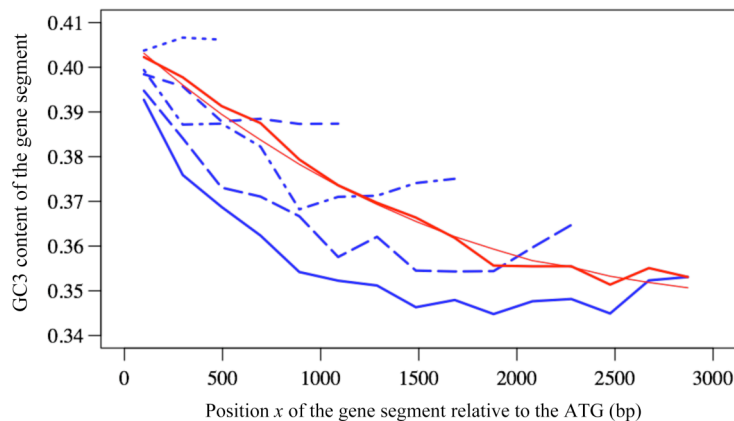


A model for the influence of meiotic conversion tracts on gene GC content

Marie-Claude Marsolier-Kergoat

Institut de Biologie et de Technologies de Saclay, CEA/Saclay, Gif-sur-Yvette, France

A strong correlation between GC content and recombination rate is observed in many eukaryotes, which is thought to be due to conversion events linked to the repair of meiotic double-strand breaks [1]. In several organisms, meiotic gene conversion tract lengths follow a geometric distribution. I have shown that this behavior leads to a simple analytical model for the evolution and the equilibrium state of the GC content of sequences devoid of meiotic double-strand break sites [2]. In the yeast *Saccharomyces cerevisiae*, meiotic double-strand breaks are practically excluded from protein-coding sequences. A good fit was observed between the predictions of the model and the variations of the average GC content of the third codon position (GC3) of *S. cerevisiae* genes. Moreover, recombination parameters that can be extracted by fitting the data to the model coincide with experimentally determined values. These results thus indicate that meiotic recombination plays an important part in determining the fluctuations of GC content in yeast coding sequences. The model also accounted for the different patterns of GC variations observed in the genes of *Candida* species that exhibit a variety of sexual lifestyles, and hence a wide range of meiotic recombination rates. Finally, the variations of the average GC3 content of human and chicken coding sequences could also be fitted by the model. These results suggest the existence of a widespread pattern of GC variation in eukaryotic genes due to meiotic recombination, which would imply the generality of two features of meiotic recombination: its association with GC-biased gene conversion and the quasi-exclusion of meiotic double-strand breaks from coding sequences.



Variations of GC3 content in *S. cerevisiae* protein-coding sequences. The observed GC3 content of 66-codon gene segments is averaged over classes of genes binned by their lengths and plotted as a function of their position x relative to the ATG (blue lines, from the set of the shortest genes _dotted line_ to the set of the longest genes _solid line_). The thick, red line and the thin, red line represent, respectively, the mean observed GC3 content and the mean theoretical GC3 content, as determined by the model, calculated for all genes.

1. Duret, L. and N. Galtier, *Biased gene conversion and the evolution of mammalian genomic landscapes*. *Annu Rev Genomics Hum Genet*, 2009. **10**: p. 285-311.
2. Marsolier-Kergoat, M.C., 2011. *A simple model for the influence of meiotic conversion tracts on GC content*. *PLoS One*. **6**(1): p. e16109.

