC TCTTACTGTC ATGCCATCCG TAAGATGCTT TTCTGTGACT GGTGAGTACT
     9481 CAACCAAGTC ATTCTGAGAA TAGTGTATGC GGCGACCGAG TTGCTCTTGC CCGGCGTCAA
     9541 TACGGGATAA TACCGCGCCA CATAGCAGAA CTTTAAAAGT GCTCATCATT GGAAAACGTT
     9601 CTTCGGGGCG AAAACTCTCA AGGATCTTAC CGCTGTTGAG ATCCAGTTCG ATGTAACCCA
     9661 CTCGTGCACC CAACTGATCT TCAGCATCTT TTACTTTCAC CAGCGTTTCT GGGTGAGCAA
     9721 AAACAGGAAG GCAAAATGCC GCAAAAAAGG GAATAAGGGC GACACGGAAA TGTTGAATAC
     9781 TCATACTCTT CCTTTTTCAA TATTATTGAA GCATTTATCA GGGTTATTGT CTCATGAGCG
     9841 GATACATATT TGAATGTATT TAGAAAAATA AACAAATAGG GGTTCCGCGC ACATTTCCCC
     9901 GAAAAGTGCC ACCTGACGCG CCCTGTAGCG GCGCATTAAG CGCGGCGGGT GTGGTGGTTA
     9961 CGCGCAGCGT GACCGCTACA CTTGCCAGCG CCCTAGCGCC CGCTCCTTTC GCTTTCTTCC
    10021 CTTCCTTTCT CGCCACGTTC GCCGGCTTTC CCCGTCAAGC TCTAAATCGG GGGCGCCCTT
    10081 TAGGGTTCCG ATTTAGTGCT TTACGGCACC TCGACCCCAA AAAACTTGAT TAGGGTGATG
    10141 GTTCACGTAG TGGGCCATCG CCCTGATAGA CGGTTTTTCG CCCTTTGACG TTGGAGTCCA
    10201 CGTTCTTTAA TAGTGGACTC TTGTTCCAAA CTGGAACAAC ACTCAACCCT ATCTCGGTCT
    10261 ATTCTTTTGA TTTATAAGGG ATTTTGCCGA TTTCGGCCTA TTGGTTAAAA AAATGAGCTG
    10321 ATTTAACAAA AATTTAACGC GAATTTTAAC AAAATATTAA CGCTTACAAT TTCCATTCGC
    10381 CATTCAGGCT GCGCAACTGT TGGGAAGGGC GATCGGTGCG GGCCTCTTCG CTATTACGCC
    10441 AGCTGGCGAA AGGGGGATGT GCTGCAAGGC GATTAAGTTG GGTAACGCCA GGGTTTTCCC
    10501 AGTCACGACG TTGTAAAACG ACGGCCAGTG AGCGCGCGTA ATACGACTCA CTATAGGGCG
    10561 AATTGGATTT
//
```

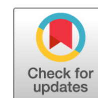
APPENDIX C. PUBLICATION

Some of this work was published as part of the following article:

Saura, A., Zakharova, A., **Klocek, D.**, Gerasimov, E. S., Butenko, A., Macedo, D. H., Servienė, E., Zagirova, D., Meshcheryakova, A., Rogozin, I. B., Serva, S., Kostygov, A. Y., & Yurchenko, V. (2022). Elimination of LRVs Elicits Different Responses in Leishmania spp. *mSphere*, 7(4), e0033522. https://doi.org/10.1128/msphere.00335-22

AMERICAN SOCIETY FOR MICROBIOLOGY | mSphere®

RESEARCH ARTICLE



Elimination of LRVs Elicits Different Responses in *Leishmania* spp.

Andreu Saura,[a] Alexandra Zakharova,[a] Donnamae Klocek,[a] Evgeny S. Gerasimov,[a] Anzhelika Butenko,[a,b,c] Diego H. Macedo,[a*] Elena Servienė,[d] Diana Zagirova,[a] Anastasia Meshcheryakova,[a,e] Igor B. Rogozin,[f] Saulius Serva,[g] Alexei Yu. Kostygov,[a] Vyacheslav Yurchenko[a]

[a]Life Science Research Centre, Faculty of Science, University of Ostrava, Ostrava, Czech Republic

[b]Institute of Parasitology, Biology Centre, Czech Academy of Sciences, České Budějovice (Budweis), Czech Republic

[c]Faculty of Science, University of South Bohemia, České Budějovice (Budweis), Czech Republic

[d]Laboratory of Genetics, Institute of Botany, Nature Research Centre, Vilnius, Lithuania

[e]Department of Pathophysiology and Allergy Research, Center of Pathophysiology, Infectiology and Immunology, Medical University of Vienna, Vienna, Austria

[f]National Center for Biotechnology Information, National Library of Medicine, National Institutes of Health, Bethesda, Maryland, USA

[g]Department of Biochemistry and Molecular Biology, Institute of Biosciences, Vilnius University, Vilnius, Lithuania

Andreu Saura and Alexandra Zakharova contributed equally and authors' order was based on alphabetical order of last names.

ABSTRACT Leishmaniaviruses (LRVs) have been demonstrated to enhance progression of leishmaniasis, a vector-transmitted disease with a wide range of clinical manifestations that is caused by flagellates of the genus *Leishmania*. Here, we used two previously proposed strategies of the LRV ablation to shed light on the relationships of two *Leishmania* spp. with their respective viral species (*L. guyanensis*, LRV1 and *L. major*, LRV2) and demonstrated considerable difference between two studied systems. LRV1 could be easily eliminated by the expression of exogenous capsids regardless of their origin (the same or distantly related LRV1 strains, or even LRV2), while LRV2 was only partially depleted in the case of the native capsid overexpression. The striking differences were also observed in the effects of complete viral elimination with 2'C-methyladenosine (2-CMA) on the transcriptional profiles of these two *Leishmania* spp. While virtually no differentially expressed genes were detected after the LRV1 removal from *L. guyanensis*, the response of *L. major* after ablation of LRV2 involved 87 genes, the analysis of which suggested a considerable stress experienced even after several passages following the treatment. This effect on *L. major* was also reflected in a significant decrease of the proliferation rate, not documented in *L. guyanensis* and naturally virus-free strain of *L. major*. Our findings suggest that integration of *L. major* with LRV2 is deeper compared with that of *L. guyanensis* with LRV1. We presume this determines different effects of the viral presence on the *Leishmania* spp. infections.

IMPORTANCE *Leishmania* spp. represent human pathogens that cause leishmaniasis, a widespread parasitic disease with mild to fatal clinical manifestations. Some strains of leishmaniae bear leishmaniaviruses (LRVs), and this has been shown to aggravate disease course. We investigated the relationships of two distally related *Leishmania* spp. with their respective LRVs using different strategies of virus removal. Our results suggest the South American *L. guyanensis* easily loses its virus with no important consequences for the parasite in the laboratory culture. Conversely, the Old-World *L. major* is refractory to virus removal and experiences a prominent stress if this removal is nonetheless completed. The drastically different levels of integration between the studied *Leishmania* spp. and their viruses suggest distinct effects of the viral presence on infections in these species of parasites.

KEYWORDS *Leishmania guyanensis*, *Leishmania major*, LRV1, LRV2, capsid

Editor Ira J. Blader, University at Buffalo

Copyright © 2022 Saura et al. This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International license.

Address correspondence to Vyacheslav Yurchenko, vyacheslav.yurchenko@osu.cz, or Alexei Yu. Kostygov, kostygov@gmail.com.

*Present address: Diego H. Macedo, Stockholm University, Sweden.

The authors declare no conflict of interest.

Received 11 July 2022
Accepted 21 July 2022
Published 9 August 2022

Leishmaniasis remains a public health concern affecting over 1.2 million people worldwide annually (1). It manifests in a repertoire of symptoms ranging from self-healing lesions in the case of cutaneous forms to fatal organ failures in visceral leishmaniasis (2). Even though the clinical picture of the disease usually depends on the infecting *Leishmania* species and the immune status of the host, our understanding of the molecular factors modulating the etiology of leishmaniases remains rather limited (3). One such a factor is the presence of double-stranded RNA (dsRNA) *Leishmania* RNA viruses (LRVs, genus *Leishmaniavirus*) of the family *Totiviridae*. Most totiviruses infect fungi (4), while some have been documented from animals (5–7) and protists (8–11). The LRVs suppress the anti-leishmanial immune response of the vertebrate host and, thus, provide a survival advantage to the parasites (12, 13). The two best studied species, LRV1 and LRV2, infect *Leishmania* of the New World (subgenus *Viannia*) and the Old World (subgenus *Leishmania*), respectively (14). Recently, two other *Leishmaniavirus* species, LRV3 and LRV4, have been described in *Blechomonas* spp., distant relatives of *Leishmania* parasitizing fleas (15). The dsRNA of LRV1 facilitates chronic inflammation and spread of *L. guyanensis* to secondary sites (16–18). It is generally assumed *Leishmania* and LRV coevolve (14, 19, 20), although occasional horizontal viral transfer events have also been reported (21). The genome of LRVs contains four open reading frames (ORFs), two of which (ORF2 and 3) encode the capsid and RNA-dependent RNA-polymerase (RDRP), respectively (22, 23).

As LRV presence is considered clinically important (24–26), different strategies of viral elimination were proposed in order to make *Leishmania* less virulent. One of the early approaches relied on hygromycin B treatment of *L. guyanensis*: parasites transfected with pX63-HYG plasmid and kept under antibiotic selection for several weeks lost the virus (27). The phenomenon was explained by specific inhibition of viral gene translation on the background of hygromycin B resistance of *Leishmania* strains. The resultant strain, *L. guyanensis* pX63-HYG, became a “gold standard” in all LRV1-related experiments for many years (16, 28). Another strategy was based on the chemical inhibition of viral replication by 2′C-methyladenosine triphosphate (2-CMA) (29, 30). Specific targeting of RDRP by this chemical led to the elimination of the virus without affecting *Leishmania* fitness. The last approach relied on an early observation that LRV2 capsid overexpression in *L. major* resulted in a significant and stable reduction of viral load (31). The self-assembled virus-like particles have inhibited *Leishmaniavirus* replication in a “dominant negative” manner; in other words, overexpression of native viral capsid proteins substantially interfered with essential processes in host cells. This is highly reminiscent of a description of this phenomenon in classical genetics (32). In line with the observed inhibition of replication, later studies have elegantly demonstrated LRV-facilitated leishmaniasis can be prevented by immunization with its viral capsid (33). A similar approach has been used to successfully eliminate L-A and M dsRNA viruses (family *Totiviridae*) of *Saccharomyces cerevisiae* (34–36).

In the current study, we systematically investigated the dominant negative effect of the capsid protein expression on the fate of LRVs in *L. guyanensis* and *L. major*, the specificity of the underlying mechanism, and the response of the flagellates to virus removal. While the elimination of LRV1 from *L. guyanensis* does not lead to any perceptible consequences for the parasite, LRV2 loss from *L. major* substantially changes the transcription profile and manifests in an attenuated cell division.

RESULTS

Establishment of the *L. guyanensis* lines expressing LRV1 capsid or its derivatives. We employed the standard pLEXSY-based conventional approach to integrate genes encoding LRV1 capsid or its derivatives (Cap-23 and Cap-105) into the 18S rRNA locus of this species as in Zakharova et al. and Ishemgulova et al. (37, 38) (Fig. 1A). The successful integration and capsid expression were confirmed by genomic PCR (Fig. S1), Western blotting (Fig. 1B), and Reverse transcription-quantitative polymerase chain reaction (RT-qPCR) (Fig. 1C to E). As a negative control in RT-qPCR and Western blotting experiments, a cell line of *L. guyanensis* cured of LRV1 (labeled LRV1–) was used (29, 30). Notably, the expression of the rRNA-integrated capsids was comparable to that of

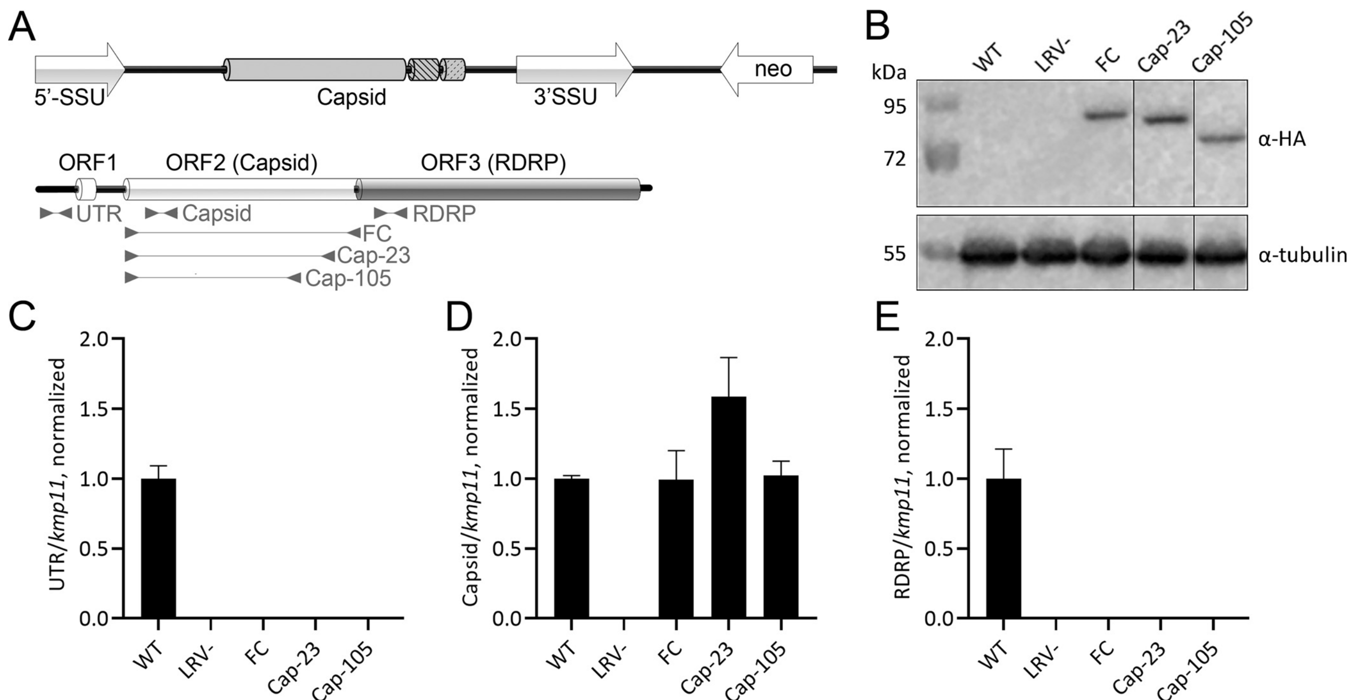


FIG 1 Expression of truncated capsids in *L. guyanensis* M4147. (A) Strategy for generation of the truncated capsids. Top: LRV1-4 integrated into the SSU-rRNA locus. Known capsid domains are represented by different shading and hatching. Bottom: genome organization of LRV1-4 with indicated open reading frames (ORF). Arrowheads indicate the RT-qPCR primer sets ("UTR," "Capsid," and "RDRP") used in expression analyses and PCR primers used to generate wild-type (FC, full capsid), Cap-23, and Cap-105 constructs. (B) Western blotting confirmation of capsids' expression. Sizes are in kDa. (C to E) RT-qPCR analysis of viral load and capsid expression in cultures overexpressing either full capsid or truncated capsid isoforms. Wild-type (WT) and *L. guyanensis* cured of virus (LRV–) were used as positive and negative controls, respectively. Data presented as normalized means and standard deviations of three independent biological replicates.

the endogenous LRV1 as judged by the RT-qPCR analysis with primers annealing to both endogenous and exogenous capsid RNAs ("Capsid" in Fig. 1A lower panel and Fig. 1D).

Overexpression of the full-length or truncated capsid eliminates LRV1 from *L. guyanensis*. Next, we investigated the effect of exogenous capsid (or its derivatives) expression on the fate of endogenous LRV1. Similar to what has been reported for *L. major* and LRV2 (31), the expression of the full-length capsid of LRV1 has eliminated the endogenous virus from *L. guyanensis* (Fig. 1C and E). Note the primer sets used in these analyses ("UTR" and "RDRP"; Fig. 1A lower panel) detected only the endogenous LRV1. The same effect was documented for the Cap-23 and Cap-105 constructs (Fig. 1C and E). Of note, while the elimination of endogenous LRV2 from *L. major* in a previous study was only partial (31), the ablation of LRV1 from *L. guyanensis* was complete.

Dominant-negative effect of the capsid overexpression on LRV1 and LRV2. We used two LRV-positive *Leishmania* spp. (*L. guyanensis* M4147 and *L. major* T44g) and overexpressed capsids of their own viruses (LRV1-4 and LRV2, respectively) as well as those of phylogenetically distant LRV1s from *L. guyanensis Lg*2014 and *L. braziliensis* LEM2700 (21). Successful integration and capsid expression were confirmed by RT-qPCR and Western blotting. While RNA levels of exogenous capsids were similar, the protein levels significantly varied up to the virtually undetectable in the case of LRV1 *Lbr*LEM2700 capsid expressed in *L. major* (Fig. S2). This suggests differential stability of the capsid proteins depending on their sequences (Fig. 2A) and *Leishmania* species/strain, caused, for example, the capsid proteins' ability to trigger autophagy response in *L. major* (and not in *L. guyanensis*).

Expression of all the exogenous (either LRV1 or LRV2) capsids eliminated the endogenous virus from *L. guyanensis* and none of them was able to exert the same effect in

FIG 2 Dominant-negative effect of the capsid overexpression in different *Leishmania* spp. (A) Sequences of the LRV capsids, used in this work, aligned with MAFFT using G-INS-I method and visualized in Jalview using ClustalX color scheme. Amino acids identical to the top sequence are replaced with dots. (B) RT-qPCR analysis of viral load in *L. guyanensis* M4147 and *L. major* T44g cultures expressing different capsid proteins. WT, wild type; 1-4, LRV1-4 of *L. guyanensis* M4147; 2014, LRV1 of *L. guyanensis Lg*2014; 2700, LRV1 of *L. braziliensis Lbr*LEM2700; LRV2 of *L. major* T44g; pLEXSY, empty plasmid control. Data presented as normalized means and standard deviations of three independent biological replicates. *P*-values are denoted as follows: ns, not significant; *, $P \leq 0.05$; **, $P \leq 0.01$; ***, $P \leq 0.001$; ****, $P \leq 0.0001$.

L. major. However, the LRV2 levels in *L. major* significantly decreased (*P* value ≤ 0.05) in the presence of additional LRV2 capsid (Fig. 2B).

Whole-transcriptome analysis of virus-positive and -negative *L. major* and *L. guyanensis* strains. Prompted by the data on stability of the LRV2-*L. major* T44g association, we decided to investigate whether the effect of LRV1/2 ablation is the same in *L. guyanensis* and in *L. major*. For that we used virus-negative isogenic lines of *L. major* T44g and *L. guyanensis* M4147 established by 2-CMA treatment. As a control, we also treated the natively virus-free strain *L. major* LV39 in the same way. After six passages with 2-CMA, followed by another six passages in drug-free medium, no LRV RNA could be detected by RT-qPCR in either of the three strains (Fig. S3).

Previously, it has been demonstrated LRV1-4 ablation from *L. guyanensis* M4147 does not influence the culture growth (Fig. 2 in [29]). In contrast, our data demonstrated that

elimination of LRV2 from *L. major* T44g results in a slower multiplication rate ($P = 0.0078$), although the density reached on day 9 was the same as in the wild-type strain. Notably, the *L. major* LV39 strain used as a control was dividing at a higher rate, which was not affected by 2-CMA treatment (Fig. 3A).

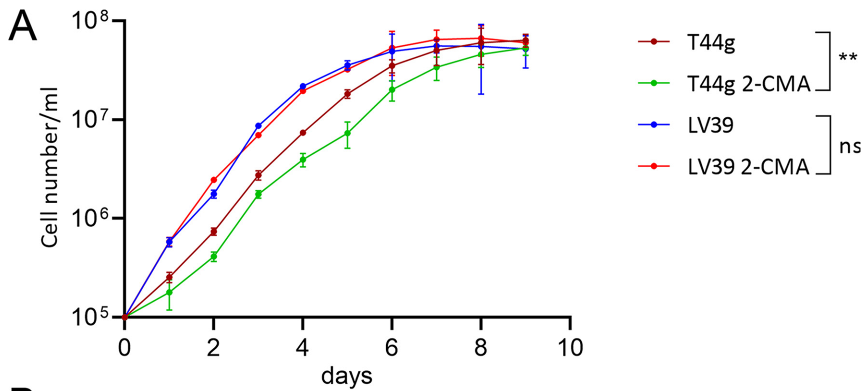
To get insight into the molecular mechanisms behind the differences in the response to viral removal, we compared whole transcriptome profiles of virus-positive and virus-negative *L. major* T44g and *L. guyanensis* M4147. We detected only two differentially expressed genes, which are downregulated in cells upon LRV1-4 ablation when comparing *L. guyanensis* M4147 cell lines with and without LRV1-4. These genes encode a putative subunit of the gamma-tubulin complex (*Lgu_360054900*) and a proteasome activator protein PA26 (*Lgu_350012100*). In stark contrast, 87 differentially expressed genes (67 and 20 up- and downregulated, respectively) were found in LRV-negative *L. major* T44g cells (Table S1). Notably, nine of the proteins, encoded by these genes (all upregulated), formed a well-defined cluster of interaction partners in the STRING analysis (Fig. 3B; number of nodes: 87; number of edges 24; PPI enrichment *P* value 7.59×10^{-11}). We found all genes from this cluster are all upregulated with approximate fold change of $\sim 2 \times$. The gene ontology (GO) enrichment and KEGG analyses identified statistically significant overrepresentation of transcripts involved in cell response to various stimuli, autophagy, and nucleosome assembly among upregulated genes (Fig. S4), while expression of four zinc finger motif-contacting proteins, leucine-rich protein, cyclin, and arginase were found among 20 downregulated genes. The downregulation of cyclin is in a good agreement with the growth kinetics of 2-CMA treated *L. major* cells (Fig. 3A). Also, a reduced expression of the membrane-bound acid phosphatase 2 (Table S1) is noteworthy, as we have recently demonstrated ablation of a related enzyme (phosphatidate phosphatase 2-like protein) has profound effect in *Leishmania* biology (37).

Thus, the ablation of LRV2 from *Leishmania major* T44g has more dramatic outcomes than elimination of LRV1-4 from *L. guyanensis* M4147. This also correlates with the fact that LRV1-4 (but not LRV2) could be completely eliminated by overexpression of capsid of either LRV1 or LRV2 origin.

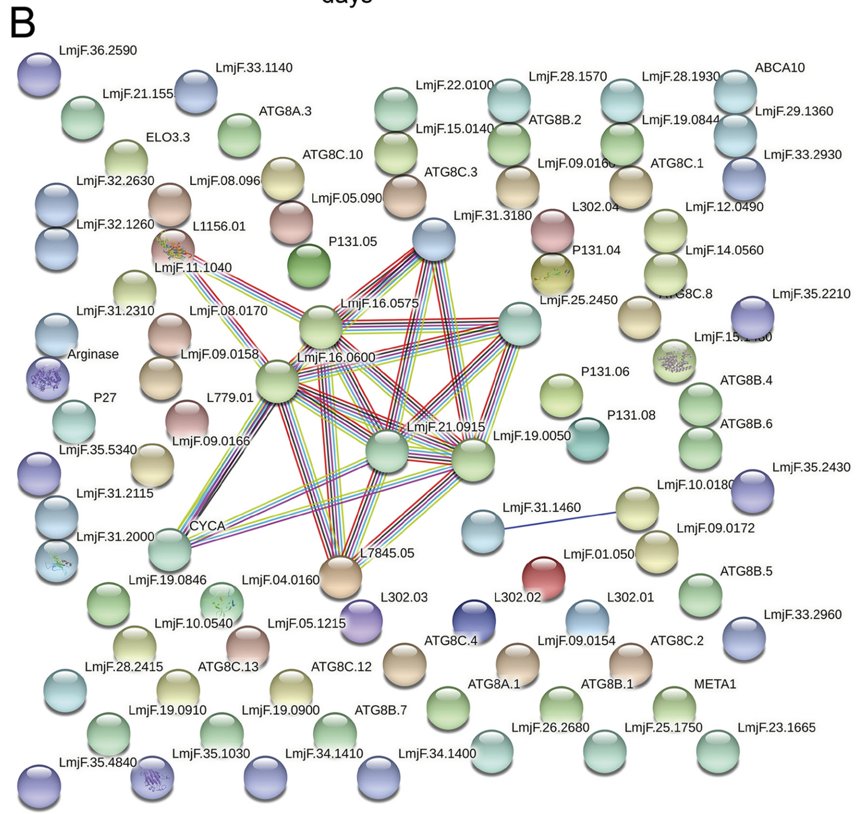
DISCUSSION

The relationships between the LRVs and their flagellate hosts have been previously assessed regarding the importance of the viruses for the virulence of *Leishmania* spp. (16–18, 39, 40). Here, we attempted to get an insight into these associations from a different perspective: the extent of mutual adaptation within them. For this purpose, we used two different strategies of viral removal, which have been proposed before: dominant negative effect by capsid overexpression and 2-CMA treatment. The first approach is based on the disruption of the ideal 120:2 ratio between the capsid and fused (capsid-RDRP) protein, which results in a decreasing proportion of correctly assembled virions and subsequent viral loss (31, 41). We also tested two C-terminally truncated versions that preserved all functional capsid elements (21, 42), because it was demonstrated virus loss can be observed even with an abridged capsid protein in the case of the related L-A virus of yeasts (35). The results in the *L. guyanensis*-LRV1 system led to complete viral elimination, regardless of the experimental construct used, suggesting the C-terminal part of the capsid protein is not necessary for the assembly. Moreover, such an effect could be achieved even with nonnative capsids, which originated either from distantly related LRV1 strains, or, unexpectedly, another viral species—LRV2. These results imply the specificity of the capsid protein interactions is rather limited. In contrast, all the tested exogenous capsids were not able to remove LRV2 from *L. major*. Nevertheless, exogenous expression of the native capsid resulted in a significant decrease in LRV2 level in this species, although apparently to a lesser extent than was reported before for *L. major* strain MHOM/SU/73/5-ASKH (31). Importantly, it was previously demonstrated the decrease in the viral load was associated with LRV absence from the majority of leishmanial cells. The refractoriness of *L. major* to viral removal suggests a more intimate and stable relationship of this species with its virus than that of *L. guyanensis* with LRV1 or

A



B



Edges:

Known Interactions	Predicted Interactions	Others
from curated databases	gene neighborhood	textmining
experimentally determined	gene fusions	co-expression
	gene co-occurrence	protein homology

Functional enrichments in your network

Biological Process (Gene Ontology)

GO-term	description	count in network	strength	false discovery rate
GO:0006914	Autophagy	9 of 27	1.5	1.48e-07
GO:0006334	Nucleosome assembly	4 of 19	1.3	0.0449
GO:0006950	Response to stress	12 of 265	0.64	0.0187

Molecular Function (Gene Ontology)

GO-term	description	count in network	strength	false discovery rate
GO:0046982	Protein heterodimerization activity	7 of 29	1.36	6.41e-05

Cellular Component (Gene Ontology)

GO-term	description	count in network	strength	false discovery rate
GO:0000421	Autophagosome membrane	9 of 11	1.89	7.42e-11
GO:0000786	Nucleosome	7 of 28	1.38	5.74e-06
GO:0031410	Cytoplasmic vesicle	9 of 129	0.82	0.00063

FIG 3 Comparison of virus-positive and virus-negative *L. major* T44g and *L. guyanensis* M4147. (A) Growth curves (see Materials and Methods for experimental details). **, *P* value ≤ 0.01. (B) STRING-based

(Continued on next page)

differences in autophagy-related response to exogenous capsids in these two species. This is further supported by the observation that while the growth of *L. guyanensis* after viral elimination remained unaffected, this resulted in a significant decrease of the proliferation rate in *L. major*. It appears virus removal is deleterious to *L. major* and, as judged by the transcriptome analysis, cells experience considerable stress even after several passages following the loss of the virus. This is rather surprising because until now the main advantage of bearing LRVs was regarded to be associated with the interaction of viruses with the immune system of vertebrate hosts, enhancing the progression of the infection (43). Indeed, it is unclear how the presence of a virus can be important under *in vitro* conditions. The most plausible explanation is the deep *L. major*-LRV2 integration. The usefulness of the virus for the fitness of *L. major* is questionable given a considerable proportion of virus-free strains in natural populations (21, 44). Regardless of the exact nature of this relationship, virus-bearing *L. major* seemingly tuned its cellular processes to coexist with the virus and removal of the latter apparently causes a "phantom pain," i.e., disturbs an established balance.

Interestingly, LRV2 of *L. major* compared with LRV2 from *L. aethiopica* or LRV1s from *L. guyanensis* and *L. braziliensis* demonstrates a higher ratio of nonsynonymous to synonymous substitutions in both capsid and RDRP proteins. This pattern suggests a few sites in these proteins may be under positive selection (21). Therefore, the LRV2s may also intensively adapt to various strains.

The phenomenon of tight LRV2–*L. major* association discovered here poses new questions on the exact underlying molecular mechanisms, which deserve a further scrutiny. It is possible there are continuous or accidental interactions between virus proteins and the host genome transcription/translation processes that cause substantial changes in gene expression. This, in turn, should have an essential effect on *L. major* interactions with its sandfly vectors and mammalian hosts. Thus, our study highlights the profound difference between *L.* guyanensis-LRV1 and *L. major*–LRV2 associations, which presumably determines different impacts of the viral presence on the virulence of these leishmaniae.

MATERIALS AND METHODS

Strains, cultivation, viral elimination, and growth kinetics. The wild-type strains *Leishmania guyanensis* MHOM/BR75/M4147 (in the text referred to as M4147, LRV1-positive [45]), *L. major* MRHO/UZ/2003/IsvT44g (in the text referred to as T44g, LRV2-positive [21]), and *L. major* MRHO/UZ/59/P (in the text referred to as LV39, virus-negative [16]) were cultivated in M199 (MilliporeSigma, Burlington, USA) supplemented with 2 μg/mL hemin (Jena Bioscience, Jena, Germany), 10% heat-inactivated fetal bovine serum (FBS, BioSera Europe, Nuaillé, France), 2 μg/mL biopterin, 100 units/mL of penicillin, and 100 μg/mL of streptomycin (all from Life Technologies/Thermo Fisher Scientific, Carlsbad, USA) at 23°C. Total genomic DNA was isolated from 5×10^7 cells using GeneJET Genomic DNA purification kit (Thermo Fisher Scientific, Carlsbad, USA) and used for 18S rRNA gene amplification and species identity confirmation as described previously (46).

To cure viruses from *L. guyanensis* and *L. major*, cells of the virus-bearing strains were passaged six times in Schneider's Drosophila medium supplemented with 10% FBS, penicillin, streptomycin as above, and 10 mM 2-CMA (29, 30). The virus-negative strain *L. major* LV39 was used as a specificity control in the 2-CMA treatment experiments. To evaluate the efficiency of elimination, the viral load was assayed by quantitative reverse transcription-PCR (RT-qPCR, see below) after six recovery passages in chemical-free medium. Afterwards, the parasites were cultured in complete M199, as defined above. Growth kinetics were analyzed for 9 days from a starting density of 1×10^5 parasites per milliliter. Cell number was counted using a hemocytometer every 24 h as described previously (47) in three biological replicates for each strain/condition.

Genetic manipulations and transfections. Three LRV1 capsid-containing constructs for integration into *L. guyanensis* 18S rRNA locus were designed based on the position of known structural elements (42, 48): the full-length capsid (FC) ended at the frameshift, a Cap-23 variant terminated 23 amino acids

FIG 3 Legend (Continued)

protein–protein network reconstruction. For the known interactions, turquoise line indicates those that came from the curated databases and crimson line indicates those that were experimentally determined; for predicted interactions, green line indicates gene neighborhood, red line indicates gene fusion, blue line gene cooccurrence, black line indicates coexpression, and light blue line indicates protein homology. Empty and filled nodes denote proteins with unknown and known or predicted three-dimensional (3D) structure, respectively. Data on functional enrichment (Biological process, Molecular function, and Cellular component) are tabulated at the bottom.

upstream of it, and a Cap-105 version terminated immediately downstream of the annotated functional domains. See Table S2 for all primer sequences. These three capsid sequences were amplified from cDNA of *L. guyanensis* M4147 and cloned into pLEXSY-Neo2.1 (Jena Bioscience). The same was done for the full-length capsid sequence of LRV2 from *L. major* T44g. In addition, the LRV1 capsid sequences from *L. guyanensis Lg*2014 (NCBI accession number KY750611, labeled 2014) and *L. braziliensis Lbr*LEM2700 (NCBI accession number KX808483, labeled 2700) were synthesized at GeneCust (Boynes, France), and cloned directly into pLEXSY-Neo2.1. The choice of *Lg*2014 and *Lbr*LEM2700 was determined by their phylogenetic remoteness from LRV1-4 of *L. guyanensis* M4147 (labeled 1–4) (21).

For transfection, 5×10^7 *Leishmania* spp. cells were electroporated with 2 to 5 μg of SwaI-linearized plasmids using Nucleofector-2b (Lonza Bioscience, Basel, Switzerland) and program X-001. Transfected cells were incubated in complete M199 medium at 23°C: initially without antibiotic for 16 h and then with 50 to 100 μg/mL of Neomycin (VWR, Radnor, USA) for 3 weeks.

Isolation of RNA, cDNA synthesis, and RT-qPCR. Total RNA was isolated from 5×10^7 cells in three biological replicates using the RNeasy minikit (Qiagen, Hilden, Germany) following the manufacturer's recommendations. The cDNA was synthesized with random hexamer primers using the Transcriptor First Strand cDNA Synthesis Kit (Roche Life Science, Penzberg, Germany) following the manufacturer's instructions. Capsid and RDRP expression were measured by RT-qPCR as described previously (49) using LightCycler 480 (Roche Life Science). All experiments were performed in biological (three randomly selected populations) and technical triplicates. Expression levels of genes of interest were normalized to the housekeeping gene kinetoplast membrane protein-11 (KMP11) (50) and shown relative to the wild type.

Western blotting. To validate capsid expression in different *Leishmania* spp. populations, lysates from approximately 1×10^7 cells were probed with anti-HA and anti-tubulin antibodies (both from MilliporeSigma) at 1:1,000 and 1:5,000 dilutions, respectively, as in Kraeva et al. (51).

Differential expression analysis of LVR-ablated and wild-type *Leishmania* spp. Transcriptomes of *L. major* T44g wild-type and LRV2-ablated cells after six passages in 2-CMA-free media were sequenced in four independent biological replicates each. Transcriptomes of *L. guyanensis* M4147 wild-type and LRV1-ablated cells after six passages in 2-CMA-free media were sequenced in two series yielding seven independent biological replicates for each line in total. All samples were sequenced in paired-end mode on Illumina NovaSeq with read length of 150 bp. Reads were trimmed with Trimmomatic v. 0.39 (52), using "SLIDINGWINDOW: 10:25" and "TRAILING: 25" trimming functions and mapped on reference genome sequence with Bowtie2 (53) using "–sensitive-local –no-unal." The reference sequence and annotation for *L. major* (strain Friedlin) were taken from the TriTrypDB release 54 (54). The reference sequence and annotation for *L. guyanensis* (strain 204) were downloaded from the NCBI (accession number GCA_003664525). Read counting was performed with BEDTools v. 2.30 (55). Analysis of differential expression was performed in R v. 3.5.1 with EdgeR package v. 3.26 (56), genes with overall low counts were filtered out with "filterByExpr (min.count =10, min.total.count =30)" function. Differentially expressed gene lists were generated using FDR-corrected *P*-value cut-off 0.01 and fold change cut-off 2. GO enrichment analysis was done with the g:Profiler2 v. 0.2.1 package (57).

Network reconstruction. The obtained differentially expressed genes were incorporated into the STRING v. 11.5 (58) and network reconstruction was performed on the basis of the corresponding proteins. Specific and meaningful protein–protein associations were indicated by edges. The interactions from the curated databases and those that were experimentally determined were included. The protein–protein interactions (PPI) enrichment *P* value was used to verify whether observed number of edges is significant and the nodes are not random.

Statistical analyses. The statistical analyses were performed using GraphPad Prism v. 9 (GraphPad Software, San Diego, USA). The two-tailed Student's *t* test was applied for the analysis of RT-qPCR data. The growth curves were analyzed using two-tailed Wilcoxon matched-pairs signed rank test.

SUPPLEMENTAL MATERIAL

Supplemental material is available online only.

FIG S1, PDF file, 0.1 MB.

FIG S2, PDF file, 0.2 MB.

FIG S3, PDF file, 0.1 MB.

FIG S4, PDF file, 0.9 MB.

TABLE S1, XLSX file, 0.01 MB.

TABLE S2, XLSX file, 0.01 MB.

ACKNOWLEDGMENTS

This work was supported by the Grant Agency of Czech Republic 20-22689S to V.Y. and the European Regional Development Funds (CZ.02.1.01/0.0/0.0/16_019/0000759) to V.Y., A.S., A.Y.K., and A.B.; grants SGS/PřF/2022 from the University of Ostrava and RRC/10/2019 from the Moravskoslezský kraj research initiative to A.Z., who was also supported by the Ministry of Education, Youth and Sports of the Czech Republic in the frame of the Operational Program "Research, Development and Education" (CZ.02.2.69/

0.0/0.0/19_073/0016939). I.B.R. was supported by the Intramural Research Program of the National Library of Medicine at the National Institutes of Health. The funders had no role in study design, data collection and interpretation, or the decision to submit the work for publication. Individual author contributions were as follows: planning of the study, V.Y.; genetic manipulations, A.S., A.Z., D.K., D.H.M., and D.Z.; data analysis and bioinformatics, E.S.G., A.B., E.S., A.M., I.B.R., S.S., and A.Y.K.; and writing of the manuscript, V.Y. and A.Y.K., with additions from all other authors.

REFERENCES

1. WHO. 2020. Leishmaniasis. https://www.who.int/en/news-room/fact-sheets/detail/leishmaniasis. Accessed July 02, 2022.
2. Bruschi F, Gradoni L. 2018. The leishmaniases: old neglected tropical diseases. Springer, Cham, Switzerland. https://doi.org/10.1007/978-3-319-72386-0.
3. Kaye P, Scott P. 2011. Leishmaniasis: complexity at the host-pathogen interface. Nat Rev Microbiol 9:604–615. https://doi.org/10.1038/nrmicro2608.
4. Ghabrial SA, Castón JR, Jiang D, Nibert ML, Suzuki N. 2015. 50-plus years of fungal viruses. Virology 479-480:356–368. https://doi.org/10.1016/j.virol.2015.02.034.
5. Poulos BT, Tang KF, Pantoja CR, Bonami JR, Lightner DV. 2006. Purification and characterization of infectious myonecrosis virus of penaeid shrimp. J Gen Virol 87:987–996. https://doi.org/10.1099/vir.0.81127-0.
6. Zhai Y, Attoui H, Mohd Jaafar F, Wang HQ, Cao YX, Fan SP, Sun YX, Liu LD, Mertens PP, Meng WS, Wang D, Liang G. 2010. Isolation and full-length sequence analysis of *Armigeres subalbatus* totivirus, the first totivirus isolate from mosquitoes representing a proposed novel genus (*Artivirus*) of the family *Totiviridae*. J Gen Virol 91:2836–2845. https://doi.org/10.1099/vir.0.024794-0.
7. Løvoll M, Wiik-Nielsen J, Grove S, Wiik-Nielsen CR, Kristoffersen AB, Faller R, Poppe T, Jung J, Pedamallu CS, Nederbragt AJ, Meyerson M, Rimstad E, Tengs T. 2010. A novel totivirus and piscine reovirus (PRV) in Atlantic salmon (*Salmo salar*) with cardiomyopathy syndrome (CMS). Virol J 7:309. https://doi.org/10.1186/1743-422X-7-309.
8. Xin C, Wu B, Li J, Gong P, Yang J, Li H, Cai X, Zhang X. 2016. Complete genome sequence and evolution analysis of *Eimeria stiedai RNA virus 1*, a novel member of the family *Totiviridae*. Arch Virol 161:3571–3576. https://doi.org/10.1007/s00705-016-3020-7.
9. Janssen ME, Takagi Y, Parent KN, Cardone G, Nibert ML, Baker TS. 2015. Three-dimensional structure of a protozoal double-stranded RNA virus that infects the enteric pathogen *Giardia lamblia*. J Virol 89:1182–1194. https://doi.org/10.1128/JVI.02745-14.
10. Grybchuk D, Kostygov AY, Macedo DH, d'Avila-Levy CM, Yurchenko V. 2018. RNA viruses in trypanosomatid parasites: a historical overview. Mem Inst Oswaldo Cruz 113:e170487.
11. Wang AL, Wang CC. 1986. Discovery of a specific double-stranded RNA virus in *Giardia lamblia*. Mol Biochem Parasitol 21:269–276. https://doi.org/10.1016/0166-6851(86)90132-5.
12. Hartley MA, Ronet C, Zangger H, Beverley SM, Fasel N. 2012. *Leishmania RNA virus*: when the host pays the toll. Front Cell Infect Microbiol 2:99. https://doi.org/10.3389/fcimb.2012.00099.
13. Brettmann EA, Shaik JS, Zangger H, Lye LF, Kuhlmann FM, Akopyants NS, Oschwald DM, Owens KL, Hickerson SM, Ronet C, Fasel N, Beverley SM. 2016. Tilting the balance between RNA interference and replication eradicates *Leishmania RNA virus 1* and mitigates the inflammatory response. Proc Natl Acad Sci U S A 113:11998–12005. https://doi.org/10.1073/pnas.1615085113.
14. Cantanhêde LM, Mata-Somarribas C, Chourabi K, Pereira da Silva G, Dias das Chagas B, de Oliveira RPL, Cortes Boite M, Cupolillo E. 2021. The maze pathway of coevolution: a critical review over the *Leishmania* and its endosymbiotic history. Genes 12:657. https://doi.org/10.3390/genes12050657.
15. Grybchuk D, Kostygov AY, Macedo DH, Votypka J, Lukes J, Yurchenko V. 2018. RNA viruses in *Blechomonas* (Trypanosomatidae) and evolution of *Leishmaniavirus*. mBio 9:e01932-18. https://doi.org/10.1128/mBio.01932-18.
16. Ives A, Ronet C, Prevel F, Ruzzante G, Fuertes-Marraco S, Schutz F, Zangger H, Revaz-Breton M, Lye LF, Hickerson SM, Beverley SM, Acha-Orbea H, Launois P, Fasel N, Masina S. 2011. *Leishmania* RNA virus controls the severity of mucocutaneous leishmaniasis. Science 331:775–778. https://doi.org/10.1126/science.1199326.
17. Rossi M, Castiglioni P, Hartley MA, Eren RO, Prevel F, Desponds C, Utzschneider DT, Zehn D, Cusi MG, Kuhlmann FM, Beverley SM, Ronet C, Fasel N. 2017. Type I interferons induced by endogenous or exogenous viral infections promote metastasis and relapse of leishmaniasis. Proc Natl Acad Sci U S A 114:4987–4992. https://doi.org/10.1073/pnas.1621447114.
18. Hartley MA, Bourreau E, Rossi M, Castiglioni P, Eren RO, Prevel F, Couppie P, Hickerson SM, Launois P, Beverley SM, Ronet C, Fasel N. 2016. *Leishmaniavirus*-dependent metastatic leishmaniasis is prevented by blocking IL-17A. PLoS Pathog 12:e1005852. https://doi.org/10.1371/journal.ppat.1005852.
19. Barrow P, Dujardin JC, Fasel N, Greenwood AD, Osterrieder K, Lomonossoff G, Fiori PL, Atterbury R, Rossi M, Lalle M. 2020. Viruses of protozoan parasites and viral therapy: is the time now right? Virol J 17:142. https://doi.org/10.1186/s12985-020-01410-1.
20. Widmer G, Dooley S. 1995. Phylogenetic analysis of *Leishmania RNA virus* and *Leishmania* suggests ancient virus-parasite association. Nucleic Acids Res 23:2300–2304. https://doi.org/10.1093/nar/23.12.2300.
21. Kostygov AY, Grybchuk D, Kleschenko Y, Chistyakov DS, Lukashev AN, Gerasimov ES, Yurchenko V. 2021. Analyses of *Leishmania*-LRV co-phylogenetic patterns and evolutionary variability of viral proteins. Viruses 13:2305. https://doi.org/10.3390/v13112305.
22. Scheffter S, Widmer G, Patterson JL. 1994. Complete sequence of *Leishmania RNA virus 1–4* and identification of conserved sequences. Virology 199:479–483. https://doi.org/10.1006/viro.1994.1149.
23. Scheffter SM, Ro YT, Chung IK, Patterson JL. 1995. The complete sequence of *Leishmania RNA virus* LRV2-1, a virus of an Old World parasite strain. Virology 212:84–90. https://doi.org/10.1006/viro.1995.1456.
24. Adaui V, Lye LF, Akopyants NS, Zimic M, Llanos-Cuentas A, Garcia L, Maes I, De Doncker S, Dobson DE, Arevalo J, Dujardin JC, Beverley SM. 2016. Association of the endobiont double-stranded RNA virus LRV1 with treatment failure for human leishmaniasis caused by *Leishmania braziliensis* in Peru and Bolivia. J Infect Dis 213:112–121. https://doi.org/10.1093/infdis/jiv354.
25. Abtahi M, Eslami G, Cavallero S, Vakili M, Hosseini SS, Ahmadian S, Boozhmehrani MJ, Khamesipour A. 2020. Relationship of *Leishmania RNA Virus* (LRV) and treatment failure in clinical isolates of *Leishmania major*. BMC Res Notes 13:126. https://doi.org/10.1186/s13104-020-04973-y.
26. Bourreau E, Ginouves M, Prevot G, Hartley MA, Gangneux JP, Robert-Gangneux F, Dufour J, Sainte-Marie D, Bertolotti A, Pratlong F, Martin R, Schutz F, Couppie P, Fasel N, Ronet C. 2016. Presence of *Leishmania RNA Virus 1* in *Leishmania guyanensis* increases the risk of first-line treatment failure and symptomatic relapse. J Infect Dis 213:105–111. https://doi.org/10.1093/infdis/jiv355.
27. Ro YT, Scheffter SM, Patterson JL. 1997. Hygromycin B resistance mediates elimination of *Leishmania* virus from persistently infected parasites. J Virol 71:8991–8998. https://doi.org/10.1128/JVI.71.12.8991-8998.1997.
28. Lye LF, Owens K, Shi H, Murta SM, Vieira AC, Turco SJ, Tschudi C, Ullu E, Beverley SM. 2010. Retention and loss of RNA interference pathways in trypanosomatid protozoans. PLoS Pathog 6:e1001161. https://doi.org/10.1371/journal.ppat.1001161.
29. Kuhlmann FM, Robinson JI, Bluemling GR, Ronet C, Fasel N, Beverley SM. 2017. Antiviral screening identifies adenosine analogs targeting the endogenous dsRNA *Leishmania RNA virus 1* (LRV1) pathogenicity factor. Proc Natl Acad Sci U S A 114:E811–E819. https://doi.org/10.1073/pnas.1619114114.
30. Robinson JI, Beverley SM. 2018. Concentration of 2'C-methyladenosine triphosphate by *Leishmania guyanensis* enables specific inhibition of *Leishmania RNA virus 1 via* its RNA polymerase. J Biol Chem 293:6460–6469. https://doi.org/10.1074/jbc.RA117.001515.
31. Widmer G. 1995. Suppression of *Leishmania RNA virus* replication by capsid protein overexpression. J Virol 69:4122–4126. https://doi.org/10.1128/JVI.69.7.4122-4126.1995.
32. Herskowitz I. 1987. Functional inactivation of genes by dominant negative mutations. Nature 329:219–222. https://doi.org/10.1038/329219a0.

33. Castiglioni P, Hartley MA, Rossi M, Prevel F, Desponds C, Utzschneider DT, Eren RO, Zangger H, Brunner L, Collin N, Zehn D, Kuhlmann FM, Beverley SM, Fasel N, Ronet C. 2017. Exacerbated leishmaniasis caused by a viral endosymbiont can be prevented by immunization with its viral capsid. PLoS Negl Trop Dis 11:e0005240. https://doi.org/10.1371/journal.pntd.0005240.
34. Lukša J, Ravoitytė B, Konovalovas A, Aitmanaitė L, Butenko A, Yurchenko V, Serva S, Servienė E. 2017. Different metabolic pathways are involved in response of *Saccharomyces cerevisiae* to L-A and M viruses. Toxins 9:e233. https://doi.org/10.3390/toxins9080233.
35. Aitmanaitė L, Konovalovas A, Medvedevas P, Servienė E, Serva S. 2021. Specificity determination in *Saccharomyces cerevisiae* killer virus systems. Microorganisms 9:236. https://doi.org/10.3390/microorganisms9020236.
36. Yao W, Bruenn JA. 1995. Interference with replication of two double-stranded RNA viruses by production of N-terminal fragments of capsid polypeptides. Virology 214:215–221. https://doi.org/10.1006/viro.1995.9938.
37. Zakharova A, Albanaz ATS, Opperdoes FR, Škodová-Sveráková I, Zagirova D, Saura A, Chmelová L, Gerasimov ES, Leštinová T, Bečvář T, Sádlová J, Volf P, Lukeš J, Horváth A, Butenko A, Yurchenko V. 2022. *Leishmania guyanensis* M4147 as a new LRV1-bearing model parasite: phosphatidate phosphatase 2-like protein controls cell cycle progression and intracellular lipid content. PLoS Negl Trop Dis 16:e0010510. https://doi.org/10.1371/journal.pntd.0010510.
38. Ishemgulova A, Kraeva N, Hlaváčová J, Zimmer SL, Butenko A, Podešvová L, Leštinová T, Lukeš J, Kostygov A, Votýpka J, Volf P, Yurchenko V. 2017. A putative ATP/GTP binding protein affects *Leishmania mexicana* growth in insect vectors and vertebrate hosts. PLoS Negl Trop Dis 11:e0005782. https://doi.org/10.1371/journal.pntd.0005782.
39. Eren RO, Reverte M, Rossi M, Hartley MA, Castiglioni P, Prevel F, Martin R, Desponds C, Lye LF, Drexler SK, Reith W, Beverley SM, Ronet C, Fasel N. 2016. Mammalian innate immune response to a *Leishmania*-resident RNA virus increases macrophage survival to promote parasite persistence. Cell Host Microbe 20:318–328. https://doi.org/10.1016/j.chom.2016.08.001.
40. Zangger H, Hailu A, Desponds C, Lye LF, Akopyants NS, Dobson DE, Ronet C, Ghalib H, Beverley SM, Fasel N. 2014. *Leishmania aethiopica* field isolates bearing an endosymbiontic dsRNA virus induce pro-inflammatory cytokine response. PLoS Negl Trop Dis 8:e2836. https://doi.org/10.1371/journal.pntd.0002836.
41. Dinman JD, Wickner RB. 1992. Ribosomal frameshifting efficiency and gag/gag-pol ratio are critical for yeast M1 double-stranded RNA virus propagation. J Virol 66:3669–3676. https://doi.org/10.1128/JVI.66.6.3669-3676.1992.
42. Procházková M, Füzik T, Grybchuk D, Falginella FL, Podešvová L, Yurchenko V, Vácha R, Plevka P. 2021. Capsid structure of *Leishmania RNA Virus 1*. J Virol 95:e01957-20. https://doi.org/10.1128/JVI.01957-20.
43. de Carvalho RVH, Lima-Junior DS, da Silva MVG, Dilucca M, Rodrigues TS, Horta CV, Silva ALN, da Silva PF, Frantz FG, Lorenzon LB, Souza MM, Almeida F, Cantanhede LM, Ferreira RGM, Cruz AK, Zamboni DS. 2019. *Leishmania RNA virus* exacerbates leishmaniasis by subverting innate immunity via TLR3-mediated NLRP3 inflammasome inhibition. Nat Commun 10:5273. https://doi.org/10.1038/s41467-019-13356-2.
44. Saberi R, Fakhar M, Hajjaran H, Ataei-Pirkooh A, Mohebali M, Taghipour N, Ziaei Hezarjaribi H, Moghadam Y, Bagheri A. 2020. Presence and diversity of *Leishmania RNA virus* in an old zoonotic cutaneous leishmaniasis focus, northeastern Iran: haplotype and phylogenetic based approach. Int J Infect Dis 101:6–13. https://doi.org/10.1016/j.ijid.2020.08.033.
45. Widmer G, Comeau AM, Furlong DB, Wirth DF, Patterson JL. 1989. Characterization of a RNA virus from the parasite *Leishmania*. Proc Natl Acad Sci U S A 86:5979–5982. https://doi.org/10.1073/pnas.86.15.5979.
46. Yurchenko V, Lukeš J, Xu X, Maslov DA. 2006. An integrated morphological and molecular approach to a new species description in the Trypanosomatidae: the case of *Leptomonas podlipaevi* n. sp., a parasite of *Boisea rubrolineata* (Hemiptera: Rhopalidae). J Eukaryotic Microbiology 53:103–111. https://doi.org/10.1111/j.1550-7408.2005.00078.x.
47. Ishemgulova A, Hlaváčová J, Majerová K, Butenko A, Lukeš J, Votýpka J, Volf P, Yurchenko V. 2018. CRISPR/Cas9 in *Leishmania mexicana*: a case study of LmxBTN1. PLoS One 13:e0192723. https://doi.org/10.1371/journal.pone.0192723.
48. Lee SE, Suh JM, Scheffter S, Patterson JL, Chung IK. 1996. Identification of a ribosomal frameshift in *Leishmania RNA virus* 1–4. J Biochem 120:22–25. https://doi.org/10.1093/oxfordjournals.jbchem.a021387.
49. Záhonová K, Hadariová L, Vacula R, Yurchenko V, Eliáš M, Krajčovič J, Vesteg M. 2014. A small portion of plastid transcripts is polyadenylated in the flagellate *Euglena gracilis*. FEBS Lett 588:783–788. https://doi.org/10.1016/j.febslet.2014.01.034.
50. Basu R, Bhaumik S, Basu JM, Naskar K, De T, Roy S. 2005. Kinetoplastid membrane protein-11 DNA vaccination induces complete protection against both pentavalent antimonial-sensitive and -resistant strains of *Leishmania donovani* that correlates with inducible nitric oxide synthase activity and IL-4 generation: evidence for mixed Th1- and Th2-like responses in visceral leishmaniasis. J Immunol 174:7160–7171. https://doi.org/10.4049/jimmunol.174.11.7160.
51. Kraeva N, Ishemgulova A, Lukeš J, Yurchenko V. 2014. Tetracycline-inducible gene expression system in *Leishmania mexicana*. Mol Biochem Parasitol 198:11–13. https://doi.org/10.1016/j.molbiopara.2014.11.002.
52. Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30:2114–2120. https://doi.org/10.1093/bioinformatics/btu170.
53. Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. Nat Methods 9:357–359. https://doi.org/10.1038/nmeth.1923.
54. Aslett M, Aurrecoechea C, Berriman M, Brestelli J, Brunk BP, Carrington M, Depledge DP, Fischer S, Gajria B, Gao X, Gardner MJ, Gingle A, Grant G, Harb OS, Heiges M, Hertz-Fowler C, Houston R, Innamorato F, Iodice J, Kissinger JC, Kraemer E, Li W, Logan FJ, Miller JA, Mitra S, Myler PJ, Nayak V, Pennington C, Phan I, Pinney DF, Ramasamy G, Rogers MB, Roos DS, Ross C, Sivam D, Smith DF, Srinivasamoorthy G, Stoeckert CJ, Jr., Subramanian S, Thibodeau R, Tivey A, Treatman C, Velarde G, Wang H. 2010. TriTrypDB: a functional genomic resource for the Trypanosomatidae. Nucleic Acids Res 38:D457–62. https://doi.org/10.1093/nar/gkp851.
55. Quinlan AR, Hall IM. 2010. BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics 26:841–842. https://doi.org/10.1093/bioinformatics/btq033.
56. Robinson MD, McCarthy DJ, Smyth GK. 2010. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 26:139–140. https://doi.org/10.1093/bioinformatics/btp616.
57. Raudvere U, Kolberg L, Kuzmin I, Arak T, Adler P, Peterson H, Vilo J. 2019. g:Profiler: a web server for functional enrichment analysis and conversions of gene lists (2019 update). Nucleic Acids Res 47:W191–W198. https://doi.org/10.1093/nar/gkz369.
58. Szklarczyk D, Gable AL, Lyon D, Junge A, Wyder S, Huerta-Cepas J, Simonovic M, Doncheva NT, Morris JH, Bork P, Jensen LJ, Mering CV. 2019. STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. Nucleic Acids Res 47:D607–D613. https://doi.org/10.1093/nar/gky1131.