Supplementary materials

Evolution and Functional Classification of Vertebrate Gene Deserts

Ivan Ovcharenko, Gabriela G. Loots, Marcelo A. Nobrega, Ross C. Hardison, Webb Miller and Lisa Stubbs

Figure S1. G+C content versus SINE content correlation in the gene deserts.

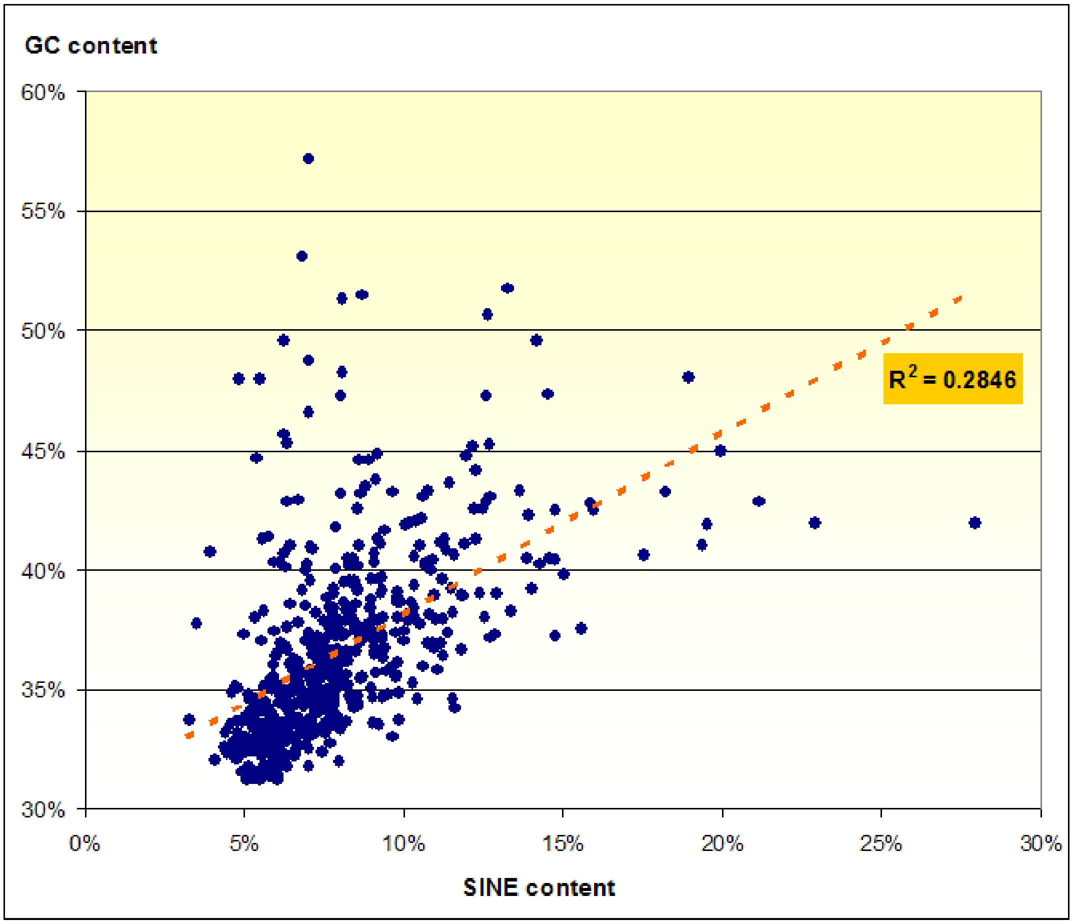

Figure S2. Histogram of the phastCons (phylogenetic hidden Markov model) conservation of the stable (in red) and flexible (in blue) gene deserts. The percentage of total regions falling into a particular density category is displayed on the vertical axis. The horizontal axis depicts the sequence coverage by the .7/100bps phastCons elements.

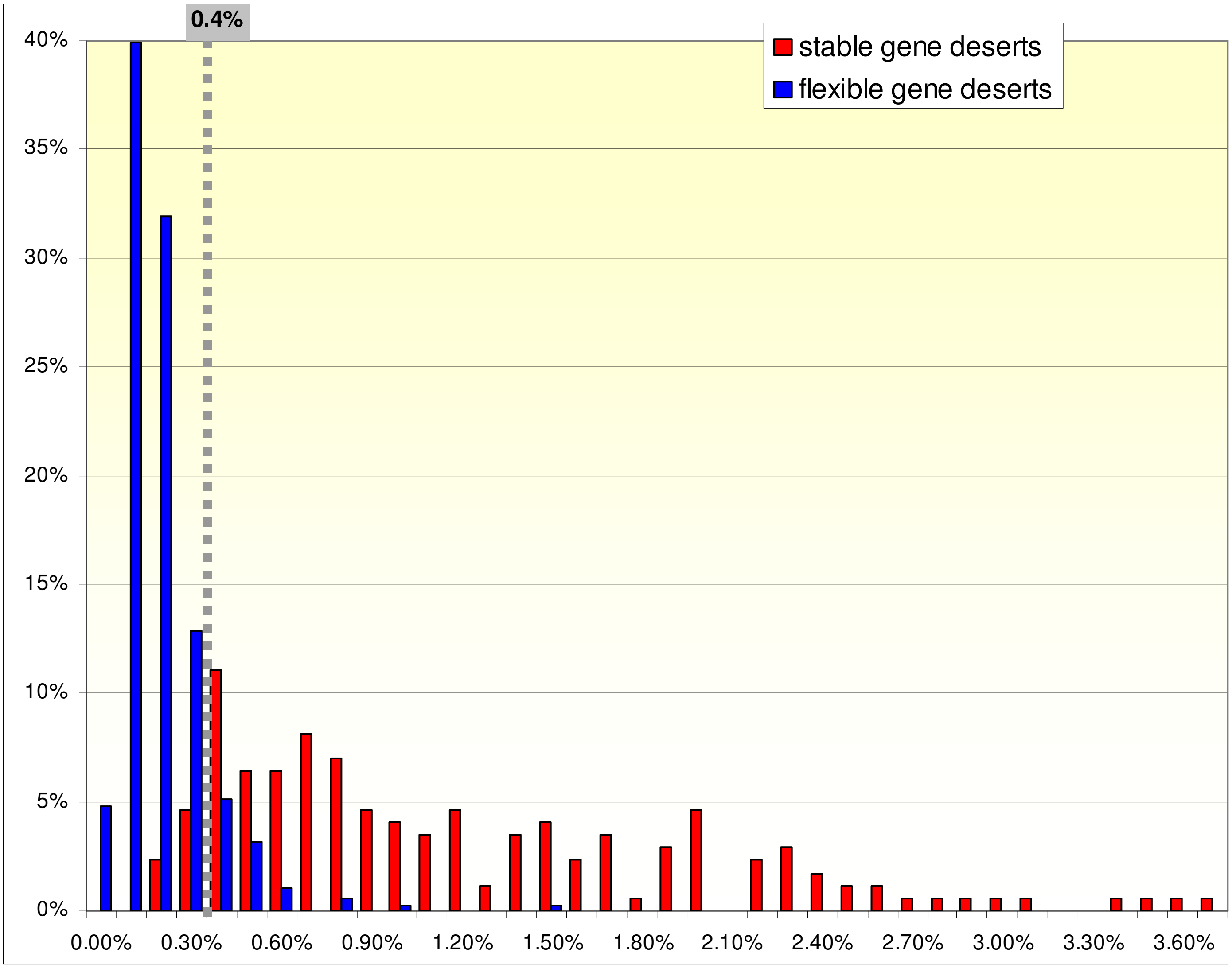

Figure S3. Histogram of the RE density of the stable (in red) and flexible (in blue) gene deserts. The percentage of total regions falling into a particular density category is displayed on the vertical axis. The horizontal axis depicts the RE density in the number of RE elements per 1Mb of sequence.

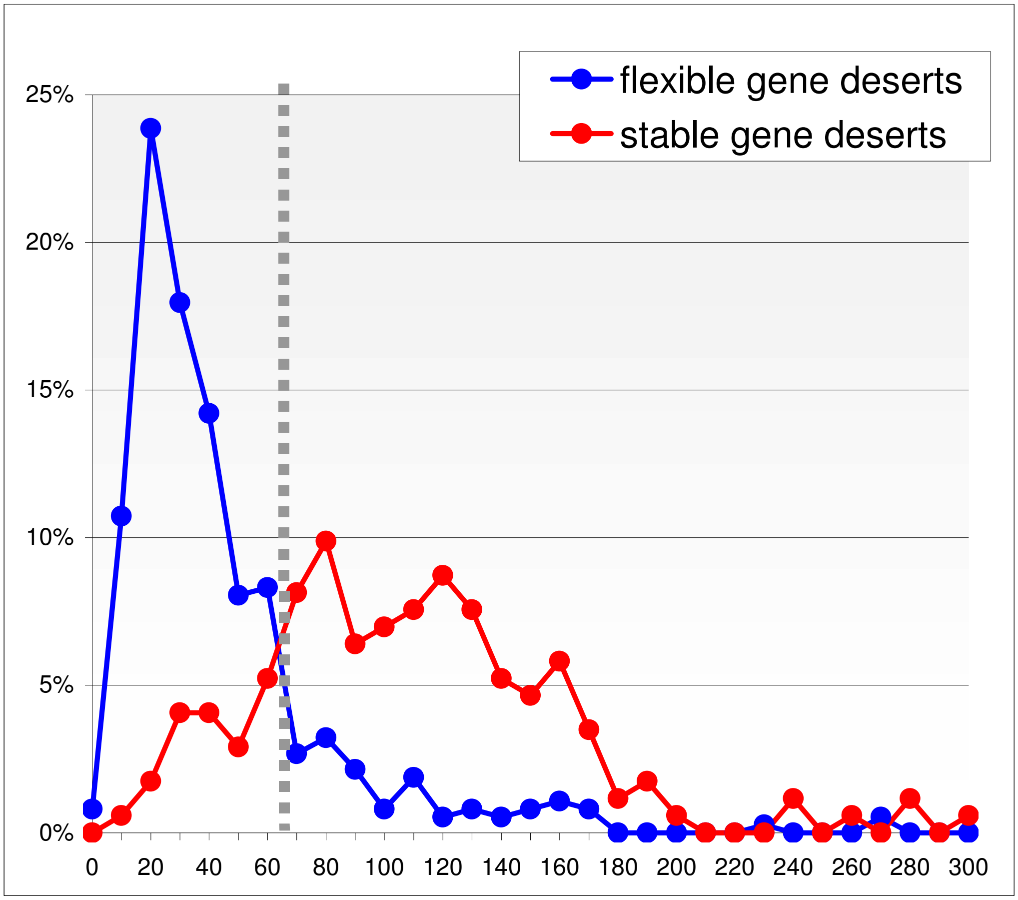

Table S1. SINE and LINE content of different genomic intervals. Regular intergenic intervals were partitioned into two categories by the G+C content: (1) with G+C content less than 40% (as marked by "<40% G+C") and (2) with G+C content greater than 40% (as marked by ">40% G+C").

	LINE content	SINE content
gene deserts	26.39%	7.74%
flexible gene deserts	27.34%	7.34%
stable gene deserts	24.07%	8.72%
average	21.17%	13.63%
regular intergenic "<40% G+C"	20.19%	17.68%
regular intergenic	17.70%	22.45%
regular intergenic ">40% G+C"	11.92%	24.62%
gene-rich	13.97%	24.79%