

Neocentromeres and human artificial chromosomes: An unnatural act

Huntington F. Willard*

Center for Human Genetics and Research Institute, University Hospitals of Cleveland and Department of Genetics, Case Western Reserve University School of Medicine, Cleveland, OH 44106

In the beginning—even before the Human Genome Project—there were chromosomes. Notwithstanding their primary role in determining the transmission of genes and alleles during mitosis and meiosis and their clear relevance to major forms of human pathology such as cancer and birth defects, chromosomes in complex eukaryotic organisms remain relatively poorly understood. The molecular basis for functional chromosomal elements such as telomeres, origins of DNA replication and the centromere are under investigation in a host of organisms ranging from the unicellular yeast *Saccharomyces cerevisiae* to humans. This task has been both complicated and motivated in part by the fascinating observation that the cellular, genomic, and molecular mechanisms of some of these functional elements are remarkably variable from organism to organism (1). Nowhere is this more evident than for the centromere, the chromosomal locus responsible for attachment of the chromosome to the mitotic and meiotic spindle apparatus and thus for segregation of chromosomes to daughter cells during cell division. Centromeres can be small, unitary elements (as in *S. cerevisiae*), localized repetitive elements stretching on for millions of base pairs (as in many plants and animals, including humans), or dispersed elements distributed along the length of the chromosome (as in the nematode *Caenorhabditis elegans*) (1). And most recently, as illustrated by the paper of Saffery *et al.* (2) in this issue of PNAS, it has become apparent that, even in organisms characterized by a localized centromere, other chromosomal sequences that are located far away from the normal centromere and that normally have nothing to do with centromere function can fulfill the role of a centromere. These so-called “neocentromeres,” once formed, appear to behave functionally in all respects as if they are *bona fide* centromeres. A lingering debate in the field has been whether the existence of neocentromeres, especially in humans, implies the existence of many different sequences equally capable of forming a functional centromere, or whether it reflects a strong selection for sequences—no matter

how inefficient—capable of conferring mitotic stability (3, 4). Saffery *et al.* (2) have addressed this question by attempting to generate human artificial chromosomes with neocentromeres. Neocentromeres were first described in maize as subtelomeric heterochromatic “knobs” that could be transformed into functioning centromeres capable of mobilization on the spindle (5). However, unlike maize neocentromeres, which are comprised primarily of tandemly repeated satellite DNA with homology to maize centromere DNA (6), human neocentromeric DNA is not repetitive (7) and appears not to share obvious homology with the 171-bp repeated DNA family, alpha satellite, that characterizes normal human centromeres (3). Furthermore, several dozen neocentromeres have been documented in the human genome, from many different chromosomal regions (8–10). Together, these observations have been used to argue in favor of the model that the basis for centromere activity in human chromosomes is epigenetic, rather than determined solely by primary sequence (3, 4, 8, 11). The basis for such an epigenetic mark is not clear and may reflect a functional determinant created by secondary structure (4, 12, 13) or some other characteristic of centromere behavior such as late DNA replication (3, 14).

Although neocentromeres are the most striking example of the functional plasticity of centromeres in higher organisms, they are by no means the only example. Human centromeres, for example, can adopt a wide range of apparently equally functional configurations (4). A typical centromere appears to consist of several million base pairs of alpha satellite DNA, although there is substantial variation both in the amount of alpha satellite as well as its primary sequence from chromosome to chromosome. Dicentric

chromosomes—generally presumed to be unstable because of their susceptibility to chromosome breakage (15)—are in fact frequently stabilized by centromere inactivation, the epigenetic modification of one of the two centromeres to render it incapable of forming a kinetochore and binding to spindle microtubules (4, 16). Although the molecular basis of

centromere inactivation is unknown, it is accompanied by absence of centromere and kinetochore proteins that are normally associated with functionally active

Neocentromeres appear to be indistinguishable from normal, alpha satellite-based centromeres in terms of the chromatin and kinetochore complex that provides functional activity.

centromeres (16). Among these is centromere protein A (CENP-A), a centromere-specific histone H3 variant that normally associates with alpha satellite DNA and marks centromeric chromatin (17). CENP-A homologues have been described at yeast, worm, and fly centromeres, indicating an evolutionarily highly conserved chromatin component at functional centromeres despite the wild variation in centromeric DNA content among eukaryotic chromosomes (1, 18). It is thus particularly notable that neocentromeres are associated with all of the known, functionally implicated centromere proteins (9, 10, 19). Furthermore, in the chromatin of the one neocentromere examined to date, CENP-A is associated over a nearly half million base pair region of genomic DNA that is, at least at the level of primary sequence, not detectably different from that of surrounding non-CENP-A-binding DNA (20). Thus, neocentromeres appear to be indistinguishable from normal, alpha satellite-based centromeres in terms of the chromatin and kinetochore complex that provides functional activity.

See companion article on page 5705.

*Address reprint requests to: Research Institute, University Hospitals of Cleveland, Lakeside 1400, 11100 Euclid Avenue, Cleveland, OH 44106. E-mail: Willard@UHRI.org.

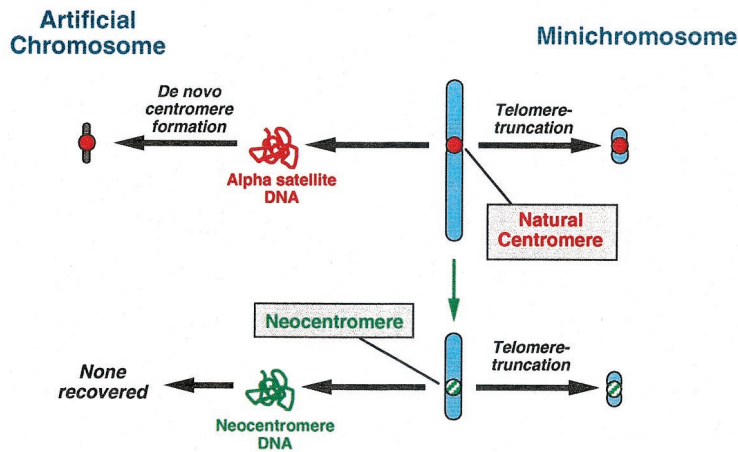


Fig. 1. Formation of human artificial chromosomes and engineered minichromosomes containing natural centromeres (Top) or neocentromeres (Bottom). Minichromosomes, generated by telomere-associated truncation (27, 28), can be formed from either centromeric or neocentromeric chromatin that already possesses the epigenetic modifications necessary for functional centromere competence (18). In contrast, human artificial chromosomes can be generated *de novo* from isolated centromeric DNA (alpha satellite; refs. 23, 24, and 30), but not from isolated neocentromeric DNA (2). Thus, at the level of DNA, alpha satellite and neocentromeric DNA are not equivalent at centromere formation.

It is this premise that Saffery *et al.* (2) set out to explore in the context of human artificial chromosomes. The concept of artificial chromosomes was introduced nearly twenty years ago in *S. cerevisiae* (21), referring to construction of a fully functional chromosome from its component parts. Since then, artificial chromosomes have been used to explore many features of chromosome biology, especially in *S. cerevisiae* and *Schizosaccharomyces pombe* (22). More recently, an artificial chromosome system has been developed in human cells, based on the ability of alpha satellite DNA to provide *de novo* centromere function (23, 24). Indeed, the observation that alpha satellite DNA can seed formation of *de novo* centromeres in artificial chromosomes while other sequences cannot (25) provides the strongest argument that alpha satellite in fact comprises the functional centromere of normal human chromosomes (26). Whereas formation of mitotically stable artificial chromosomes from individual components provides an assay for the function of centromeric DNA, an alternative strategy employs engineering of already existing, natural chromosomes to provide information on the function of individual elements in their chromatin state. In this strategy, telomere-associated chromosome truncation is used to delete most or all of the chromosome arms, thus generating stable “minichromosomes” that can be as small as several million base pairs (27, 28). (By comparison, even the smallest normal human chromosome is some 50 million base pairs in length.) This approach has been used (28) to carefully engineer chromosomes for the purposes of mapping and characterizing aspects of normal centromere structure and

behavior. With respect to centromere function, an important distinction between the two assays is that one tests the functionality of already competent centromeric chromatin, whereas the other tests the functionality of purified DNA, which must presumably be packaged into chromatin before becoming functional (refs. 3 and 4; Fig. 1).

Saffery *et al.* (2) have used both approaches to examine the centromere competence of neocentromere sequences from a well characterized chromosome 10-derived marker chromosome called mardel(10). In this chromosome, the neocentromere has been previously localized to a fully sequenced ≈ 80 -kb region originating from 10q25 (7, 29). First, Saffery *et al.* used the neocentromeric DNA as a candidate centromere in a human artificial chromosome assay. In over 450 clones examined, no artificial chromosomes were detected, in contrast to alpha satellite-containing input sequences that routinely generate artificial chromosomes in up to 50% of clones tested (23–25). Similarly, they were unable to generate an artificial chromosome by using a modified yeast artificial chromosome approach (30), substituting neocentromeric DNA for alpha satellite. The inability of neocentromeric sequences to form artificial chromosomes by using either approach argues strongly that these sequences—at least in the form of naked DNA—are less efficient than alpha satellite DNA at forming a *de novo* functional centromere.

To test the ability of chromatin at the mardel(10) neocentromere to function, they then used the minichromosome assay, truncating mardel(10) with a telomere-containing construct containing targeting DNA from the chromosome arms (2). The result-

ing minichromosomes were estimated to be <1 – 2 million base pairs in size, among the smallest human minichromosomes yet engineered. Extensive genomic characterization and functional testing provided direct evidence for the function of the mardel(10) neocentromere.

The conflicting results obtained by Saffery *et al.* (2) when using the artificial chromosome assay and the minichromosome assay likely reflect that the two assays measure different aspects of centromere assembly and function (Fig. 1). The biological relevance of the initial steps of centromere assembly in the *de novo* assay is open to some question, as naked centromeric DNA presumably does not exist in Nature. Nonetheless, remodeling of centromeric chromatin likely occurs during development and during epigenetic resolution of unusual centromere configurations, as must occur during centromere inactivation in dicentric chromosomes or in activation of neocentromeres (3, 4, 11). Formation of *de novo* centromeres in the artificial chromosome assay thus may provide a suitable model for exploring aspects of and requirements for centromere remodeling, which likely involves incorporation of the centromere-specific histone variant CENP-A (18). In contrast to the *de novo* assay, the engineering of minichromosomes does not involve generation or remodeling of centromeric chromatin; rather, it utilizes an already functional centromere (or neocentromere) as the basis for engineering smaller derivatives that are still capable of normal chromosome function. One particularly intriguing aspect of centromere organization and function raised by neocentromeres and by the aforementioned variation among centromeric DNAs from different species and organisms is the extent to which a centromere from one organism can substitute for that of another (1).

Lastly, as pointed out by Saffery *et al.* (2), the generation of minichromosomes based on neocentromeres provides an alternative to strategies that consider the use of artificial chromosomes as potential nonviral gene therapy or gene transfer vectors (23, 26, 31). In principle, it may be possible to make neocentromere-based minichromosomes (2) that are smaller than alpha satellite-based minichromosomes (27, 28, 32). However, it will be necessary to test this premise directly, as the lower limit of neocentromeric DNA required hasn't been established. Notwithstanding the conceptual advantages of this approach, it is important to recall just how inefficient chromosome engineering in mammalian cells is compared with other, high-efficiency systems (21). Natural or unnatural, the potential of human artificial chromosomes may not be fully realized until more tractable approaches for chromosome engineering, handling, and transfer are developed and explored.

I thank Brenda Grimes, Katie Rudd, and Bala Balakumaran for helpful discussions. Work on

artificial chromosomes in our laboratory is supported by a Franklin Delano Roosevelt Re-

search Award from the March of Dimes Birth Defects Foundation.

1. Tyler-Smith, C. & Florida, G. (2000) *Cell* **102**, 5–8.
2. Saffery, R., Wong, L. H., Irvine, D. V., Bateman, M. A., Griffiths, B., Cutts, S. M., Cancilla, M. R., Cendron, A. C., Stafford, A. J. & Choo, K. H. A. (2001) *Proc. Natl. Acad. Sci. USA* **98**, 5705–5710. (First Published May 1, 2001; 10.1073/pnas.091468498)
3. Willard, H. F. (1998) *Curr. Opin. Genet. Dev.* **8**, 219–225.
4. Choo, K. H. A. (2000) *Trends Cell Biol.* **10**, 182–188.
5. Rhoades, M. M. & Vilkomerson, H. (1942) *Proc. Natl. Acad. Sci. USA* **28**, 433–443.
6. Alfenito, M. R. & Birchler, J. A. (1993) *Genetics* **135**, 589–597.
7. Barry, A. E., Howman, E. V., Cancilla, M. R., Saffery, R. & Choo, K. H. (1999) *Hum. Mol. Genet.* **8**, 217–227.
8. Choo, K. H. A. (1997) *Am. J. Hum. Genet.* **61**, 1225–1233.
9. Depinet, T. W., Zackowski, J. L., Earnshaw, W. C., Kaffe, S., Sekhon, G. S., Stallard, R., Sullivan, B. A., Vance, G. H., Van Dyke, D. L., Willard, H. F., *et al.* (1997) *Hum. Mol. Genet.* **6**, 1195–1204.
10. Warburton, P. E., Dolled, M., Mahmood, R., Alonso, A., Li, S., Naritomi, K., Tohma, T., Nagai, T., Hasegawa, T., Ohashi, H., *et al.* (2000) *Am. J. Hum. Genet.* **66**, 1794–1806.
11. Karpen, G. H. & Allshire, R. C. (1997) *Trends Genet.* **13**, 489–496.
12. Koch, J. (2000) *Hum. Mol. Genet.* **9**, 149–154.
13. Florida, G., Zatterale, A., Zuffardi, O. & Tyler-Smith, C. (2000) *EMBO Rep.* **1**, 489–493.
14. Csink, A. K. & Henikoff, S. (1998) *Trends Genet.* **14**, 200–204.
15. McClintock, B. (1939) *Proc. Natl. Acad. Sci. USA* **25**, 405–416.
16. Sullivan, B. A. & Schwartz, S. (1995) *Hum. Mol. Genet.* **4**, 2189–2197.
17. Vafa, O. & Sullivan, K. F. (1997) *Curr. Biol.* **7**, 897–900.
18. Sullivan, K. F. (2001) *Curr. Opin. Genet. Dev.* **11**, 182–188.
19. Saffery, R., Irvine, D. V., Griffiths, B., Kalitsis, P., Wordeman, L. & Choo, K. H. (2000) *Hum. Mol. Genet.* **9**, 175–185.
20. Lo, A. W., Magliano, D. J., Sibson, M. C., Kalitsis, P., Craig, J. M. & Choo, K. H. (2001) *Genome Res.* **11**, 448–457.
21. Murray, A. W. & Szostak, J. W. (1983) *Nature (London)* **305**, 189–193.
22. Carbon, J. & Clarke, L. (1990) *New Biol.* **2**, 10–19.
23. Harrington, J. J., van Bokkelen, G., Mays, R. W., Gustashaw, K. & Willard, H. F. (1997) *Nat. Genet.* **15**, 345–355.
24. Ikeno, M., Grimes, B., Okazaki, T., Nakano, M., Saitoh, K., Hoshino, H., McGill, N. I., Cooke, H. J. & Masumoto, H. (1998) *Nat. Biotechnol.* **16**, 431–439.
25. Ebersole, T. A., Ross, A., Clark, E., McGill, N., Schindelbauer, D., Cooke, H. & Grimes, B. (2000) *Hum. Mol. Genet.* **9**, 1623–1631.
26. Willard, H. F. (2000) *Science* **290**, 1308–1309.
27. Farr, C. J., Bayne, R. A., Kipling, D., Mills, W., Critcher, R. & Cooke, H. J. (1995) *EMBO J.* **14**, 5444–5454.
28. Heller, R., Brown, K. E., Burgtorf, C. & Brown, W. R. (1996) *Proc. Natl. Acad. Sci. USA* **93**, 7125–7130.
29. DuSart, D., Cancilla, M. R., Earle, E., Mao, J. I., Saffery, M., Tainton, K. M., Kalitsis, P., Martyn, J., Barry, A. E. & Choo, K. H. A. (1997) *Nat. Genet.* **16**, 144–153.
30. Henning, K. A., Novotny, E. A., Compton, S. T., Guan, X. Y., Liu, P. P. & Ashlock, M. A. (1999) *Proc. Natl. Acad. Sci. USA* **96**, 592–597.
31. Brown, W. R., Mee, P. J. & Hong Shen, M. (2000) *Trends Biotechnol.* **18**, 218–223.
32. Yang, J. W., Pendon, C., Yang, J., Haywood, N., Chand, A. & Brown, W. R. (2000) *Hum. Mol. Genet.* **9**, 1891–1902.