

PublisherInfo		
PublisherName	:	BioMed Central
PublisherLocation	:	London
PublisherImprintName	:	BioMed Central

Evolution of genomic GC variation

ArticleInfo		
ArticleID	:	4306
ArticleDOI	:	10.1186/gb-2002-3-10-reports0058
ArticleCitationID	:	reports0058
ArticleSequenceNumber	:	25
ArticleCategory	:	Paper report
ArticleFirstPage	:	1
ArticleLastPage	:	4
ArticleHistory	:	RegistrationDate : 2002-8-28 Received : 2002-8-28 OnlineDate : 2002-11-25
ArticleCopyright	:	BioMed Central Ltd2002
ArticleGrants	:	

Diane P Genereux

Summary

Bioinformatics and classical genetics reveal a positive correlation between recombination rate and GC content at silent sites in coding regions of *Saccharomyces cerevisiae*

Significance and context

GC content is known to vary greatly between different genomic regions in many eukaryotes. Mechanisms previously proposed to explain this variation include selection, mutational bias and biased recombination-associated DNA repair. To determine how such variation might have arisen, Birdsell reports a positive correlation between local recombination rates and 'GC3' - the GC content of third-base codon positions - within coding regions in the budding yeast *Saccharomyces cerevisiae*. Many third codon positions are 'silent', meaning that they do not have a determining role in which amino acid is integrated into a protein. These third codon positions are referred to as 'silent' as they are free from the selection pressures that act to maintain DNA sequences with a direct role in determining amino-acid sequences. Here, Birdsell shows that the documented GC-bias of the system that repairs strand breaks produced during recombination can account for the correlation between GC3 and recombination rate.

Key results

Three approaches were used to look at the correlation of the GC content (as judged by the frequency of GC3 silent sites) with local recombination rate. First, available microarray data were used to determine the correlation between GC3 content and recombination rate for 6,143 individual yeast open reading frames (ORFs). Second, data from 12 previous studies revealed a significant GC-bias in the repair of heteroduplexes in mitotic cells. Third, Birdsell used a comparative approach to determine whether a change in the regional recombination rate can produce a change in GC3 content. Orthologous sequences for the *Fxy* gene from two mouse species, rat and human were aligned, and the ancestral sequence deduced. The gene occurs in a low-recombination region of the genomes of rat, human and one of the mouse species. In the other mouse species, however, a genomic rearrangement has relocated the gene to an area that experiences high recombination rates. Consistent with the predictions of the GC3-recombination hypothesis, the GC content of the mouse gene occurring in the high-recombination region was significantly larger than that of any of the other orthologs, relative to the inferred ancestral sequence.

Birdsell explored four hypotheses to explain this correlation. First, pre-existing high-GC regions might encourage recombination. This hypothesis was rejected, as recombination rate is correlated with frequency of GC3, but not with overall GC content. Second, selection for transcriptional efficiency might drive the correlation, but Birdsell found no correlation between frequency of GC3 or recombination and the codon-adaptation index (CAI). The null expectation is that a given amino acid will be encoded equally frequently by all possible codons. But selection for transcriptional efficiency will tend to increase the frequency of those codons best represented in the tRNA repertoire. The CAI measures the departure of the observed codon frequencies from the null expectation, and can thus serve as an indicator of selection for transcriptional efficiency. Third, mutational bias has been proposed as an explanation of the variation in GC3 frequency throughout the genome. However, this does not account for the correlation between GC3 content and recombination rate, or the low GC content of introns and intergenic locations. The fourth hypothesis, that of G/C-biased gene conversion as a result of GC-biased heteroduplex repair occurring during recombination, avoids many of these pitfalls by providing a causal relationship between local recombination rates and GC3. Under this hypothesis, mismatch repair systems would preferentially insert G or C at sites where strand breakage occurs during meiosis and mitosis.

Conclusions

GC-biased heteroduplex mismatch repair has many advantages over other explanations, but it, too, fails to explain the low GC3 of non-coding regions. To address this problem, Birdsell proposes a "constraint model". Coding regions are typically subject to much stronger selection to preserve sequence than are introns and intergenic regions. Therefore, they tend to be less divergent within the populations. This comparatively low divergence within populations makes them more likely to pair up to form heteroduplexes and undergo recombination. Thus, the model concludes, if biased DNA repair is operating, higher recombination rates will result in an increase in GC3 content in coding regions relative to non-coding regions.

Reporter's comments

Birdsell's study achieves high precision through its correlation of recombination rates and GC3 content within individual ORFs, and the constraint model merits further investigation to establish its validity. As he notes, general correlation of GC3 to recombination may make coding-region GC3 content a useful - if rough - proxy for recombination rates. Moreover, the capacity of biased gene conversion to counteract AT-biased mutation suggests it as a selective pressure favoring sexual recombination.

Table of links

References

1. Birdsell JA: Integrating genomics, bioinformatics and classical genetics to study the effects of recombination on genome evolution. *Mol Biol Evol.* 2002, 19: 1181-1197.