Appendix

Table of Contents

Appendix Table S1: Orthologous genes of TbKKTs in the genome of *L. major*: LmKKTs 1

Appendix Figure S1: Chromosome 29: comparison of the results obtained using fold-change *versus* subtraction methods for ChIP-seq analysis. 2

Appendix Figure S2: PCR failed to amplify intergenic regions A and B. 3

Appendix Table S2: List of the primers used to amplify intergenic regions A and B 3

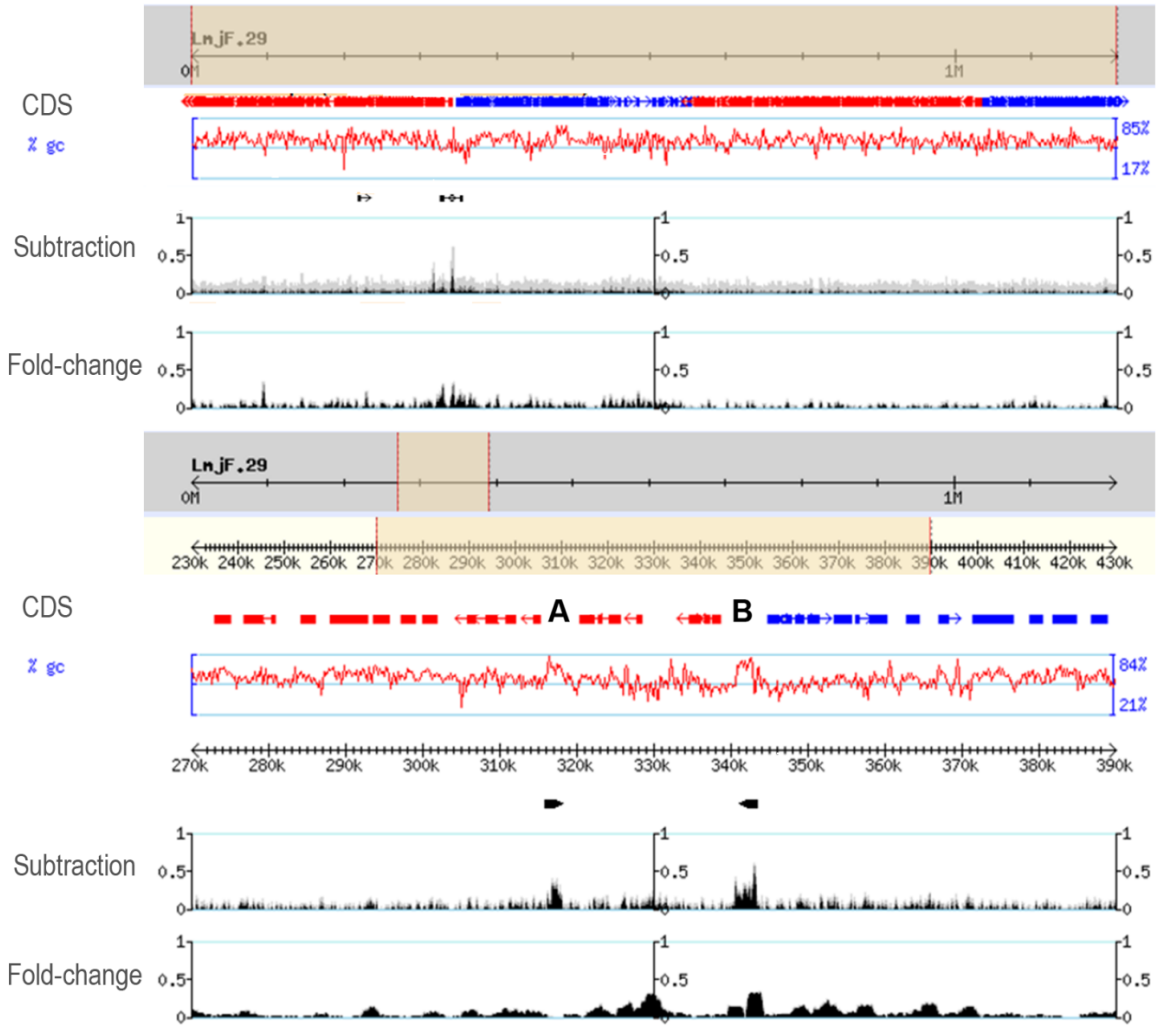
Appendix Figure S3: DNA motifs characterizing the centromeres of *T. brucei* (TbKKT3 binding sites) 4

Appendix TableS2 : Presence and Localization Retroposons in the Leishmania major centromeric sequences 5

Appendix Table S1: Orthologous genes of TbKKTs in the genome of *L. major*: LmKKTs

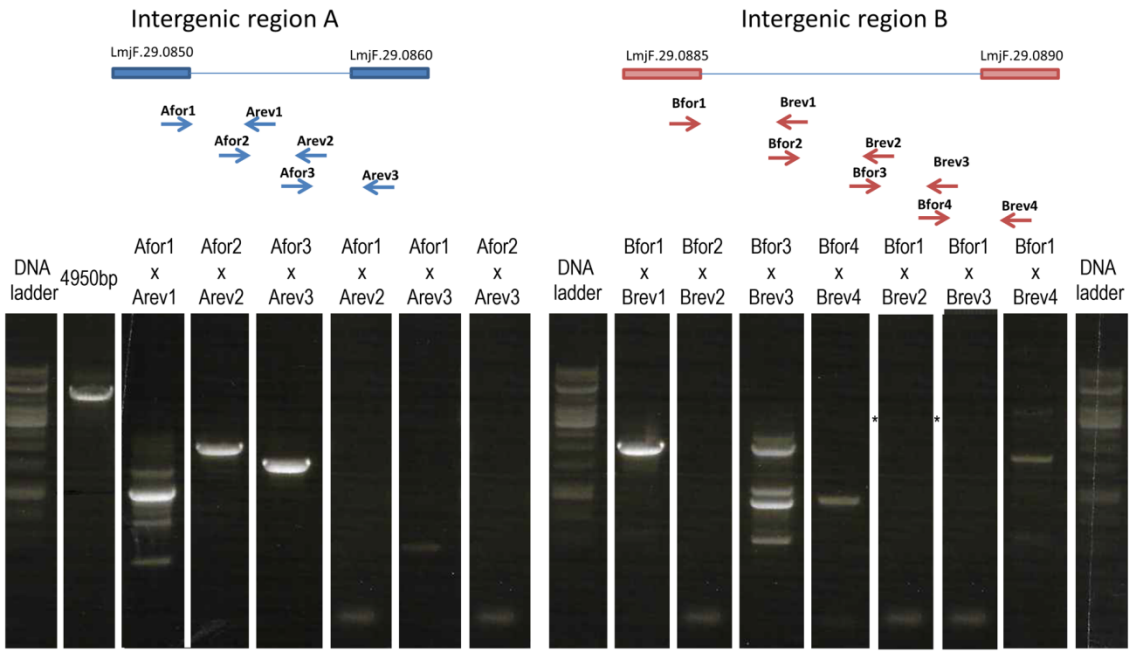
Name	*T.brucei*	*L.major*	Similarity (%)	Identity(%)
KKT1	Tb927.10.6330	LmjF.36.1900	53	36.2
KKT2	Tb927.11.10520	LmjF.36.5350	61.9	46.9
KKT3	Tb927.9.10920	LmjF.35.4050	51	38.5
KKT4	Tb927.8.3680	LmjF.10.0300	43	29.9
KKT5	Tb927.7.4850	LmjF.06.0200	57	36
KKT6	Tb927.6.1210	LmjF.12.0080	67.8	50.5
KKT7	Tb927.11.1030	LmjF.27.0430	36.6	24.4
KKT8	Tb927.4.5110	LmjF.31.2750	46.1	30
KKT9	Tb927.8.1150	LmjF.02.0610	42.8	26.2
KKT10*	Tb927.11.12410	LmjF.09.0410 LmjF.09.0400	74.3 79	65 70
KKT11	Tb927.7.2110	LmjF.22.0120	48.6	25.3
KKT12	Tb927.8.1680	LmjF.24.1400	51	29
KKT13	Tb927.7.4860	LmjF.06.0210	26	17.9
KKT14	Tb927.10.7240	LmjF.36.2800	39.6	27.5
KKT15	Tb927.6.3760	LmjF.30.2520	51.5	36.2
KKT16	Tb927.11.1000	LmjF.27.0400	45.1	24.6
KKT17	Tb927.3.2330	LmjF.25.2220	52.8	36
KKT18	Tb927.9.3800	LmjF.01.0350	43.6	28.7
KKT19*	Tb927.11.12420	LmjF.09.0410 LmjF.09.0400	71.1 74.2	62.7 66.1
KKT20	Tb927.8.4760	LmjF.10.1227	31.8	18.5

* TbKKT19 and TbKKT10 share 82.5% similarity and 79.8% identity



Appendix Figure S1: Chromosome 29: comparison of the results obtained using fold-change *versus* subtraction methods for ChIP-seq analysis.

Chromosome #29 is the only chromosome for which a clear peak was not obtained by ChIP-seq; on the contrary, it gives a very "noisy" dataset, in particular when using the fold-change-method to discriminate noise from signal. A simple subtraction after normalization to separate noise from signal; and indeed the subtraction method allowed us to distinguish a major peak and a minor peak, both being separated by 25 kb and sharing sequence similarities. Top panel: whole chromosome; lower panel: zoom-in view, A: intergenic region LmjF.29:315250..320300 and B: intergenic region LmjF.29:338691..344823.



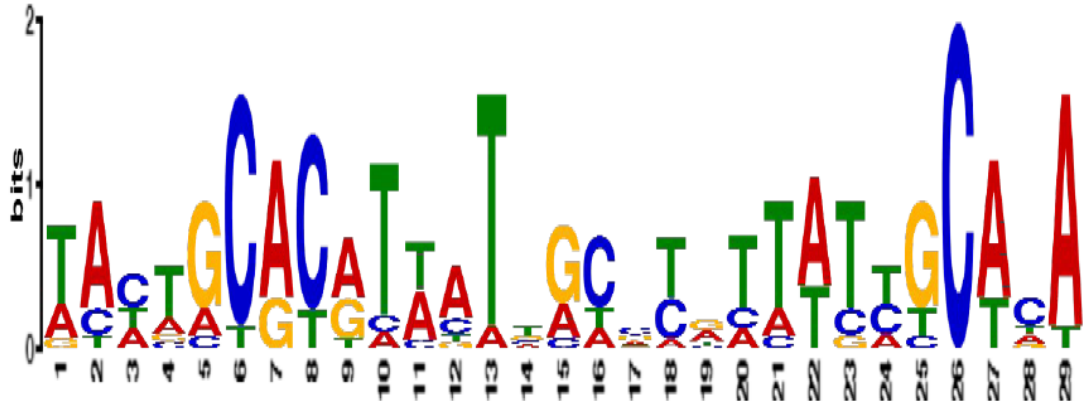
Appendix Figure S2: PCR failed to amplify intergenic regions A and B.

Chr. #29 has the same statistics regarding mapped reads as all other chromosomes, so there is no unfortunate sequencing bias that could explain this. One possible explanation is that the centromeric region of chromosome 29 may has not been properly assembled in the sequence available in TritrypDB. This same explanation had been alleged by Akiyoshi & Gull (Cell 2014) for Chr. 9, 10 and 11 of *T. brucei*. This hypothesis is further supported by the fact that both regions LmjF.29:315250..320300 and LmjF.29:338691..344823 (so-called intergenic region A and B) share sequence similarities. Indeed, a large 2244-bp region is perfectly duplicated, but inverted, in the positions LmjF.29:316100..318343 and LmjF.29:342885..340642; another smaller fragment of 114 bp was found in LmjF.29:20722..208354 (Appendix Figure S1, top panel). PCR primers were designed to map the intergenic regions A and B. The expected product sizes were 2000 bp and 5000 bp depending on the primer pairs. A 4950bp positive control was successfully amplified (Lane 2). Some of the PCR products were successfully amplified (Lane 4 (Afor2xArev2), 5 (Afor3xArev3), 10 (Bfor1xBrev1), 13 (Bfor4xBrev4) and the sequence verified by sequencing. For Bfor1xBrev2, a faint band (**) was visible at the expected size (4000bp) but could not be sequenced. All the other primer pairs failed to give a specific product at the expected size. For the list of primers see Appendix Table S2).

Appendix Table S2: List of the primers used to amplify intergenic regions A and B

	SSR A		SSR B
A-for1	ccgacatcgtgcatgatctcc	B-for1	taattgcagcttcccgcagc
A-rev1	cagcatctttcgtcggcagc	B-rev1	acggcgtccatccttcaagg
A-for2	cccgagcacgaaaaagtcacg	B-for2	tccagacgccaaagctgacg
A-rev2	cgggaagctctctcgatgagg	B-rev2	gagcgaaagcaaccacgtcc
A-for3	ccggtcgagtcatcgtttcg	B-for3	tgggatgccgcggcttg
A-rev3	ccgcacaatctcattgcagcc	B-rev3	gtgtgggcacattcctcagc
		B-for4	tcacttagggcgtcgcttgg
		B-rev4	ccccgagaacttcctgactcc

Appendix Figure S3: DNA motifs characterizing the centromeres of *T. brucei* (TbKKT3 binding sites)



Appendix TableS2 : Presence and Localization Retroposons in the Leishmania major centromeric sequences

Chromosome number	Coordonates of centromeres	Presence of Retroposon	LmSIDER 1	LmSIDER 2	LmDIRE	Coordinates of Retroposons and Remark
>LmjF.01	LmjF.01:268500..268899	No				
>LmjF.02	LmjF.02:264600..269599	Yes	1			LmjF.02: 265759..266339 (+)
>LmjF.03	LmjF.03:247700..253299	Yes		1		LmjF.03: 252497..253066 (+)
>LmjF.04	LmjF.04:126000..131499	No				
>LmjF.05	LmjF.05:361500..366799	No				
>LmjF.06	LmjF.06:120889..129601	No				
>LmjF.07	LmjF.07:209000..213799	Yes	2	1	1	LmjF.07: LmSIDER2 LmjF.07: 209000..209299 (-) and LmSIDER1-LmDIRE (3.7 kb)-LmSIDER1 LmjF.07: 20964 209000+4679 (+)
>LmjF.08	LmjF.08:486100..492999	Yes		2		LmjF.08: 486608..487168 (+) and LmjF.08: 491788..492360 (-)
>LmjF.09	LmjF.09:272500..276899	No				
>LmjF.10	LmjF.10:295800..300699	No				
>LmjF.11	LmjF.11:159000..162599	Yes	1	1		LmSIDER1 LmjF.11: 159131..160049 (+) and LmSIDER2 LmjF.11: 162118..162599 (+)
>LmjF.12	LmjF.12:286300..290999	Yes*			1	1 LmSIDER2 and 1 LmSIDER1 sequence 30 bp upstream the centromeric sequence and one LmSIDER2 40 bp downstream the centromeric sequence
>LmjF.13	LmjF.13:143300..145099	Yes*	1			LmSIDER1: 1117 bp downstream the centromeric sequence
>LmjF.14	LmjF.14:158300..162799	Yes		1		LmjF.14: 162101..1617469 (+)
>LmjF.15	LmjF.15:323300..329099	Yes		1		LmjF.15: 325989..326264 (-)
>LmjF.16	LmjF.16:340100..343899	No				
>LmjF.17	LmjF.17:340300..343799	Yes		3		3 LmSIDER2 copies: 2 upstream the sequence, the second has 89 bp in the centromeric séquence , the third copy is 820 bp down to the centromeric sequence
>LmjF.18	LmjF.18:445000..448999	Yes	3		1	3 LmSIDER1 copies: 2 upstream the sequence, the second has 48 bp in the centromeric séquence ,the third copy is 479 bp downstream to the centromeric sequence LmDIRE: LmjF.18: 445000..448004
>LmjF.19	LmjF.19:625500..631099	Yes		1		LmSIDER2:downstrem the centromerci (163 bp included in the centromerci sequence)
>LmjF.20	LmjF.20:523700..527299	Yes		1		LmjF.20: 526597..527160 (-)
>LmjF.21	LmjF.21:225400..229699	Yes		2		LmjF.21: 225215..225748 (-) and LmjF.21: 229172..229702 (+)
>LmjF.22	LmjF.22:606,000..612,000	Yes		2		LmjF.22: 606565..607003 (+) and LmjF.22: 607884..608208 (-)
>LmjF.23	LmjF.23:550300..553999	Yes*		1		LmSIDER2: 1 bp downstream the centromeric sequence

>LmjF.24	LmjF.24:467300..470099	Yes		1		LmjF.24: 469346..469815 (+)
>LmjF.25	LmjF.25:586700..590099	Yes		2		LmjF.25: 586475..587100 (-) and LmjF.25: 589148..589717 (-)
>LmjF.26	LmjF.26:607800..610199	Yes		1		LmSIDER2: upstream the centromeric sequence with 26 bp included in the centromerci sequence
>LmjF.27	LmjF.27:983200..987599	No				
>LmjF.28	LmjF.28:821000..823899	Yes		1		LmjF.28: 821950..822304 (-)
>LmjF.29	LmjF.29:340400..343399	Yes	1	1	1	LmDIRE: LmjF.29: 341015..342916 (+) and LmSIDER2: LmjF.29: 343228..343800 (+)
>LmjF.30	LmjF.30:229500..232799	No				
>LmjF.31	LmjF.31:778100..781199	Yes		1		LmjF.31: 780368..780848 (-)
>LmjF.32	LmjF.32:1169000..1172399	Yes		2		LmjF.32: 1168802.. 1169260 (+) and LmjF.32: 1170174..1170617 (+)
>LmjF.33	LmjF.33:761900..765099	Yes		1		LmjF.33: 764818..765396 (+)
>LmjF.34	LmjF.34:299200..302699	Yes		1		LmjF.34: 302258..302702 (+)
>LmjF.35	LmjF.35:546700..549899	No				
>LmjF.36	LmjF.36:1112100..1115899	Yes	1			LmjF.36: 1112172 1112412 (+)
		total	9	26	4	