

Supplementary Table 3. Comparative analysis of rhesus and human MSY genes

Stratum	Ancestral genes	Nucleotide position in rhesus MSY assembly			Coding sequence		% divergence (rhesus - human)			Detecting purifying selection			
		Start	End	Strand	ORF length	% of human ORF length	mRNA	Coding	Introns	dN/dS ¹	Fisher's exact test		
											95% C.I.	p-value	
1	<i>SRY</i>	81482	82385	R	611	100	6.60	6.04	n/a	0.55	0.29	0.98	0.06
	<i>RBMY</i>	6271571	6284071	R	695	47	11.32	9.04	12.21	0.54	0.32	0.99	0.07
	<i>RPS4Y1</i>	122092	147116	F	791	100	3.90	3.04	8.74	0.31	0.14	0.70	0.005
	<i>RPS4Y2</i>	6055088	6082902	F	791	100	5.31	4.55	9.58	0.26	0.13	0.48	<0.0001
	<i>HSFY1</i>	8006371	8008261	R	1205	100	7.76	7.79	10.55	0.62	0.42	0.91	0.02
	<i>HSFY2</i>	8384853	8386745	F	1205	100	9.16	8.70	12.11	0.81	0.55	1.19	0.3
	<i>HSFY3</i>	8553282	8555173	R	1205	100	9.29	8.78	12.11	0.77	0.53	1.15	0.2
2	<i>KDM5D</i>	5560125	5599597	R	4619	100	5.09	4.31	9.11	0.26	0.20	0.35	<0.0001
	<i>TSPY1</i>	6307721	6309591	R	566	61	11.27	9.48	15.54	0.77	0.41	1.26	0.3
	<i>TSPY2</i>	9880248	9881875	R	821	89	11.69	9.76	15.67	0.69	0.50	1.25	0.3
	<i>TSPY3</i>	9896624	9898258	R	821	89	11.69	9.76	15.67	0.69	0.50	1.25	0.3
	<i>TSPY4</i>	9913025	9914659	R	821	89	11.69	9.76	15.67	0.69	0.50	1.25	0.3
	<i>TSPY5</i>	9929422	9931056	R	821	89	11.69	9.76	15.67	0.69	0.50	1.25	0.3
3	<i>ZFY</i>	239742	287462	F	2405	100	3.71	2.08	7.63	0.17	0.09	0.30	<0.0001
	<i>DDX3Y</i>	2107890	2124485	F	1979	100	3.52	2.28	9.01	0.19	0.11	0.34	<0.0001
	<i>UTY</i>	2366565	2577315	R	4043	100	4.38	3.49	8.19	0.33	0.25	0.46	<0.0001
	<i>EIF1AY</i>	5853306	5870788	F	434	100	7.85	2.75	8.55	0.10	0.03	0.40	0.0003
	<i>CYorf15A</i>	5408162	5432524	F	395	100	6.74	5.79	9.10	0.33	0.15	0.69	0.005
	<i>CYorf15B</i>	5434583	5448447	F	539	99	6.26	5.74	8.61	0.54	0.27	1.05	0.07
	<i>USP9Y</i>	1904430	2077520	F	7682	100	3.67	3.02	8.22	0.29	0.23	0.37	<0.0001
	<i>TMSB4Y</i>	2640014	2642403	F	134	100	10.22	10.29	9.57	0.47	0.17	1.12	0.1
4	<i>AMELY</i>	659259	667900	R	578	100	4.89	5.00	7.19	1.33	0.61	2.98	0.5
	<i>NLGN4Y</i>	4143178	4465183	F	2510	102	6.90	4.04	9.95	0.09	0.05	0.13	<0.0001
	<i>TBL1Y</i>	1058868	1127299	F	1562	100	6.78	6.14	9.10	0.18	0.12	0.26	<0.0001

Stratum	Ancestral genes	Nucleotide position in rhesus MSY assembly			Coding sequence		% divergence (rhesus - human)			Detecting purifying selection			
		Start	End	Strand	ORF length	% of human ORF length	mRNA	Coding	Introns	Fisher's exact test			
										dN/dS*	95% C.I.	p-value	
5	<i>PRKY</i>	1189854	1297038	F	1076	129	8.28	6.31	9.64	0.28	0.16	0.46	<0.0001
	<i>MXRA5Y</i>	1583565	1646589	F	8405	642	7.17	7.35	8.67	0.99	0.80	1.24	1.0
	Average^{2,3} (ancestral):						6.55	5.32	9.47	0.30	0.27	0.34	<0.0001
	Average³ (ancestral; single-copy in rhesus and human only):						5.88	4.68	8.81	0.26	0.23	0.29	<0.0001
	Added genes												
	<i>DAZ1</i>	9133192	9195151	R	1796	80	9.24	7.53	11.00	0.92	0.53	1.59	0.75
	<i>DAZ2</i>	10630299	10698472	R	1742	78	9.53	11.03	11.27	0.62	0.40	0.96	0.06
	<i>CDY1</i>	8812938	8815149	R	1626	100	10.92	10.22	n/a	0.41	0.31	0.53	<0.0001
	<i>CDY2</i>	8902500	8904710	F	1625	100	10.87	10.16	n/a	0.41	0.32	0.54	<0.0001
	<i>XKRY</i>	7641603	7643120	R	749	212	16.06	14.40	n/a	1.08	0.75	1.57	0.7
	Average² (added):						11.44	10.79	11.13	0.60	0.50	0.72	<0.0001

¹dN/dS = Ratio of # of non-synonymous substitutions per non-synonymous site (dN) to # of synonymous substitutions per synonymous site (dS)

(Calculated using the CODEML program in the PAML package: Yang Z (1998) Mol Biol Evol 15:568-573)

(Fisher's exact test used to detect deviation of dN/dS ratio from neutrality or deviation from dN/dS = 1)

²Averages calculated using only one member of multicopy gene families with >99% identity

³Ancestral average does not include *MXKRA5Y*, which is a pseudogene in human