

## APPENDICES

**Appendix A. European Nucleotide Archive (ENA) accession numbers for all samples sequenced for this dissertation. For RNA-seq samples, the Array Express accession number is E-ERAD-408. The ENA accession numbers for all RNA-seq triplicate samples were from ERS763603 to ERS763656.**

<b>Sample</b>	<b>ENA Accession Number</b>
Melloy	ERS364123
LRC-L747	ERS364124
KK27	ERS364125
Ackerman	ERS364126
K26_1	ERS364127
LRC-L810	ERS364128
MN-11	ERS364129
Rupert	ERS364130
R55	ERS364131
Kubba	ERS364132
Boone	ERS364133
Azad	ERS364134
E50	ERS364135
MA-37	ERS364136
K112_1	ERS364137
Sultan	ERS364138
Melloy	ERS364123
LRC-L747	ERS364124
KK27	ERS364125
Ackerman	ERS364126
K26_1	ERS364127
LRC-L810	ERS364128
SAF-K27	ERS198200
LRC-L838	ERS198235
Adhanis I	ERS214826
L75	ERS198220
SAF-K27	ERS198200
LRC-L838	ERS198235
Rupert C1	ERS471347
Rupert C2	ERS471346
MA-37 C1	ERS763658
MN-11 C1	ERS471344
MN-11 C2	ERS471343
Kubba C1	ERS471342
L590	ERS218438
L747 HYG	ERS763657
MA-37 NEO	ERS763658
H1a	ERS763659
H1b	ERS763664
H2a	ERS763665
H2b	ERS763660
H3a	ERS763666
H3b	ERS763667
H4a	ERS763661
H4b	ERS763668
H5a	ERS763669
H5b	ERS763662

**Appendix B. Inbreeding coefficients (FIT) and clustering results for the 34 isolates analysed in Chapter 2. Isolates are ranked from highest to lowest inbreeding coefficient. The inbreeding coefficient goes from 0 to 1 and increases with decreasing heterozygosity compared to HW expectations.**

<b>Sample</b>	<b>F<sub>IT</sub></b>	<b>Cluster</b>
LA28	0.7996	Cluster 1
L747	0.7361	Cluster 2
Killicki	0.7268	Cluster 3
MN-11	0.6819	Cluster 1
L810	0.6071	Cluster 3
Tropica63	0.5603	Cluster 2
MA-37	0.5529	Cluster 1
Leep0920	0.5424	Cluster 3
Tropica75	0.5420	Cluster 2
Tropica57	0.5365	Cluster 2
AM	0.5031	Cluster 3
DBKM	0.4571	Cluster 1
CJ	0.3629	Cluster 3
E50	0.3096	Cluster 2
Azad	0.2406	Cluster 1
Boone	0.1625	Cluster 1
Gabai	0.1439	Cluster 1
Melloy	0.1340	Cluster 1
IIKK	0.1094	Cluster 1
MON497px3	0.1078	Cluster 1
Bum30	0.1064	Cluster 1
WR683	0.1063	Cluster 1
K27	0.1040	Cluster 1
Ackerman	0.1010	Cluster 1
Adhanis I	0.1005	Cluster 1
Kubba	0.09649	Cluster 1
188	0.09235	Cluster 1
Asinai III	0.09151	Cluster 1
Rachnan	0.09119	Cluster 1
L75	0.08906	Cluster 1
311W	0.08197	Cluster 1
Bag17	0.07922	Cluster 1
Rupert	0.07415	Cluster 1
Bag9	0.06428	Cluster 1

**Appendix C. Most highly expressed genes in the promastigote stage from RNA-seq of the isolates analysed in Chapter 3. Mean VSD is the mean variance-stabilized dispersion across the set of isolates. Genes are ranked from most to least expressed.**

Gene	Product	Mean VSD
LmjF.21.1860	Beta tubulin	16.41689
LmjF.36.1940	Inosine/guanosine transporter (NT2)	15.23829
LmjF.15.1230	Nucleoside transporter 1, putative	14.59714
LmjF.15.1240	Nucleoside transporter 1, putative	14.45777
LmjF.35.1890	60S ribosomal protein L5	14.36469
LmjF.32.2660	L-lysine transport protein (AAT16)	13.86768
LmjF.21.0915	Histone H2A	13.82777
LmjF.25.0720	Eukaryotic initiation factor 5a (EIF5A1)	13.82076
LmjF.11.0970	40S ribosomal protein S5	13.76264
LmjF.32.3130	Ribosomal protein L3	13.7539
LmjF.13.0450	Alba-like nucleic acid binding protein	13.64462
LmjF.31.0350	Aminoacid transporter (AAT1.4)	13.42285
LmjF.36.6300	Glucose transporter 1 (GT1)	13.52832
LmjF.22.0420	40S ribosomal protein S15	13.58769
LmjF.32.3900	60S ribosomal protein L2/L8	13.52428
LmjF.35.0240	60S ribosomal protein L30	13.53991
LmjF.32.0750	RNP1-like RNA binding protein	13.52044
LmjF.06.1260	Pteridine transporter, putative	13.40097
LmjF.36.0980	40S ribosomal protein S10	13.46978
LmjF.14.1160	Enolase	13.44339
LmjF.24.2070	40S ribosomal protein S8	13.42711
LmjF.27.1380	60S acidic ribosomal protein P0	13.37493
LmjF.05.0500	ATPase alpha subunit	13.30601
LmjF.25.0910	Cyclophilin a (CYPA)	13.24228
LmjF.13.0570	40S ribosomal protein S12	13.21298
LmjF.28.2555	40S ribosomal protein S17	13.21329
LmjF.06.0570	60S ribosomal protein L23a	13.18204
LmjF.27.1245	Carboxypeptidase-like protein	13.15227
LmjF.36.1260	Fructose-1,6-biphosphate aldolase (ALD)	13.09008
LmjF.07.1160	Aminoacid transporter (AAT19)	12.75687

**Appendix D. List of DE genes in L810 compared to all other isolates (excluding MN-11 C2 and Boone). Only genes that had significant p-values and that had a log-fold change greater than 2 or less than -2 were considered.**

Gene	Log-FC	P-value	Product
LmjF.02.0400	2.081753	2.88E-10	Transmembrane hypothetical protein, unknown function
LmjF.04.0310	4.434332	0.00E+00	beta-fructofuranosidase, putative
LmjF.04.0320	2.754353	2.36E-175	beta-fructofuranosidase, putative
LmjF.04.1050	10.20729	0.00E+00	acyltransferase-like protein, copy 2
LmjF.09.0003	6.314717	1.00E+00	hypothetical protein, conserved
LmjF.09.0690	2.939361	3.84E-20	hypothetical protein, conserved
LmjF.17.0086	5.799281	7.53E-186	elongation factor 1-alpha
LmjF.17.0730	3.260571	3.68E-87	transmembrane hypothetical protein, conserved
LmjF.17.0733	4.468265	2.24E-05	hypothetical protein, conserved
LmjF.17.0740	2.220649	3.42E-129	Transmembrane hypothetical protein, conserve
LmjF.20.0830	2.340252	3.81E-115	phosphopantetheinyl transferase-like protein
LmjF.20.1175	5.285097	1.77E-21	hypothetical protein
LmjF.21.0015	2.434106	9.56E-35	histone H4
LmjF.26.0640	2.126129	1.79E-109	10 kDa heat shock protein, putative
LmjF.28.0350	2.237566	8.52E-47	Transmembrane hypothetical protein, unknown function
LmjF.30.2460	2.768505	7.25E-132	heat shock 70-related protein 1, mitochondrial precursor, putative
LmjF.30.2480	2.23119	8.076193e-316	heat shock 70-related protein 1, mitochondrial precursor, putative
LmjF.30.2490	3.514708	7.40E-121	heat shock 70-related protein 1, mitochondrial precursor, putative
LmjF.31.0420	2.987786	8.51E-99	Transmembrane hypothetical protein, unknown function
LmjF.31.0470	2.598772	6.07E-72	hypothetical protein, conserved
LmjF.31.2680	3.515456	1.00E+00	hypothetical protein, unknown function
LmjF.32.1380	4.66376	5.53E-111	hypothetical protein, conserved
LmjF.32.1390	2.155002	4.45E-274	hypothetical protein, conserved
LmjF.33.1630	2.074137	3.27E-211	cyclophilin 4, putative
LmjF.35.0450	10.170533	3.96E-14	hypothetical protein, unknown function
LmjF.35.2600	2.589783	5.11E-14	hypothetical protein, unknown function
LmjF.03.0010	-2.06585	8.30E-208	hypothetical protein
LmjF.04.1210	-2.220235	9.88E-114	casein kinase I, putative
LmjF.05.0360	-2.059694	4.23E-03	ATP-dependent RNA helicase, putative
LmjF.13.0100	-2.635757	1.43E-220	excreted hypothetical protein, unknown function
LmjF.16.0480	-2.353516	0.00E+00	fucose kinase, putative
LmjF.17.0190	-2.709674	6.31E-226	receptor-type adenylate cyclase, putative
LmjF.22.0750	-2.252003	1.26E-98	CCCH-type zinc finger hypothetical protein, conserved
LmjF.23.0730	-2.481566	1.17E-68	RNA-binding protein, putative
LmjF.23.1665	-2.321696	0.00E+00	Transmembrane hypothetical phosphatidic acid phosphatase type-2
LmjF.25.1730	-2.018702	0.00E+00	signal peptide-containing hypothetical protein, conserved
LmjF.26.2680	-2.075803	5.60E-182	Excreted hypothetical protein, unknown function
LmjF.28.1570	-2.253956	0.00E+00	hydrolase, alpha/beta fold family, putative
LmjF.28.2660	-2.437439	5.97E-245	hypothetical methionine sulphoxide reductase B protein, conserved
LmjF.31.2100	-2.711269	7.20E-223	hypothetical protein, unknown function
LmjF.31.2460	-2.050621	1.09E-116	lipase, putative
LmjF.33.0520	-2.486323	0.00E+00	d-xylulose reductase, putative
LmjF.35.0710	-2.056016	2.44E-186	Nuclease-like hypothetical protein, conserved
LmjF.35.2810	-2.036244	5.89E-298	MFS-1 transporter, putative
LmjF.35.4240	-2.17067	2.01E-76	hypothetical protein, conserved
LmjF.35.5030	-2.534736	0.00E+00	hypothetical protein, conserved
LmjF.36.2760	-2.464428	2.39E-168	MFS general substrate transporter protein

**Appendix E. List of DE genes between the two clones of isolate Rupert that were compared in Chapter 3 that also overlap with CNVs identified in this study. LogFC is the log-fold change in expression between the two clones, representing the difference between Rupert C2 and Rupert C1. Genes with positive logFC are more highly expressed in Rupert C2 than Rupert C1. Log-2-ratio is the ratio between the average read depth for that gene in Rupert C2 and Rupert C1. Positive Log-2-ratio means that a CNV has either amplified the gene in Rupert C2 or deleted a copy of the gene in Rupert C1.**

Gene	Product	Log-FC	CNV region	Log-2-ratio
LmjF.06.0080	ATP-binding cassette protein subfamily G, member 1	0.5524548	CNVR_113	-0.920944
LmjF.06.0090	ATP-binding cassette protein subfamily G, member 2	0.2920734	CNVR_114	-0.9580283
LmjF.14.1100	kinesin K39	-0.2676972	CNVR_110	0.7792663
LmjF.15.1480	cAMP specific phosphodiesterase	0.4970286	CNVR_37	-0.9574448
LmjF.17.0190	receptor-type adenylate cyclase, putative	0.6477952	CNVR_8	-2.746126
LmjF.22.0600	acetyltransferase (GNAT) domain containing protein	-0.63869	CNVR_61	2.383092
LmjF.22.1390	hypothetical protein, conserved	0.3337642	CNVR_71	-0.774964
LmjF.23.0210	ATP-binding cassette protein subfamily C, member 1	1.3873056	CNVR_53	1.786927
LmjF.23.0223	hypothetical protein, conserved	1.1552017	CNVR_53	1.786927
LmjF.23.0225	hypothetical protein, conserved	0.8470822	CNVR_53	1.786927
LmjF.23.0230	hypothetical protein, conserved	0.8052322	CNVR_53	1.786927
LmjF.23.0240	terbinafine resistance locus protein (yip1)	1.486045	CNVR_53	1.786927
LmjF.23.0250	multidrug resistance protein A	1.2144793	CNVR_53	1.786927
LmjF.23.0260	argininosuccinate synthase, putative	1.0082369	CNVR_53	1.786927
LmjF.23.0270	pteridine reductase 1	1.313278	CNVR_53	1.786927
LmjF.23.0280	zinc finger, C3HC4 type	1.539302	CNVR_53	1.786927
LmjF.23.0290	checkpoint protein HUS1	0.451299	CNVR_53	1.786927
LmjF.23.0300	tryptophanyl-tRNA synthetase	0.8019783	CNVR_53	1.786927
LmjF.23.0310	hypothetical protein, conserved	0.6507995	CNVR_53	1.786927
LmjF.23.0320	hypothetical protein, conserved	-0.5984561	CNVR_53	1.786927
LmjF.23.0340	(H+)-ATPase G subunit, putative	-2.2208945	CNVR_53	1.786927
LmjF.23.0220	ATP-binding cassette protein subfamily C, member 2	1.3804016	CNVR_54	-1.347129
LmjF.24.1010	Mnd1 family, putative	2.2562565	CNVR_10	2.126852
LmjF.24.1020	F-box domain/Galactose oxidase containing protein	2.5872305	CNVR_10	2.126852
LmjF.24.1030	dynein light chain	2.581097	CNVR_10	2.126852
LmjF.24.1040	hypothetical protein, unknown function	2.2095486	CNVR_10	2.126852
LmjF.24.1050	SNF2 family protein	2.4974836	CNVR_10	2.126852
LmjF.24.1060	AKAP7 2'5' RNA ligase-like domain containing protein	3.2916332	CNVR_10	2.126852
LmjF.24.1070	hypothetical protein, conserved	2.3595209	CNVR_10	2.126852
LmjF.24.1080	DNAJ domain protein, putative	-0.3976447	CNVR_10	2.126852
LmjF.24.1100	pre-mRNA-splicing factor ATP-dependent RNA helicase	-0.8934357	CNVR_10	2.126852
LmjF.24.1125	hypothetical protein, conserved	-0.2420312	CNVR_14	-0.9032868
LmjF.25.1443	hypothetical protein, conserved	0.264172	CNVR_44	-1.198483
LmjF.27.2340	sucrose hydrolase-like protein	-0.7033248	CNVR_16	-1.525739
LmjF.27.2350	vesicle-associated membrane protein	-0.9150209	CNVR_16	-1.525739
LmjF.27.2360	hypothetical protein, conserved	-0.6365974	CNVR_16	-1.525739
LmjF.27.2370	hypothetical protein, conserved	-0.4573555	CNVR_16	-1.525739

LmjF.27.2380	hypothetical protein, unknown function	-0.6501815	CNVR_16	-1.525739
LmjF.27.2390	TPR-repeat protein, putative	-0.5023864	CNVR_16	-1.525739
LmjF.27.2400	heat shock protein DNAJ, putative	-0.7452128	CNVR_16	-1.525739
LmjF.27.2410	aldo-keto reductase-like protein	-0.3671546	CNVR_17	-0.9651632
LmjF.27.2420	hypothetical protein, conserved	-0.2953249	CNVR_17	-0.9651632
LmjF.27.2430	hypothetical protein, conserved	-0.7725951	CNVR_17	-0.9651632
LmjF.27.2440	3-oxoacyl-ACP reductase, putative	-0.856198	CNVR_19	-0.9582698
LmjF.27.2450	hypothetical protein, conserved	-0.3068082	CNVR_19	-0.9582698
LmjF.27.2460	hypothetical protein, conserved	-0.6411425	CNVR_20	-0.8594242
LmjF.27.2590	dynein heavy chain	0.3560877	CNVR_26	-0.8456079
LmjF.28.1540	hypothetical protein, unknown function	0.3502462	CNVR_141	-1.515351
LmjF.33.2020	hypothetical protein, unknown function	-0.3044996	CNVR_127	-1.244153
LmjF.33.2760	hypothetical protein, conserved	-0.4947443	CNVR_133	-0.7255231
LmjF.33.2955	hypothetical protein	-0.4253012	CNVR_134	-0.8459617
LmjF.34.1520	p25-alpha, putative	-0.3007447	CNVR_1	1.194699
LmjF.34.1530	p25-alpha, putative	0.7736714	CNVR_1	1.194699
LmjF.35.5170	hypothetical protein, conserved	2.7907753	CNVR_85	0.7553396
LmjF.35.5180	hypothetical protein, conserved	1.0829859	CNVR_86	0.641145
LmjF.35.5190	NIMA-related kinase	1.4470836	CNVR_87	0.7353458
LmjF.35.5240	hypothetical protein, unknown function	-0.2359127	CNVR_93	0.7487856
LmjF.35.5290	Lsm12 protein, putative	0.6631023	CNVR_96	0.7330686
LmjF.35.5300	hypothetical protein, conserved	0.8586335	CNVR_96	0.7330686
LmjF.35.5320	SAC3/GANP/Nin1/mts3/eIF-3 p25 family protein, putative	0.4844933	CNVR_97	0.6892603
LmjF.35.5330	isopentenyl-diphosphate delta-isomerase (type II)	0.7756338	CNVR_98	0.6620098
LmjF.35.5350	AAT27.1, amino acid permease	0.9528822	CNVR_99	0.6878376
LmjF.36.1520	NIMA-related protein kinase	-0.7818998	CNVR_46	-0.9723839