Journal of Biology

BioMed Central

Research news

# Making sense of centromeres

Pete Moore

Published: 31 August 2004

*Journal of Biology* 2004, **3:**16

The electronic version of this article is the complete one and can be found online at http://jbiol.com/content/3/4/16



**Comparative analysis of the proteins that bind exclusively at the centromere provides evidence of an evolutionary battle that may make sense of sex.**

At the pinched waist of each eukaryotic chromosome is a region that is both elusive and enigmatic. Despite considerable effort, and multiple announcements of completed genome sequences, this zone stubbornly refuses to reveal its complete sequence, and what little we know of it at first sight runs counter to standard theories of evolution. The region in question is, of course, the centromere.

Back in the 1880s, scientists worked out that centromeres played a critical part in helping cells get their fair share of chromosomes during cell division, and we now know that it is at these sites that spindle microtubules attach. "It's not just another intriguing organelle that we would like to understand; it is central to eukaryotic biology," claims Steven Henikoff, researcher in the Basic Sciences Division of the Fred Hutchinson Cancer Research Center, Seattle, and senior author of the study of centromere protein evolution published in *Journal of Biology* [1] (see the 'The bottom line' box for a summary of the work). In prokaryotes, chromosome segregation at division occurs simultaneously with DNA replication, whereas in eukaryotes the two processes occur at different points in a complex cell cycle. And while the rest of the chromosome's DNA is packaged away and consequently 'silenced' during mitosis, the centromere alone remains active in directing chromosomal movement.

**The bottom line**

- Centromeric DNA is highly repetitive and is therefore difficult to sequence. It is also highly variable between species, which is surprising for a region with an essential function.
- Comparative analysis of centromere-specific DNA-binding proteins, such as CenH3 and CENP-C, across species can provide insights into the evolution of centromeres.
- Although CenH3 is adaptively evolving in *Arabidopsis* and *Drosophila*, there is no evidence of this in grasses and mammals. Instead, there is strong evidence of adaptive evolution in distinct regions of the centromere-binding protein CENP-C in plants and mammals.
- In contrast, yeast CENP-C is under negative selection, perhaps reflecting the simpler organization and lower inter-species variability of the yeast centromeric DNA to which it binds.
- 'Centromere drive', or the unequal transmission of competing centromeres in female meiosis, may account for the rapid evolution of complex centromeres in plants and animals. CENP-C, and some CenH3 proteins, may undergo positive selection to suppress this centromere drive. The absence of this drive process in yeast accounts for its centromeric stability.

**Background**

- **Centromeres**, the DNA sequences that bind proteins of the kinetochore, and hence spindle microtubules, at cell division, have a highly conserved function throughout eukaryotes. The centromeres of animals and seed plants are typically composed of repetitive **satellite** sequences. But paradoxically, despite their conserved function, the sequences in these satellite sequences evolve rapidly.
- While the majority of the chromosomal DNA is wrapped around octamers built of four **histones**, H2A, H2B, H3 and H4, centromeres have a unique version of H3 commonly referred to as **CenH3** (or CENP-A).
- The pressures under which a protein is evolving can be assessed by measuring the rate of mutations that change individual bases within a sequence but still leave the codon specifying the same amino acid (**synonymous substitutions**), and comparing this with the rate at which mutations lead to new amino acids being substituted in the protein (**nonsynonymous substitutions**).
- Earlier studies showed that while the histone H3 in the main bulk of the chromosome is highly conserved, the centromeric H3 proteins in *Drosophila* and *Arabidopsis* are adaptively evolving. The rapidly evolving DNA and adaptively evolving CenH3 proteins seemed to provide evidence of evolutionary conflict between '**centromere drive**' - the competition for transmission among the various centromeres involved in female meiosis - and the need for centromere parity for full fertility in male meiosis.
- Most plants, animals and yeasts also employ a large DNA-binding protein at the centromere named **CENP-C**, which is characterized by a single 24-amino-acid motif - the CENPC motif.

"The centromere is unique in eukaryotic biology - there is nothing else like it - and when we look at its evolution we find that there is nothing like it either," says Henikoff.

While the exact detail of the DNA sequence at the **centromeres** (see the 'Background' box) is unknown, it is clear that the DNA is highly variable at this region both in sequence and in amount - so much so that centromere-specific DNA in the human Y chromosome varies in size by one order of magnitude between people, a feature that is also sometimes seen in other chromosomes. "The paradox is that normal expectations of evolutionary biology say that a region with such a critical and highly conserved function should have a stable sequence," says Kevin Sullivan, in the Department of Cell Biology at The Scripps Research Institute, who has been working on the structure and function of centromere proteins for 15 years. One would expect the DNA sequence to be passed on almost unchanged from individual to individual and even from species to species. But in reality, flies, yeast, plants and mammals have highly individualistic versions of centromere DNA.

## Meiotic solution

To explain this, Henikoff and colleague Harmit Malik proposed a radical theory [2]. Maybe, they suggested, there is a Darwinian competition going on during female meiosis, the process that yields the eggs. A normal cell contains chromosomes in pairs, with one member of that pair coming from the male parent, the other from the female. In female meiosis the two chromosomes first duplicate to make four chromatids, but then three chromatids are effectively thrown away (as 'polar bodies' in mammals, or non-functional 'megaspores' in plants), while lucky number four becomes packaged ready for use in sexual reproduction. It has always been a puzzle that sexual reproduction creates a scenario in which a parent throws its genes into a new individual with only a 50:50 chance that that individual will pass them on. But what if one of the parents built a centromere that made it more likely that its chromatid would win out in meiotic selection? It would now have a 100% chance of launching its genes into the future. If this were to occur, we would expect to see a genetic 'arms race', with individuals within a species competing to create ever more effective centromeres. You'd see this in rapidly evolving centromeric DNA - the very observation that triggered this line of thought.

The idea behind this **centromere drive** model is that each of the four copies of a chromosome goes to a distinct area of the cell and the spindles might be induced to favor pulling one towards the zone that becomes the egg. "It makes a lot of cell biological sense, the idea that the cellular geometry of the spindles in female meiosis influences the fate of the products - it's a great theory," comments Sullivan. This solution to the paradox creates its own problem, however. Male meiosis produces four sets of chromosomes, but in this case all are used. Any imbalance in the types of centromere could cause problems with division. The hypothesis

therefore conjectures that centromeric proteins would need to evolve rapidly to counteract this potential problem.

## Questioning histones

Excited by the implications of this research, Henikoff's colleagues Paul Talbert and Terri Bryson began searching for evidence. Clearly there was no point in trying to squeeze more data from the DNA of the centromeres: major sequencing projects had already drawn a blank there. The alternative approach, however, was to examine the centromeric proteins. Within the centromere, the usual **histone** H3 is replaced by a variant, centromeric H3 (**CenH3**). The obvious starting point was to ask whether there were signs that CenH3 was actively adapting, and to check this out in many different species.

Evolutionary molecular biologists like Caro-Beth Stewart of the University of Albany have shown that you can say whether a protein has evolved adaptively, neutrally, or negatively by measuring the rate of **synonymous** and **nonsynonymous substitutions** in the sequence between species [3]. She likens the task to monitoring speeding cars on a freeway. "After looking at the traffic for long enough you can spot the general regular speed limit, but then among the vehicles there will be a car that is going faster. If this is just for a short burst then it will make no overall difference to that car's progress, but if it consistently speeds then it will be statistically different from the rest and the traffic cops are very likely to spot it," she explains.

By the time that Henikoff's team started the current work [1], they already knew that CenH3 was evolving adaptively in *Drosophila* and *Arabidopsis* [4,5]. They then compared CenH3s in mice and rats. Contrary to expectation, this comparison showed negative selection: in these species, CenH3s were being actively conserved. The same was the case for the Chinese hamster, chimpanzee and human. Switching to plants, they came to the same conclusion: in maize and sugarcane, CenH3 showed overall negative selection.

## Rescued by CENP-C

At this point lesser mortals might have turned tail and torn up their hypothesis. But Henikoff's team

### Behind the scenes

*Journal of Biology* asked Paul Talbert about the inspiration and outlook for his work on centromere evolution.

**What motivated this work, and how long did it take?**

In previous work, our lab discovered adaptive evolution in the CenH3 genes of *Drosophila* and *Arabidopsis*. To explain this unusual finding a model of centromere evolution was developed in which centromeres compete in female meiosis for preferential transmission - a type of meiotic drive we call 'centromere drive'. CenH3s were hypothesized to evolve adaptively to suppress this process. The model predicts that other kinetochore proteins might be evolving adaptively to suppress the meiotic drive of centromeres. When we looked at the *Arabidopsis Cenpc* gene, we were struck by the lack of conservation with the published maize *Cenpc* genes, and wondered if this rapid evolution might be adaptive, as with CenH3. We decided that a more thorough comparison of plant and animal *Cenpc* and *CenH3* genes was warranted. It took us about two years from the initial observation to the completed manuscript.

**What were your initial reactions to your findings?**

Of course we were pleased to find adaptive evolution in CENP-C in accordance with the expectation of the centromere-drive model. The occurrence of positive selection in plants and animals but not yeasts was a satisfying confirmation of the predictions of the model.

**How have the results been perceived by others?**

The centromere-drive model has received some support, but has also generated some controversy. Although the model has always predicted that multiple proteins might act to suppress centromere drive, we expect that the demonstration that CENP-C is more consistently under positive selection than CenH3 should help persuade critics that genetic conflict at animal and plant centromeres is widespread. To our knowledge, no one has proposed an alternative to the centromere-drive model that explains recurrent positive selection in essential conserved kinetochore proteins.

**What are the next steps, and what does the future hold?**

We need to do some more direct tests of the centromere-drive model. Already there is evidence that Robertsonian translocations (centromere fusions) are subject to meiotic drive in women and cause reduced fertility in men. One prediction of the centromere-drive model is that the rapid divergence of centromeres and kinetochore proteins can be a mechanism of post-zygotic reproductive isolation in speciation of animals and seed plants. Identification of 'speciation' genes should help in determining whether some or all of these genes suppress centromere drive.

remained convinced of their basic premise and instead started to look at other DNA-binding proteins in the centromere. One obvious target was **CENP-C**, a poorly conserved centromere protein that is less conserved over its entire length than CenH3 but contains a 24 amino-acid motif known as the CENPC motif. Human CENP-C had previously been shown to bind centromeric DNA and to be needed for centromeres to function successfully. The CENP-C protein is bigger than CenH3 and quite possibly makes more contacts with DNA. Although its function is unclear, it does co-localize with CenH3 within the active heart of centromeres. Once again the team started measuring the rate of protein evolution.

Starting with rodents, they found clear evidence that while the CENPC motif was under negative selection, most of the amino-terminal portion of the protein was under positive selection. Similar findings came from humans and chimpanzees, and in the mustard and grass families CENP-C was much more prone to positive selection than was CenH3. "What is particularly exciting is that when we look at organisms where the CenH3 is not adaptively evolving, CENP-C is," says Henikoff. The adaptive evolution of CenH3 in *Drosophila* is quite possibly due to the fact that it does not have CENP-C. At the same time, Sullivan wonders whether you could now identify which of the 30 or so known centromeric proteins are in contact with the DNA by looking for regions of adaptive evolution alone.

If the idea is that at least one key centromere protein must evolve adaptively with centromeric DNA this could explain not only how and why centromeric DNA evolves so rapidly, but also why hybrids between species are generally infertile. It would go a long way to explaining not only the role of sex, but also the origin of distinct species (see the 'Behind the scenes' box for more of the rationale for the work). Sullivan would love to see the theory tested by taking *Drosophila simulans* and transforming it with the corresponding CenH3 from *D. melanogaster* and seeing if the species barrier then breaks down.

## Is this the answer?

The idea takes a moment to think through. Peter Langridge, CEO and Director of the Australian Centre for Plant Functional Genomes at the University of Adelaide, has views symptomatic of many. Initially less than convinced, he is warming to the idea. He points out that the fertility of a higher eukaryote is largely determined by the ability of the eggs to become fertilized. "Most plants can tolerate 90% male sterility without a real drop in fertility," he notes. "Having read the paper it seems obvious that such a meiotic drive process would exist, and the observation and arguments that the binding of the centromeric proteins controls this process seem logical. The model provides a good explanation for the variation in centromeric sequence and this was an issue that had puzzled me." That said, Langridge found himself thinking about other possible explanations. Could the centromeric sequences and proteins be influencing some other process, such as recombination? Suppressing recombination at the centromere might present a selective advantage to protect some genes from recombination. "If increasing variability in the sequences of the DNA and proteins at the centromere destabilized protein-centromere interactions and thus lowered the recombination rate, could one get the same results as Henikoff's team?" he asks. "This is the great thing about this paper: it got me thinking in all sorts of weird directions," says Langridge. As with all scientific theories, these models require further testing and rigorous scrutiny from the scientific community. Whatever the outcome of that research, Henikoff and colleagues' new results have provided some much-needed insight into the inscrutable centromere.

## References

1. Talbert PB, Bryson TD, Henikoff S: **Adaptive evolution of centromere proteins in plants and animals.** *J Biol* 2004, **3:**18.
2. Henikoff S, Ahmad K, Malik HS: **The centromere paradox: stable inheritance with rapidly evolving DNA.** *Science* 2001, **293:**1098-1102.
3. Messier W, Stewart CB: **Episodic adaptive evolution of primate lysozymes.** *Nature* 1997, **385:**151-154.
4. Malik HS, Henikoff S: **Adaptive evolution of Cid, a centromere-specific histone in *Drosophila*.** *Genetics* 2001, **157:**1293-1298.
5. Talbert PB, Masuelli R, Tyagi AP, Comai L, Henikoff S: **Centromeric localization and adaptive evolution of an *Arabidopsis* histone H3 variant.** *Plant Cell* 2002, **14:**1053-1066.

*Pete Moore is a science writer based in Gloucestershire, UK. E-mail: moorep@mja-uk.org*