

Anniversary of the discovery/isolation of the yeast centromere by Clarke and Carbon

Kerry Bloom

Biology Department, University of North Carolina, Chapel Hill, NC 27599

ABSTRACT The first centromere was isolated 35 years ago by Louise Clarke and John Carbon from budding yeast. They embarked on their journey with rudimentary molecular tools (by today's standards) and little knowledge of the structure of a chromosome, much less the nature of a centromere. Their discovery opened up a new field, as centromeres have now been isolated from fungi and numerous plants and animals, including mammals. Budding yeast and several other fungi have small centromeres with short, well-defined sequences, known as point centromeres, whereas regional centromeres span several kilobases up to megabases and do not seem to have DNA sequence specificity. Centromeres are at the heart of artificial chromosomes, and we have seen the birth of synthetic centromeres in budding and fission yeast and mammals. The diversity in centromeres throughout phylogeny belie conserved functions that are only beginning to be understood.

Monitoring Editor

Doug Kellogg
University of California,
Santa Cruz

Received: Jan 15, 2015

Revised: Jan 28, 2015

Accepted: Feb 18, 2015

It was ~35 years ago that centromere DNA was first discovered by Louise Clarke and John Carbon working at the University of California, Santa Barbara (Clarke and Carbon, 1980, 1985). To appreciate the pioneering aspect of this work, it is important to consider the questions at the time, and specifically the understanding of what constituted a chromosome. Chromosomes were visible in stained preparations (hence "chromosome" for "colored bodies"), but the physical proof that a chromosome was one linear DNA molecule was not established until the separation of intact chromosomal DNA by pulsed-field gradient gel electrophoresis (Schwartz and Cantor, 1984). The alternative hypothesis at the time was that each chromosome was built like a cable, with many strands wound around each other. Clarke and Carbon based their project on the kinetics of DNase cleavage (Gall, 1963) and nucleic acid reassociation kinetics (Britten and Kohne, 1968; Wetmur and Davidson, 1968). Gall had been examining lampbrush chromosomes, with loops radiating from a main axis. In a classic quantitative study, Gall used the rate of DNase digestion and a visual breakage assay to deduce that each chromatid contained one very long DNA double helix. Britten and Davidson were busy deducing the architecture of the genome (fraction of single-copy genes and middle repetitive and highly repetitive sequences), and from the size of the genome could infer the

amount of DNA/chromosome. Bacterial genomes were in the range of several million base pairs and, in the case of *Escherichia coli*, were contained in a single circular molecule. Clarke and Carbon reasoned that if the eukaryotic chromosome was a single, continuous molecule, one should be able to transform cells containing centromere-linked mutations, identify the gene by complementation, and use nucleic acid reassociation (hybridization) to walk from one side of the centromere to the other.

Molecular cloning was in its infancy, with complementation and yeast transformation just established a few years before (Hinnen *et al.*, 1978). There were no shuttle vectors for plasmid amplification in eukaryotes, no PCR, no genome sequences, and no software for DNA data analysis. Of most importance, there was no bioassay for the centromere. How would they know when they got there?

Undaunted by what would be difficult to justify in the current funding climate, Clarke and Carbon isolated genes on either side of the centromere on yeast chromosome III (*LEU2* and *PGK1*). One of the tractable features of yeast is that all products of meiotic cell division are contained in a single sac, the ascus. Homologous chromosomes segregate from each other in meiosis I, dictating that centromere-linked genes on nonsister chromatids segregate in meiosis I. This leads to a characteristic arrangement of genes in the ascus and a method to map the centromere and consequently centromere-linked genes. *LEU2* was not isolated by complementation of the yeast *leu2* mutation, as this preceded the establishment of yeast transformation in 1978. Instead, Ratzkin and Carbon (1977) reasoned that common biosynthetic pathways in prokaryotic and eukaryotic cells have common enzymes and therefore common genes. The yeast *LEU2* gene was isolated by complementation of an auxotrophic mutant in *E. coli* (*leuB*; Ratzkin and Carbon, 1977).

DOI:10.1091/mbc.E14-11-1512

Address correspondence to: Kerry Bloom (kerry_bloom@unc.edu).

© 2015 Bloom. This article is distributed by The American Society for Cell Biology under license from the author(s). Two months after publication it is available to the public under an Attribution–Noncommercial–Share Alike 3.0 Unported Creative Commons License (<http://creativecommons.org/licenses/by-nc-sa/3.0>).

"ASCB," "The American Society for Cell Biology," and "Molecular Biology of the Cell" are registered trademarks of The American Society for Cell Biology.

As proof of principle, they also identified the yeast *HIS2* gene by complementation of *hisB* *E. coli* mutants. On the other side of the centromere on chromosome III lies *PGK1*. *PGK1* catalyzes a step in the glycolytic pathway and is highly expressed in budding yeast. Carbon *et al.* (1978) had devised an immunological strategy to detect recombinant *E. coli* clones expressing various yeast enzymes, and Hitzeman *et al.* (1980) used this to isolate the *PGK1* gene.

The walk begins by radiolabeling each gene and hybridizing the labeled probes to identify *E. coli* colonies that contain some or all of each fragment. The library is made with randomly sheared fragments of the genomic DNA, and thus, on average, each colony contains differing pieces of the same region of DNA. Once a colony is identified, the plasmid DNA is isolated, mapped using restriction enzymes, and itself radiolabeled to take the next step. Each step can go in either direction, and thus progress to the centromere is twice the effort of the walk. One of the genomic landmarks discovered in this walk was the retrotransposon (yeast TY2; Kingsman *et al.*, 1981). Transposable elements had just been discovered and found to be dispersed throughout the genome (Cameron *et al.*, 1979). In addition, >100 “delta” sequences, which are repeated at the termini of TY elements, are also scattered in the genome, footprints of transposition. Repeated DNA is the curse of overlap hybridization strategists. Once a repeated region is encountered, many colonies “light up,” and there is little hope that one can “walk across” the repeat with the tools in hand.

Fortunately, many in the field were interested in John and Louise’s progress. John was on the phone with Lee Hartwell (University of Washington, Seattle, WA) discussing centromeres. Lee mentioned that one of the cell division cycle mutants (*cdc* mutants; Hartwell *et al.*, 1973), *CDC10*, was closely linked to the centromere on chromosome III. Clarke isolated a clone complementing the temperature-sensitive *cdc10* mutation. This clone contained an 8-kb fragment that overlapped with clones in the laboratory from the *LEU2* region. The hunt was on. *CDC10* is so close to the centromere that it was hard from genetic crossing-over data to distinguish whether it was on the side of *LEU2* or the other side. It was possible that this clone contained the highly sought centromere.

Sure enough, plasmids containing these sequences were stably maintained after many generations of nonselective growth in mitosis and Mendelian segregation in meiosis. These remain the hallmarks for centromere function. Mitotic stability was exploited by Hsiao and Carbon (1981) for the direct isolation of additional centromeres. They transformed a yeast DNA library into cells and simply grew the transformed cells in the absence of genetic selection for the complementing gene on the plasmid. This would have been heretical just 2 years before. After many rounds of nonselective growth, cells were plated on selective media. In this way, two of the known centromeres (*CEN3* and *CEN11*) and several others were identified without any chromosome walking, subcloning, or overlap hybridization. Students of mitosis are forever indebted to Clarke and Carbon’s bold experiment and Craig Chinault, an early chromosome walker (Chinault and Carbon, 1979).

The isolation of centromere from budding yeast marked the birth of several fields. Clarke went on to identify centromere DNA from *Schizosaccharomyces pombe* and defined regional centromeres. Carbon was interested in centromere function and went about isolating centromere DNA-binding proteins. Again, one must appreciate the scope of the problem. There are 16 chromosomes, each with a single centromere. Because an excess of centromeres leads to genetic instability (McClintock, 1938, 1953), there was every reason to believe that centromere-binding proteins would be present in extremely low abundance (1 set of proteins/centromere, ~16 per

cell). A breakthrough in the biochemical isolation of the DNA binding complex came when J. Lechner found that a chaperone protein (casein) was required to facilitate sequence-specific binding of the complex to centromere DNA and published the isolation of a 240-kDa complex, denoted *CBF3* for the three proteins in the complex (Lechner and Carbon, 1991). Two years later, the gene for the large subunit (*NDC10*, 110 kDa) of the complex was identified simultaneously by Jiang *et al.* (1993a) and Goh and Kilmartin (1993).

Ever the pioneer, Carbon isolated a second centromere-binding complex, *CBF5*, which had engaged his attention for several years (1993–1999). *CBF5* catalyzes the conversion of uridine to pseudouridine in rRNA and tRNA and is part of the H/ACA small nucleolar ribonucleoproteins. There were several genetic clues that *CBF5* had centromere function. Overexpression of *CBF5* suppresses temperature-sensitive mutants of *ndc10-1* (Jiang *et al.*, 1993b), and there is a conserved domain in *CBF5* that binds microtubules *in vitro*. It took decades to link *CBF5* back to the centromere—appropriately, through tRNAs and their genes, John’s first love (ca. 1960s to 1970s).

John and Louise are responsible for defining the point centromere, the smallest known centromere in phylogeny. As I am forever indebted to their mentorship, as well as for launching my independent career, the reader can appreciate my dismay when a colleague (J. Haber, Brandeis University, Waltham, MA) introduced me as one “who has spent a career trying to make the centromere bigger” (Bloom, 2014). The individual yeast centromeres are clustered at the plus ends of 16 kinetochore microtubules that are cylindrically arrayed around the central pole–pole microtubules of the yeast spindle (Figure 1). The entire cluster is ~250 nm in diameter and ~70 nm wide. However, sister centromeres are separated by ~800 nm in metaphase. The physical separation between sister kinetochores is highly conserved throughout phylogeny, in spite of the fact that centromere DNA spans more than four orders of magnitude (from yeast to human). The centromeric heterochromatin must be folded into a highly compact and organized structure, suited for transmitting mechanical force. An evolutionarily conserved pathway—the spindle assembly checkpoint—monitors the status of the kinetochore microtubule attachment site, including the presence or absence of a microtubule and whether tension is generated between sister kinetochores. The error-correction mechanisms and how the cell promotes stable versus unstable microtubule attachment in response to the state of each sister kinetochore are embedded in the structure of the kinetochore and the centromeric heterochromatin.

If we take a systems approach to the centromere, we find the entire yeast spindle analogous to a single mammalian kinetochore (Bloom, 2014). Even though budding yeast does not have repeat satellite sequences flanking the centromere, tRNA genes are enriched approximately twofold in the 50 kb surrounding the centromere (Snider *et al.*, 2014). The *CBF5* pseudouridine synthase, along with tRNA transcription factors (e.g., TFC3), is responsible for the enrichment of condensin to the pericentromere chromatin (Snider *et al.*, 2014). Binding of Cbf5 brings condensin proximal to the transfer DNA (tDNA) genes, which, through aggregation of multiple tDNAs, results in the gathering of pericentromeres of different chromosomes. This leads to the concentration of condensin along the spindle axis and mechanical integrity of the network. In this way, the individual centromeres in budding yeast function as an integrated unit (Stephens *et al.*, 2013). Cbf5 is important for centromere function, but at the level of centromeric heterochromatin rather than the microtubule attachment site.

From the inauspicious beginnings of a 125-bp CEN DNA, we currently face ~1 Mb of centromeric heterochromatin (800 kb: 16 chromosomes times a 50-kb region around each centromere

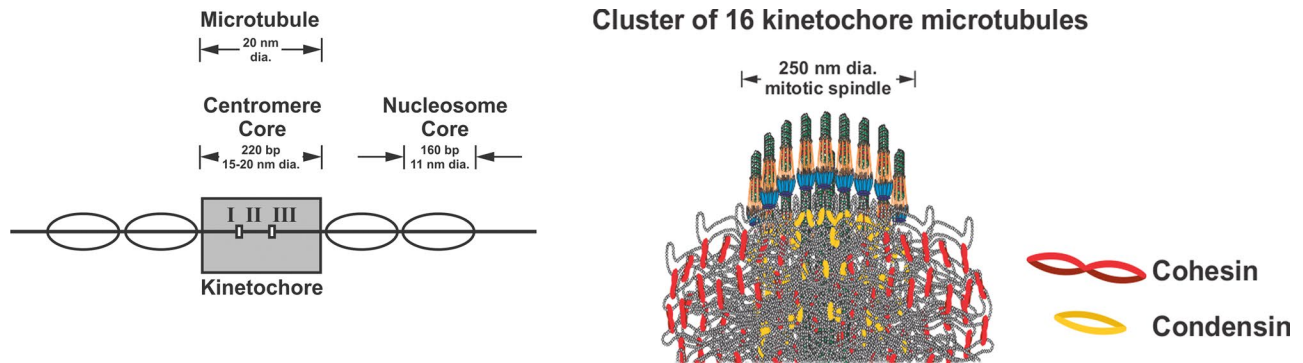


FIGURE 1: Left, the size relationship between the centromere nuclease-protected region and a single microtubule. The nucleosome is 5×11 nm, with 146 base pairs of DNA wrapped around a core of eight histone molecules. Based on the size of the nucleosome, it was estimated that the kinetochore was slightly larger, due to the increase in protected DNA (220-bp centromere vs. 160-bp nucleosome; Bloom and Carbon, 1982; Bloom *et al.*, 1983). Right, the size relationship between centromeric heterochromatin and the 16 kinetochore microtubules in the half-spindle in yeast. The 16 yeast centromeres are clustered at the plus ends of 16 kinetochore microtubules. The sister kinetochores in the other half of the spindle are not shown. The centromeres of all 16 chromosomes are cross-linked via cohesin and condensin, proteins enriched in the 50 kb surrounding each centromere (Stephens *et al.*, 2011, 2013). From a systems biology perspective, one can consider the 50 kb of each of 16 centromeres to be the equivalent of a mammalian centromere (800-kb yeast vs. 1- to 5-Mb mammalian centromere).

enriched in tDNA genes, condensin, and cohesin and proximal to the spindle axis) in metaphase. We have developed methods to visualize individual chromosomes and their centromeres in live cells (Robinett *et al.*, 1996; Straight *et al.*, 1996, 1997; Pearson *et al.*, 2001), we can track chromosome dynamics in space and time, and we can incorporate principles from polymer physics to build intuition. What is certain is that more surprises await, as we aspire to the Clarke/Carbon standard for bold experiments that build lasting foundations.

REFERENCES

- Bloom KS (2014). Centromeric heterochromatin: the primordial segregation machine. *Annu Rev Genet* 48, 457–484.
- Bloom KS, Carbon J (1982). Yeast centromere DNA is in a unique and highly ordered structure in chromosomes and small circular minichromosomes. *Cell* 29, 305–317.
- Bloom KS, Fitzgerald-Hayes M, Carbon J (1983). Structural analysis and sequence organization of yeast centromeres. *Cold Spring Harb Symp Quant Biol* 47, 1175–1185.
- Britten RJ, Kohne DE (1968). Repeated sequences in DNA. Hundreds of thousands of copies of DNA sequences have been incorporated into the genomes of higher organisms. *Science* 161, 529–540.
- Cameron JR, Loh EY, Davis RW (1979). Evidence for transposition of dispersed repetitive DNA families in yeast. *Cell* 16, 739–751.
- Carbon J, Clarke L, Chinault AC, Ratzkin B, Walz A (1978). The isolation and characterization of specific gene systems from the yeast, *Saccharomyces cerevisiae*. In: *Biochemistry and Genetics of Yeast*, ed. M Bacila, New York: Academic Press, 425–443.
- Chinault AC, Carbon J (1979). Overlap hybridization screening: isolation and characterization of overlapping DNA fragments surrounding the *leu2* gene on yeast chromosome III. *Gene* 5, 111–126.
- Clarke L, Carbon J (1980). Isolation of a yeast centromere and construction of functional small circular chromosomes. *Nature* 287, 504–509.
- Clarke L, Carbon J (1985). The structure and function of yeast centromeres. *Annu Rev Genet* 19, 29–55.
- Gall JG (1963). Kinetics of deoxyribonuclease action on chromosomes. *Nature* 198, 36–38.
- Goh PY, Kilmartin JV (1993). NDC10: a gene involved in chromosome segregation in *Saccharomyces cerevisiae*. *J Cell Biol* 121, 503–512.
- Hartwell LH, Mortimer RK, Culotti J, Culotti M (1973). Genetic control of the cell division cycle in yeast: V. Genetic analysis of *cdc* mutants. *Genetics* 74, 267–286.
- Hinnen A, Hicks JB, Fink GR (1978). Transformation of yeast. *Proc Natl Acad Sci USA* 75, 1929–1933.
- Hitzeman RA, Clarke L, Carbon J (1980). Isolation and characterization of the yeast 3-phosphoglycerokinase gene (*PGK*) by an immunological screening technique. *J Biol Chem* 255, 12073–12080.
- Hsiao CL, Carbon J (1981). Direct selection procedure for the isolation of functional centromeric DNA. *Proc Natl Acad Sci USA* 78, 3760–3764.
- Jiang W, Lechner J, Carbon J (1993a). Isolation and characterization of a gene (*CBF2*) specifying a protein component of the budding yeast kinetochore. *J Cell Biol* 121, 513–519.
- Jiang W, Middleton K, Yoon HJ, Fouquet C, Carbon J (1993b). An essential yeast protein, *CBF5p*, binds in vitro to centromeres and microtubules. *Mol Cell Biol* 13, 4884–4893.
- Kingsman AJ, Gimlich RL, Clarke L, Chinault AC, Carbon J (1981). Sequence variation in dispersed repetitive sequences in *Saccharomyces cerevisiae*. *J Mol Biol* 145, 619–632.
- Lechner J, Carbon J (1991). A 240 kd multisubunit protein complex, *CBF3*, is a major component of the budding yeast centromere. *Cell* 64, 717–725.
- McClintock B (1938). The production of homozygous deficient tissues with mutant characteristics by means of the aberrant mitotic behavior of ring-shaped chromosomes. *Genetics* 23, 315–376.
- McClintock B (1953). Induction of instability at selected loci in maize. *Genetics* 38, 579–599.
- Pearson CG, Maddox PS, Salmon ED, Bloom K (2001). Budding yeast chromosome structure and dynamics during mitosis. *J Cell Biol* 152, 1255–1266.
- Ratzkin B, Carbon J (1977). Functional expression of cloned yeast DNA in *Escherichia coli*. *Proc Natl Acad Sci USA* 74, 487–491.
- Robinett CC, Straight A, Li G, Wilhelm C, Sudlow G, Murray A, Belmont AS (1996). In vivo localization of DNA sequences and visualization of large-scale chromatin organization using lac operator/repressor recognition. *J Cell Biol* 135, 1685–1700.
- Schwartz DC, Cantor CR (1984). Separation of yeast chromosome-sized DNAs by pulsed field gradient gel electrophoresis. *Cell* 37, 67–75.
- Snider CE, Stephens AD, Kirkland JG, Hamdani O, Kamakaka RT, Bloom K (2014). Dyskerin, tRNA genes, and condensin tether pericentric chromatin to the spindle axis in mitosis. *J Cell Biol* 207, 189–199.
- Stephens AD, Haase J, Vicci L, Taylor RM 2nd, Bloom K (2011). Cohesin, condensin, and the intramolecular centromere loop together generate the mitotic chromatin spring. *J Cell Biol* 193, 1167–1180.
- Stephens AD, Snider CE, Haase J, Haggerty RA, Vasquez PA, Forest MG, Bloom K (2013). Individual pericentromeres display coordinated motion and stretching in the yeast spindle. *J Cell Biol* 203, 407–416.
- Straight AF, Belmont AS, Robinett CC, Murray AW (1996). GFP tagging of budding yeast chromosomes reveals that protein-protein interactions can mediate sister chromatid cohesion. *Curr Biol* 6, 1599–1608.
- Straight AF, Marshall WF, Sedat JW, Murray AW (1997). Mitosis in living budding yeast: anaphase A but no metaphase plate. *Science* 277, 574–578.
- Wetmur JG, Davidson N (1968). Kinetics of renaturation of DNA. *J Mol Biol* 31, 349–370.