

Selfish drivers

Steven Henikoff and Harmit S. Malik

We are entering the post-genomic era, in which complete catalogues of genetic sequences will, it is widely believed, greatly advance our understanding of biological function. Yet there are large regions of genomes (the complete genetic catalogues of organisms) for which sequencing has hardly begun. Ironically, these are the same regions that were the first to be functionally characterized. As long ago as 1880, cytologists recognized that each chromosome in a cell nucleus has a single region, the centromere, that is acted upon by the mitotic spindle, a fibrous network that pulls sister chromosomes to opposite poles at the beginning of cell division. Thus, even before it was realized that the chromosomes' arms contain the stuff of inheritance, the role of centromeres in chromosome segregation was understood.

Molecular characterization of centromeres has been slow because of technical difficulties, not because of any lack of appreciation of the importance of these structures in cell division. Centromeric DNA in complex genomes consists of tandem repeats, which are difficult to clone, amplify or sequence. Centromeric cores are the most uniformly repetitive regions, consisting of megabase-sized, homogeneous arrays of short tandem-repeat units. These are flanked by more heterogeneous repeats, including the descendants of transposable elements — sequences that can insert themselves randomly throughout genomes. The flanking repeats mediate cohesion between sister chromosomes, and simultaneous dissolution of cohesion on all chromosomes defines the beginning of the anaphase of mitosis, when sister chromosomes begin moving to opposite spindle poles.



This 1880 image shows the bulky chromosomes anchored to the spindle by their centromeres.

This functional distinction between the centromere's core, which is where the spindle microtubules attach, and its flanks, which maintain cohesion, mirrors a biochemical distinction. Half of our genetic material by weight is in the form of octamers composed of four histone proteins (H2A, H2B, H3 and H4), around which DNA is wrapped to form nucleosomes, the basic units of chromatin. Centromeric cores consist of nucleosomes with a variant H3 histone (CenH3) that replaces the usual version. Flanking sequences are packaged into 'heterochromatin', which is distinguished from gene-rich 'euchromatin' by the presence of methyl groups attached to lysine at position 9 of H3. Just how differences in H3 specify chromatin function is of substantial current interest.

Considering that centromeric cores are universally conserved in function, it seems paradoxical that their basic structural components are evolving rapidly. But centromeric repeats indeed undergo rapid evolution, evidently because of frequent homogenization events within centromeric cores. As a result, centromeric repeats differ between closely related species. Rapid evolution also occurs in CenH3s, a surprising observation given that conventional H3 is one of the most highly conserved proteins known. Indeed, this rapid evolution in both plants and animals is adaptive, apparently in response to their rapidly changing cores. Unlike other nucleoprotein machines, such as ribosomes, in which sequence conservation reflects conserved function, centromeric DNA and CenH3s continually diverge in sequence but do not vary in function. Why?

Adaptive evolution of proteins implies an 'arms race', as typified by host-parasite interactions. For example, a host immune system attacking a virus's coat protein selects for mutation of the coat protein to a resistant form, which in turn selects for mutations in the immune system that renew the attack. These adaptations result in greater rates of amino-acid replacement than would be expected in the absence of selection. In the same way, high rates of CenH3 evolution can be understood if centromeric repeats are thought of as selfish elements that constantly compete for survival, and CenH3s as host proteins that adapt in response. However, centromeres are essential for mitosis: even a single loss is highly deleterious, so that competition cannot benefit centromeres unless it involves a process by which chromosome loss does not compromise fitness.

Female meiosis is such a process. Here, the duplicated maternal and paternal centromeres separate from each other, and only one of the four meiotic products is retained

Centromeres

Centromeric DNA repeats can be thought of as selfish elements that constantly compete to beat the odds and make it into the egg at meiosis.

in the egg; the other three are degraded. There is thus the opportunity for darwinian competition between genetic variants of centromeres to reach the preferred pole and end up in the oocyte (which will form the egg). Once this one-in-four bottleneck is present, genes that reinforce it will favour their own inclusion in the oocyte if they are linked to successful centromeres.

The consequences of centromere competition would be profound. Even a slight advantage at each female meiosis will lead to rapid fixation of the winning centromere. Expanding centromeric repeats will compete during the race to reach the oocyte (known as meiotic drive). However, this has deleterious consequences. For example, misalignment of centromeres during male meiosis is thought to trigger checkpoints that interfere with spermatogenesis and cause sterility.

Driving centromeres keep winning the coin flip, and eventually the host genome will try to even the odds. CenH3s are the host factors that are in the best position to resist drive between competing centromeres. A new CenH3 allele with altered DNA-binding preference that restores meiotic parity would alleviate the deleterious effect, and thereby be driven to fixation itself. The result is an irreversible process of centromere divergence, in which both the DNA and protein components differ from their ancestors.

In this view, the sterility of interspecies hybrids that so perplexed Darwin results from a suboptimal combination of CenH3 (or another binding protein) and DNA, which would cause termination of meiosis. Darwinian competition between opposing centromeres provides a general molecular mechanism for centromere evolution that inevitably results in sterility defects in hybrids, thus accounting for the origin of species. ■

Steven Henikoff and Harmit S. Malik are at the Fred Hutchinson Cancer Research Center and the Howard Hughes Medical Institute, 1100 Fairview Avenue North, Seattle, Washington 98109, USA.

FURTHER READING

Henikoff, S., Ahmad, K. & Malik, H. S. *Science* **293**, 1098–1102 (2001).

Pardo-Manuel de Villena, F. & Sapienza, C. *Genetics* **159**, 1179–1189 (2001).

Schueler, M. G., Higgins, A. W., Rudd, M. K., Gustashaw, K. & Willard, H. F. *Science* **294**, 109–115 (2001).

Zwick, M. E., Salstrom, J. L. & Langley, C. H. *Genetics* **152**, 1605–1614 (1999).