

Different patterns of Robertsonian fusion pairing in Bovidae and the house mouse: the relationship between chromosome size and nuclear territories

Philipp L. Wesche and Terence J. Robinson

Supplementary Material

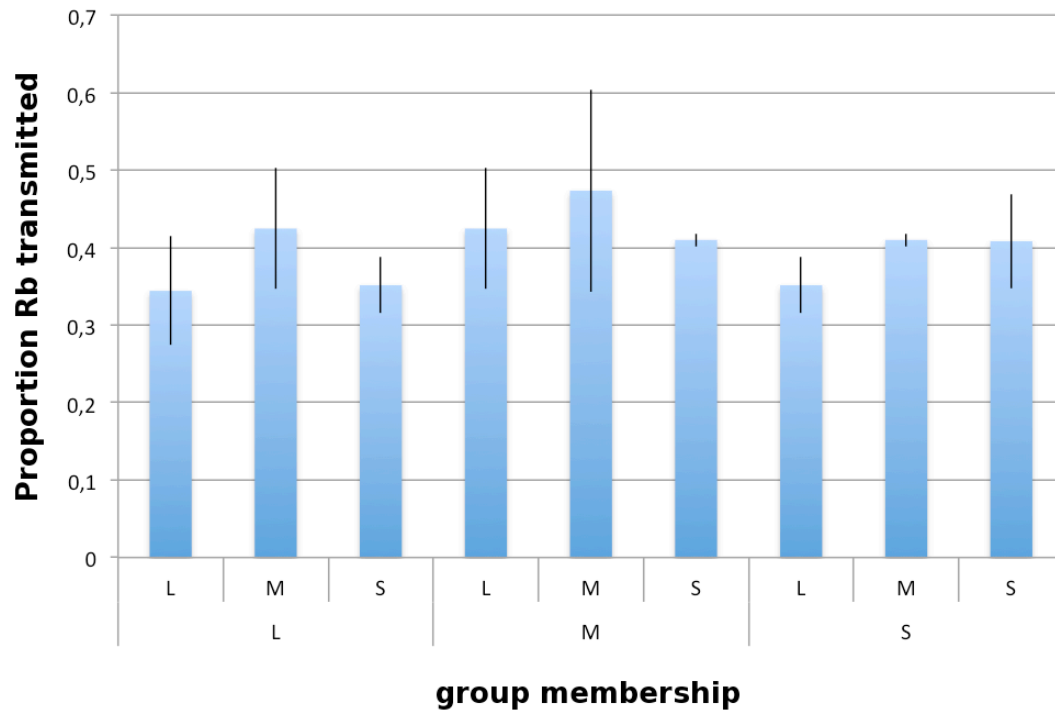


Figure S1 Proportion of metacentric house mouse chromosomes that progress into the oocyte, as a function of group membership of the acrocentric ancestor (and competitor) chromosomes. Groups are composed of the following chromosomes: L – 1-7; M – 8-15; S – 16-19. Error bars indicate standard deviation.

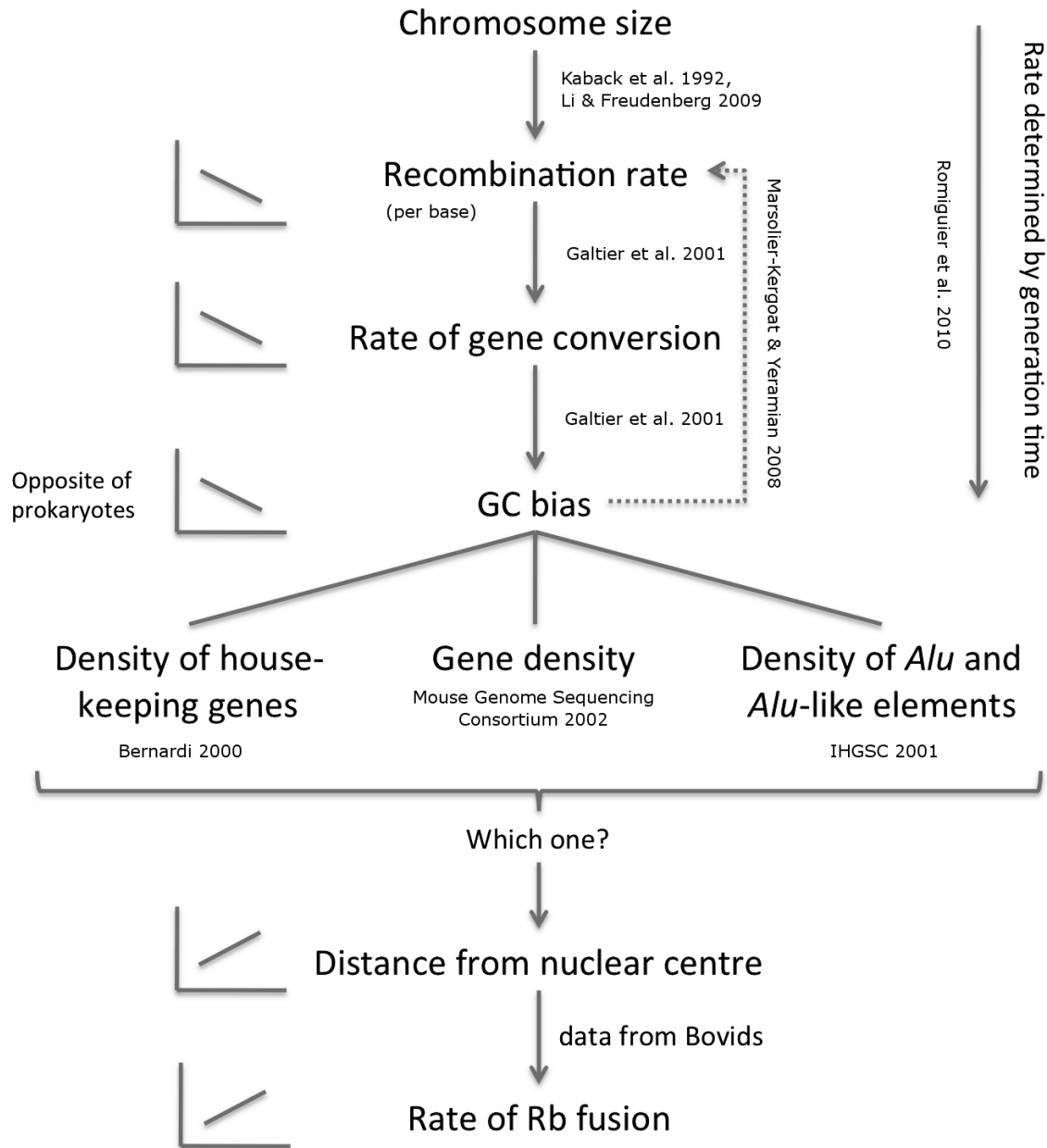


Figure S2 Tentative causality network linking chromosome size and Robertsonian fusion rate for mammals, compiled from the literature. Linkages without arrows indicate correlation rather than causation. The plausibility of this causal chain was only weakly supported by our data.

Chromosome id	Length in bp	Genes of known function	Gene density	Genes w/ high expression ("HKP")*	"HKP" density (% of known genes*)
1	197,195,432	1218	6.18×10^{-6}	218	1.11×10^{-6} (24.33)
2	181,748,087	1715	9.44×10^{-6}	271	1.49×10^{-6} (22.41)
3	159,599,783	1043	6.54×10^{-6}	178	1.12×10^{-6} (22.91)
4	155,630,120	1276	8.20×10^{-6}	203	1.30×10^{-6} (21.31)
5	152,537,259	1261	8.27×10^{-6}	199	1.30×10^{-6} (21.26)
6	149,517,037	1158	7.74×10^{-6}	220	1.47×10^{-6} (26.73)
7	152,524,553	2031	1.33×10^{-5}	292	1.91×10^{-6} (24.25)
8	131,738,871	1064	8.08×10^{-6}	177	1.34×10^{-6} (22.15)
9	124,076,172	1254	1.01×10^{-5}	170	1.37×10^{-6} (20.00)
10	129,993,255	1043	8.02×10^{-6}	138	1.06×10^{-6} (19.61)
11	121,843,856	1569	1.29×10^{-5}	251	2.06×10^{-6} (19.73)
12	121,257,530	694	5.72×10^{-6}	94	7.75×10^{-7} (19.14)
13	120,284,312	852	7.08×10^{-6}	128	1.06×10^{-6} (22.49)
14	125,194,864	769	6.14×10^{-6}	102	8.15×10^{-7} (19.39)
15	103,494,974	793	7.66×10^{-6}	131	1.27×10^{-6} (21.83)
16	98,319,150	662	6.73×10^{-6}	120	1.22×10^{-6} (21.81)
17	95,272,651	1076	1.13×10^{-5}	156	1.63×10^{-6} (20.05)
18	90,772,031	515	5.67×10^{-6}	87	9.58×10^{-7} (21.12)
19	61,342,430	708	1.15×10^{-5}	127	2.07×10^{-6} (27.89)

Table S1 Gene density calculations for the house mouse (* indicates data taken from RUIZ-HERRERA *et al.* 2010).

Dependent variable	Fitted model	Adjusted R ²	p of most recently added variable	95% confidence interval of coefficient
Participations	chrln	0.428	7.09x10 ⁻⁵	3.68x10 ⁻⁸ 9.42x10 ⁻⁸
Participations	chrln+opportunities [†]	0.438	0.232	-8.48x10 ⁻² 2.16x10 ⁻²
Participations	GC	0.00792	0.278	-105 31.5
Participations	genedens	-0.0370	0.974	-3.78x10 ⁵ 3.90x10 ⁵
Participations	chrln+GC	0.4843 [*]	0.0579	-2.19 1.24x10 ²
Participations	chrln+ genedens	0.466	0.101	-5.02x10 ⁴ 5.33x10 ⁵
Participations	GC+genedens	0.112	0.02644; 0.05134	-273 -18.4; -4.35x10 ³ 1.39x10 ⁶
Participations	chrln+GC+genedens	0.464	0.938	-6.70x10 ⁵ 6.21x10 ⁵
Log(ratio)	chrln	0.4844 [*]	1.66x10 ⁻⁵	5.78x10 ⁻⁹ 1.32x10 ⁻⁸
Log(ratio)	GC	0.0128	0.253	-14.7 4.03
Log(ratio)	genedens	-0.0368	0.938	-5.06x10 ⁴ 5.46x10 ⁴
Log(ratio)	GC+genedens	0.147	0.0149; 0.0306	-38.8 -4.59; 1.06x10 ⁴ 1.98x10 ⁵
Log(ratio)	chrln+GC	0.553	0.0318	8.41x10 ⁻¹ 1.70x10 ¹
Log(ratio)	chrln+genedens	0.5361	0.0558	-9.75x10 ² 7.35x10 ⁴
Log(ratio)	chrln+GC+genedens	0.535	0.971	-8.09x10 ⁴ 8.38x10 ⁴

* Extra digit given to clarify that the results are distinct and different.

† As opportunities are used to calculate the proportion of opportunities on which an Rb fusion occurs, opportunities does not appear in any of the models for ratio.

Table S2 Summary of GLM analyses of bovid data. Factors are genedens – gene density in genes/bp, GC – GC content, chrln – chromosome length in bp, opportunities – opportunities to fuse given shape of phylogeny. Confidence intervals are shown with the lower boundary in the line above the upper boundary, p-values are shown prior to correcting for multiple testing.

Dependent variable	Fitted model (shorthand)	Adjusted R ²	p values of most complex or most recent terms, in model order	95% confidence interval of coefficient
Combinations	GC	-0.058	0.993	-179 181
Combinations with outliers	GC	-0.0644	0.861	-164 139
Combinations	GC ²	0.163	0.0403	-3.74x10 ⁻⁴ -9.75x10 ⁻²
Combinations with outliers	GC ²	-0.117	0.731	-1.55x10 ⁻⁴ 2.16x10 ⁻⁴
Combinations	chrln ²	0.148	0.0477	-1.98x10 ⁻¹⁵ -1.17x10 ⁻¹⁷
Combinations with outliers	chrln ²	0.547	0.000356	-2.17x10 ⁻¹⁵ -7.83x10 ⁻¹⁶
Combinations	chrln ² +GC ²	0.492	0.00593; 0.00656	-2.01x10 ⁻¹⁵ -4.22x10 ⁻¹⁶ ; -3.69x10 ⁻⁴ -7.46x10 ⁻³
Combinations	genedens ²	0.403	0.00378	-4.28x10 ⁻¹¹ -1.01x10 ⁻¹¹
Combinations with outliers	genedens ²	-0.03284	0.331	-4.37x10 ⁻¹¹ 1.56x10 ⁻¹¹
Combinations	chrln ² +genedens ²	0.546	0.00490	-3.83x10 ⁻¹¹ -8.60x10 ⁻¹⁰
Combinations	chrln ² +GC ² + genedens ²	0.667	0.0306; 0.0357	-3.89x10 ⁻⁴ -2.37x10 ⁻³ ; -3.09x10 ⁻¹¹ -1.31x10 ⁻¹⁰
Combinations	HKP	-0.031	0.484	-0.718 0.357
Combinations	HKP ²	0.045	0.160	-0.0634 0.348
Combinations with outliers	HKP	0.283	0.0111	-1.21 -0.181
Combinations*	chrln ² +HKP ²	0.098	0.343	-7.32x10 ⁻¹ 4.57x10 ⁻¹
Combinations*	GC ² +HKP ²	0.128	0.284	-9.92x10 ⁻² 3.10x10 ⁻¹
Combinations*	chrln ² + GC ² +HKP ²	0.405	0.755	-1.59x10 ⁻¹ 2.13x10 ⁻¹
Combinations*	chrln ² +genedens ² + HKP ²	0.583	0.112	-3.28x10 ⁻² 2.69x10 ⁻¹
Combinations*	chrln ² +GC ² + genedens ² +HKP ²	0.630	0.418	-1.06x10 ⁻¹ 2.24x10 ⁻¹

* Additional models run to more conclusively rule out an effect of a housekeeping gene density for this dependent variable.

Table S3 Summary of GLM analyses of number of observed Rb combinations in mice. Factors are genedens – gene density, GC – GC content, chrln – chromosome length in bp, HKP – density of housekeeping genes. Confidence intervals are shown with the lower boundary in the line above the upper boundary, p-values are shown prior to correcting for multiple testing. The outliers that were omitted except where stated otherwise were chromosomes 11 and 19.

Dependent variable	Fitted model (shorthand)	Adjusted R ²	p values of most complex or most recent terms, in model order	95% confidence interval of coefficient
Populations	chrln ²	0.440	0.00196	-1.81x10 ⁻¹⁴ -5.03x10 ⁻¹⁵
Populations with outliers	chrln ²	0.645	4.06e-05	-1.46x10 ⁻¹⁴ -6.59x10 ⁻¹⁵
Populations	GC ²	-0.099	0.469	-2.30x10 ⁻⁵ 1.12x10 ⁻⁵
Populations with outliers	GC ²	-0.1077	0.835	-1.33x10 ⁻⁵ 1.09 x10 ⁻⁵
Populations	chrln ² +GC ²	0.474	0.150	-2.10x10 ⁻⁵ 3.60x10 ⁻⁴
Populations with outliers	chrln ² +GC ²	0.6312	6.91x10 ⁻⁵ ; 0.382	-1.50x10 ⁻¹⁴ -6.66x10 ⁻¹⁵ ; -1.01x10 ⁻⁵ 4.11x10 ⁻⁴
Populations	genedens ²	-0.003	0.220	-2.78x10 ⁻¹² 6.96x10 ⁻¹¹
Populations	HKP	0.151	0.069	-7.66 0.325
Populations	HKP ²	0.225	0.140	-0.410 2.62
Populations	chrln ² +HKP	0.5973	0.024	-7.08 -0.580
Populations	chrln ² +GC ² +HKP	0.6409	0.026	-6.86 -0.520

Table S4 Summary of GLM analyses of number of population occurrences of Rb fusions in mice. Factors are genedens – gene density, GC – GC content, chrln – chromosome length in bp, HKP – density of housekeeping genes. Confidence intervals are shown with the lower boundary in the line above the upper boundary, p-values are shown prior to correcting for multiple testing. The outliers, which were omitted except where stated otherwise, were chromosomes 11 and 19.

Dependent variable	Positional explanatory variable	Adjusted R ²	p value of model term	95% confidence interval of coefficient
Combinations	ls	0.509	0.108	-0.817, 0.137
Combinations	fs	0.405	0.149	-1.94, 0.476
Combinations	ess	0.406	0.149	-2.45, 0.600
Combinations	mag	-0.197	0.600	-1.28, 0.885
Combinations	smean	0.6966	0.0497	-1.44, -0.00192
Combinations	mean	0.4601	0.127	-1.61, 0.330
Populations	ls	-0.301	0.804	-5.08, 4.28
Populations	fs	0.131	0.295	-12.3, 5.30
Populations	ess	0.456	0.128	-14.6, 3.03
Populations	mag	0.0515	0.350	-7.82, 3.80
Populations	smean	-0.0845	0.468	-10.3, 6.06
Populations	mean	0.007606	0.385	-10.4, 5.39

Table S5 Summary of GLM analyses of mouse chromosome location data. All analyses were conducted with a single explanatory variable, namely the radial nuclear position (proportion of radius at which observed, standardised such that 0 at centre, 1 on the perimeter), cell types and phases are ls – lymphocyte, S phase; fs – fibroblast, S phase; ess – embryonic stem cell, S phase; and mag – macrophage, G0 phase; smean – mean of S phase data; mean – mean of all types; both means were weighted equally by type.

Supplementary file references

Bernardi, G. (2000). Isochores and the evolutionary genomics of vertebrates. *Gene* **241**, 3-17.

Galtier, N., Piganeau, G., Mouchiroud, D. & Duret, L. (2001). GC-content evolution in mammalian genomes: the biased gene conversion hypothesis. *Genetics* **159**, 907-911.

Kaback, D. B., Guacci, V., Barber, D. & Mahon, J. W. (1992). Chromosome size-dependent control of meiotic recombination. *Science* **256**, 228-232.

International Human Genome Sequencing Consortium (2001). Initial sequencing and analysis of the human genome *Nature* **409**, 860-921.

Li, W. & Freudenberg, J. (2009). Two-parameter characterization of chromosome-scale recombination rate. *Genome Research* **19**, 2300-2307.

Marsolier-Kergoat, M.-C. & Yeramian, E. (2008). GC content and recombination: reassessing the causal effects for the *Saccharomyces cerevisiae* genome. *Genetics* **183**, 31-38.

Mouse Genome Sequencing Consortium (2002). Initial sequencing and comparative analysis of the mouse genome. *Nature* **420**, 520-562.

Romiguier, J., Ranwez, V., Douzery, E. J. P. & Galtier, N. (2010). Contrasting GC-content dynamics across 33 mammalian genomes: relationship with life-history traits and chromosome sizes. *Genome Research* **20**, 1001-1009.