*Proc. Natl. Acad. Sci. USA*
Vol. 93, pp. 9997–10000, September 1996

# Commentary

# Coordinating DNA replication with cell division: Current status of the licensing concept

*Michael Botchan**

*Department of Molecular and Cell Biology, 401 Barker Hall, University of California, Berkeley, CA 94720-3204*

The replication of the genetic material during cell doubling is regulated at three apparently distinct levels. In the eukaryote nucleus, between $10^3$ and $10^5$ DNA replication initiation events occur each cell cycle (1). Though these events can occur asynchronously and can occur over an extended period of many hours, only a single initiation event is allowed to happen at any one site. Thus, on a long DNA fiber, multiple replication "eyes" that reveal start sites can be observed, but reinitiation within one eye is not detected (2). This once-and-only-once regulation of replication initiation helps maintain the relative copy numbers of genes. Furthermore, as first revealed by autoradiographic methods, the order of initiation events is determined, with certain regions replicating before others. This pattern can change during development, leading to the suggestion that such changes in the timing of initiation may result in a reprogramming of the gene expression program (3). At an apparently higher level of regulation, the entire synthesis phase (S) takes place within a discrete period, surrounded in time by two gap (G) periods, and resetting the clock to undergo a new round of DNA synthesis is dependent on passage through the mitosis (M) phase of the cell cycle. A classic demonstration of this interdependency is provided by cell fusion experiments in which heterokaryons containing combinations of S phase and G1 or G2 nuclei were observed; the DNA complement of the G1 nucleus was observed to enter S phase prematurely, while the G2 nucleus did not pass through another round of DNA replication without prior passage through mitosis (4).

Determining how these multiple forms of regulation are braided with respect to each other is a formidable problem, and surely the reflex intuition is that any model that conflates these different phenomena into a single notion would be an oversimplification. One is still unclear as to how choice of replication initiation sites may be altered in different cell lineages and why regulation at this level evolved. However, recent work focused upon the genetics of the cell cycle and upon two heteromeric proteins, the MCM [so named for their minichromosome (plasmid) maintenance functions] complex and origin recognition complexes (ORC), suggest that the mechanism that ensures that DNA replication occurs in each cell cycle once and only once is related to the mechanism that makes S phase dependent upon mitosis.

The history of the DNA replication licensing concept begins with studies focused upon amphibian eggs. The mature unfertilized *Xenopus* egg is arrested in metaphase of meiosis II, and a variety of treatments, including incubation with calcium ionophores, can mimic fertilization and synchronously activates the cell. Extracts prepared from such activated eggs have provided biochemists with a rich source of material (5) for the study of the initiation of DNA replication. Blow and Laskey (6, 7) found that when sperm chromatin is added to such extracts in the presence of protein synthesis inhibitors, the nucleus that forms around this chromatin undergoes a single round of DNA replication. M phase was blocked because, as we now know, cyclin B needs to be synthesized to trigger cell cycle progression. It could be shown that the reason that this postreplication (G2) nucleus did not reinitiate DNA synthesis was not that the extracts had run out of important components—the G2 nuclei could be reintroduced to fresh extracts and would still not replicate in the presence of protein synthesis inhibitors—but that they had become incompetent to do so. Disruption of the nuclear membrane by pricking with a needle or by reagents that compromised the integrity of the nuclear membrane led to reestablishment of competence: the G2 nuclei whose nuclear membranes were so impaired could bypass mitosis and reinitiate DNA synthesis after reintroduction into fresh extracts. This was the simple yet central observation that led to the formulation of the DNA replication "licensing" hypothesis.

Blow and Laskey speculated that a critical nondiffusable component physically marks the origin site as competent for replication and during S phase is destroyed by the act of replication. In this model, reestablishment of competence for replication (or licensing) occurs upon nuclear membrane breakdown. In the metazoan cell cycle, this breakdown occurs naturally at the onset of mitosis, and it was posited that replication licensing would occur during mitosis. The factors responsible for licensing should be in excess in the cytoplasm (at least in the early *Xenopus* extracts), chromatin-associated in G1, but lost during S phase. The power of this heuristic model was that it made strong predictions about how to look for important regulators of DNA replication and implied a straightforward biochemical assay for their identification. The model not only addresses the S phase requirement for M phase, but also why any given origin of DNA replication would fire only once in an S phase—the act of firing destroys the licensing activity at the locus.

Meanwhile genetic approaches in budding yeast uncovered new members of the DNA replication machinery. Two of these new genes, *CDC45* and *CDC54*, were identified by cold-sensitive mutation that led to arrest before the DNA replication phase (8). In another screen, *Saccharomyces cerevisiae* mutants were found that could not maintain particular plasmids (9). This latter screen was designed to find specific DNA initiator functions, applying the plausible notion that some mutant alleles would lead to a differential loss of two plasmids that are dependent upon separate replicators and initiators. Many such genes were found and they were named MCM for their minichromosome (plasmid) maintenance function. Some of these genes are important for all replication in *Sac. cerevisiae* and it is unknown why any of them affect certain replicons more than others. For example, the *mcm1-1* allele defines a transcription factor critical for a broad range of functions.

The analysis of replication function in yeast and the exploration of the licensing phenomenon came together with the observations of Hennessy *et al.* (10) and Chen *et al.* (11). Alleles of the *CDC46* gene were extragenic suppressors of *cdc45* or *cdc54* in the cold and also caused arrest at the G1/S boundary at 37°C. Moreover, the Cdc46p became visible in the nucleus immediately after anaphase and remained nuclear until replication began. The execution point of *CDC46* was found to be downstream of start but before that of the DNA

*e-mail: mbotchan@uclink2.berkeley.edu.

polymerases. The cell cycle variation of Cdc46p localization and the phenotypes of *CDC46* mutants therefore fit that of a protein involved in licensing. Chen *et al.* (11) showed that *CDC46* was the same as *MCM5* and that *MCM2, 3*, and *5* encoded for a family of structurally related proteins. The allele specific suppressive interactions between these genes were highly suggestive of a protein complex or at the least a focus of coordinated activity at time that DNA replication is initiated.

Some serendipity and an extensive search for other MCM homologous genes has yielded an enormous list of members in a wide range of organisms. Significantly, an inactivating mutation in the *Drosophila* MCM2 homologue produces phenotypes in the embryo consistent with an *in vivo* role in the regulation of DNA replication, analogous to that of its yeast counterpart (42). The nomenclature for the MCM family members is somewhat confusing, but a clarification can now be found in Chong *et al.* (12). Comparing sequences across phyla provides insight. Firstly, six distinct subgroups within the family can be discerned, which have been ascribed the names *MCM2–MCM7*; some of these names have no relationships to the nomenclature of the original *mcm* alleles. To date, in any one organism, each type of MCM gene appears to be represented only once and a total of five MCMs seems to be the maximum per creature. For example, both *Sac. cerevisiae* and humans have one copy each of MCMs 2, 3, 4, 5, and 7. The proteins each have sequences predicted for DNA-dependent ATPases (13) and have molecular masses ranging from 80 to 120 kDa. Generation of antibodies to these proteins has made it possible to demonstrate that they are indeed carried in the cell as heteromeric complexes and that these complexes are critical in a direct way for DNA replication (14–16); therefore *Xenopus* extracts depleted of the MCM complex (with anti-Mcm3 reagents) can replicate G1 nuclei from HeLa cells but not permeabilized G2 nuclei (15), and chromatin exposed to licensing-competent extracts binds *Xenopus* Mcm3 (14–16).

Chong *et al.* (16) exploited the observation that protein kinase inhibitors inactivate competent mitotic extracts for the licensing process. They prepared biochemical fractions from competent mitotic extracts and incubated them with sperm chromatin. The templates were then introduced into licensing-deficient extracts and DNA replication assays were performed. An obligatory fraction, called RLF-M (containing the MCM complex), included at least three distinct polypeptides. One of these was the *Xenopus* Mcm3p homologue, and a second was likely × Mcm2p, as it crossreacted with antibodies raised against the human Mcm2p homologue. Another required fraction not containing the heteromeric MCM complex was called RLF-B. If either of these fractions is to be called the licensing factor—somewhat ironically given the history—it would have to be the RLF-B component, as we now know that the MCM complex is constitutively nuclear in mammals and in *Xenopus* (17–20) and that the protein can freely locate into the nucleus. However, the association between chromosomes and the MCM complex can still be fit into the licensing model. The tight binding of MCM complex to chromatin occurs in telophase, as shown by confocal microscopy, and although it is nuclear in S phase, MCMs are not associated with replication foci. These observations are consistent with biochemical results, which indicate that the bulk of the MCM complex is easily dissociated from nucleoprotein complexes after replication begins. MCM chromatin binding is likely to be regulated by RFL-B, and clearly we need to know more about RLF-B and its disposition through the cell cycle. The licensing model would predict large stores of RLF-B in the *Xenopus* egg cytoplasm and its transient activation in loading the pivotal MCM complex onto chromosomes at mitosis.

Important new information concerning the composition and cell cycle behavior of the *Xenopus* MCM complex is provided by Romanowski *et al.* (20) in this issue of the *Proceedings*. Employing the yeast two-hybrid screen protocol with an XMCM3–LexA fusion protein as bait, they were able to clone the *Xenopus* MCM7 homologue. Their view of the composition of the heteromeric complex is the clearest to date: they show that it includes the *Xenopus* Mcm2, 3, 4, 5, and 7 proteins. They conclude that XMcm3 and 7 proteins stay together throughout interphase and M phase, whether on or off the chromosomes. Thus their observations are consistent with a model that has a single MCM complex whose activities can be modulated by other regulators.

Not all of the MCM complex releases from chromatin upon the initiation of S phase (21), and one wonders if the gradual release is related to the asynchronous triggering of synthesis at different DNA replication origin sites. How close is the MCM complex to the DNA replication initiation site, and is there any evidence that directly links them together? An intimate relationship is suggested from studies of replication origin binding proteins.

The ORC was first isolated from *Sac. cerevisiae* based upon its ability to bind specifically to the core DNA element common to all DNA replication origins in the yeast genome (22). A single type of such a complex is present in budding yeast, and the complex is required for replication *in vivo* (23, 24). This yeast heteromeric complex is composed of six different polypeptides, and its structure seems to be conserved in evolution, as a six-protein complex with homologous proteins has been purified from *Drosophila* embryos (25). The metazoan ORC is also important for replication initiation, as *Xenopus* extracts immunodepleted for XOrc2p lose initiation capacity (26). It is thus generally thought that the ORC proteins bind to and mark the spot that will be (or was) the replication origin. Observations also consistent with this view are that ORC seems to be a landing pad for many proteins known to be important for DNA replication. Included in this list of ORC binding factors are the MCM proteins (B. Stillman, personal communication). An indirect link between the MCM complex and ORC is also provided by genetics. The kinase Cdc7p is an evolutionarily conserved protein, and, in budding yeast, its activity is required for the G1/S transition, with an execution point very near that of the Cdc46/Mcm5p. Cdc7p is also known to be associated with the ORC complex and requires the subunit Dbf4p to target the activity to the complex (27). The recessive *bob1* mutation (28) bypasses the requirement for *CDC7* entirely, and *bob1* has been found to be an allele of *MCM5/CDC46* (Chris Hardy and R. Sclafani, personal communication). Thus a notion consistent with these findings is that a "loaded" MCM complex in association with ORC at the origin site must be phosphorylated by Cdc7p to initiate replication, thereby removing MCMs and thus the license to replicate. Mutant alleles of some components of the MCM complex seem to bypass the requirement for this phosphorylation.

The list of proteins required for initiation is not complete and at least yet one other element is involved. The Cdc6 protein of *Sac. cerevisiae* plays a key role in forming prereplicative complexes at the origin site and has properties in common with factors that should be prime candidates for involvement in licensing. For example, Cdc6p is synthesized in mitosis (29), but mutant alleles are defective in the G1/S transition. Diffley *et al.* (30) have used *in vivo* footprinting techniques to examine the protein occupancy of the DNA at the origin core site throughout the budding yeast cell cycle. Most intriguing, they find that the protective pattern in G2 cells closely resembles that of the purified ORC with DNA, and that late in mitosis (in anaphase just after the *CDC15* execution point) this protection increases by ≈50 bp. In striking analogy to the behavior of the MCM complex, this extended footprint is lost during S phase. This preinitiation footprint and its half-life are dependent upon the action of Cdc6p (31). Though Cdc6p is required for this behavior and binds ORC (24), the two are apparently not sufficient *in vitro* to mimic the *in vivo*

protection pattern. A speculation that emerges from the above correlations is that Cdc6p is the loading factor (RLF-B) for the MCM complex at anaphase. Such a pivotal role for Cdc6p would be consistent with the activities of Cdc18p, its closest relative in *Schizosaccharomyces pombe*. Overproduction of Cdc18p either by direct or indirect protocols (32, 33) leads to rereplication and thus can uncouple S phase from mitosis. If the Cdc6p/18 activity is a key focus for regulatory interactions, it would follow that a number of mutants affecting this single factor might be uncovered that would directly or indirectly lead to overreplication (34, 35). In contrast, the abundant and heteromeric MCM complex might be a more difficult target for such ends.

Given that the MCM complex, ORC, and Cdc6/18p are key players in the machine that allows for DNA replication initiation, how are their activities and their associations themselves regulated? Here the evidence is strong that the cyclin-dependent kinases (CDKs), the overseers of the master cell cycle clock, are primary (34, 35). The overarching coordination of the multiple cell cycle programs, which occur within a cell doubling (of which DNA replication is but one), is known to be achieved by a network of feedback loops that themselves determine the activity of CDKs (36, 37). The current metaphor for this multidimensional network is checkpoint control, in which, it is suggested, "retrograde" signals emanating from certain processes halt the cell cycle clock until the processes are complete. The licensing process wonderfully lends itself to such a stepwise logic, yet a key issue remains: identifying the major targets for kinase and phosphatase regulation. These major themes are summarized in the model in Fig. 1.

For the DNA replication cycle, as for other areas of checkpoint control, it is not clear what is actually being checked at the checkpoint at a biochemical level, and how the signals are really being generated is also unknown. In this context, it is interesting to note that *Sch. pombe* strains with deletions of the $Cdc18^{+}$ gene skip S phase altogether and proceed from G1 to mitosis (38). This observation leads to the hypothesis that perhaps the cell cycle feedback systems get input from preinitiation complexes; if none exist, the impetus is to go directly toward mitosis (38). However, once replication begins, signals from the synthetic apparatus itself check in. For example, treatments that block DNA synthesis in S phase halt overall cell cycle progression, and the DNA polymerase epsilon has a discrete domain that participates mysteriously in generating this retrograde signal (39).

Refocusing on the issue of how the licensing process in particular is regulated, Romanowski *et al.* (20) emphasize in their report the observation that the MCM proteins become hyperphosphorylated during S phase, reaching a peak in G2 phase. This state may therefore be downstream of CDC2–cyclin kinase activity. Concomitant with the metaphase-to-anaphase transition, the B-type cyclins are degraded and the MCM proteins are then known to be underphosphorylated, correlating with their binding to chromosomes. Thus hyperphosphorylation of the MCM proteins may prevent loading. This would not be incompatible with the notion that other phosphorylations at the G1/S transition may be required to get replication started. Other observations suggest that Cdc6p/18p (loading factor?) function is also critically controlled by the cyclin-dependent kinases. Jallepalli and Kelly (40) report that overexpression of the cyclin kinase inhibitors Rum-1 or SIC-1 lead to accumulation of Cdc18p and suggest that the stability and/or expression of this initiation factor is negatively regulated by kinase activity.

A final and perhaps premature speculation concerns the evolutionary diversity and complexity of the regulation that is brought to bear upon DNA replication initiation. It seems likely that different targets and ways to regulate the activity of the replication machine will emerge in organisms and cells with different types of doubling cycles. A comparison of the synthesis patterns for the Cdc6p/18p homologues between *Sac. cerevisiae*, *Sch. pombe*, and amphibian eggs illustrates this point. As mentioned above, Cdc6p synthesis in budding yeast occurs in mitosis in rapidly growing cells; however, in cells that have a long G1 period (for example, small daughter cells that require time to reach the critical size), a second burst of Cdc6p synthesis at the G1/S interface occurs that is required for DNA replication (29). In contrast the Cdc18p of *Sch. pombe* is synthesized only at the G1/S border (33). The differences may reflect evolutionary diversity and be related to the fact that *Sch. pombe* cells spend very little time in G1 and therefore have a very compressed period of time to license their replication

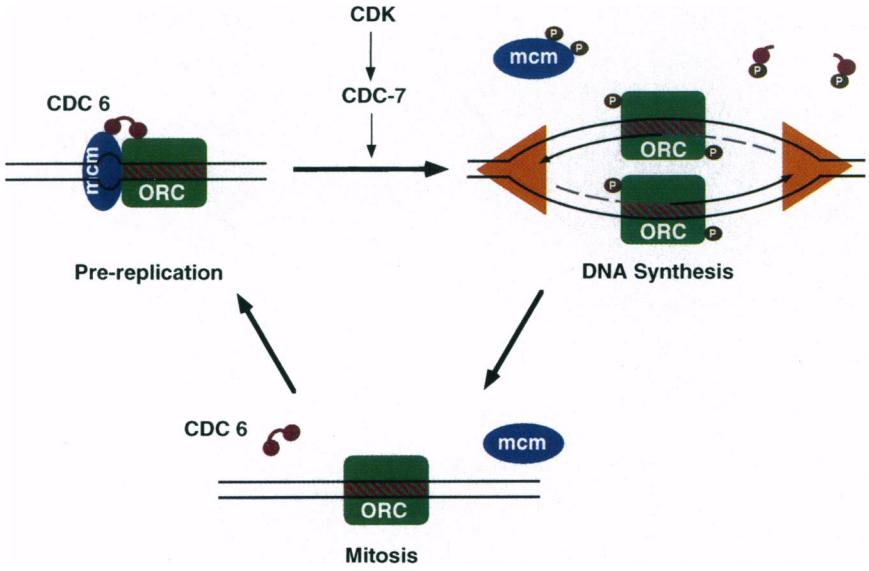


FIG. 1. Current speculative views on the initiation of DNA replication. The model indicates that DNA–origin initiation complexes undergo changes throughout the cell cycle. A fully licensed complex contains ORC (green), MCM complex (blue), and Cdc6p (red) and perhaps other unknown components. The cell cycle kinases trigger the cascade at the G1/S transition that leads to synthesis, release of the Mcm complex with its hyperphosphorylation, and the degradation of Cdc6p. Prereplication complexes probably exist in S phase at progressively lower levels, but the figure emphasizes the contrast between an origin site that has initiated synthesis from the other states. The growing fork is symbolized by the orange triangle. At mitosis, the factors are present again in the right modification states to load on the ori site.

origins for initiation. These different cell cycle styles would in turn perhaps require achieving the same ends through different circuitries. In amphibian eggs no *de novo* synthesis of RLF-B (CDC6/18?) is required at all, and large stores of the protein likely exist. This follows from the fact that multiple and complete cell cycle rounds can be observed in reconstituted frog extracts and that cyclin B mRNA is the only messenger RNA required (41).

If complexity and diversity mark the regulation of DNA synthesis in various creatures, one suspects, in contrast, that the actual mechanisms of DNA replication per se are highly conserved. This latter notion leads to the question of what the MCM complex, ORC, and CDC6/18, actually do. Licensing has been a powerful physiological and genetic concept, yet the biochemical correlates are still vague. There are a number of systems for which we understand, in detail, the mechanics of DNA replication. In general, when bidirectional DNA synthesis starts from within duplex DNA, a protein complex that marks the origin start site participates in a DNA distortion that eventually leads to duplex unwinding. It is not a far stretch to guess that the proteins involved in the replication licensing process accomplish at least that.


1. Hand, R. (1978) *Cell* **15,** 317–325.
2. Blumenthal, A. B., Kriegstein, M. J. & Hogness, D. S. (1974) *Cold Spring Harbor Symp. Quant. Biol.* **38,** 205–223.
3. Brown, D. D. (1984) *Cell* **37,** 359–365.
4. Rao, P. N. & Johnson, R. T. (1970) *Nature (London)* **225,** 159–164.
5. Lokha, M. J. & Masui, Y. (1983) *Science* **220,** 719–721.
6. Blow, J. J. & Laskey, R. A. (1986) *Cell* **47,** 577–587.
7. Blow, J. J. & Laskey, R. A. (1988) *Nature (London)* **332,** 546–548.
8. Moir, D., Stewart, S. E., Osmond, B. C. & Botstein, D. (1982) *Genetics* **100,** 547–564.
9. Maine, G. T., Sinha, P. & Tye, B. I. (1984) *Genetics* **106,** 365–385.
10. Hennessy, K. M., Clark, C. D. & Botstein, D. (1990) *Genes Dev.* **4,** 2252–2263.
11. Chen, Y., Hennessy, K. M., Botstein, D. & Tye, B. K. (1992) *Proc. Natl. Acad. Sci. USA* **89,** 10459–10463.
12. Chong, J. P. J., Thömmes, P. & Blow, J. J. (1996) *Trends Biochem. Sci.* **243,** 102–106.
13. Koonin, E. V. (1993) *Nucleic Acids Res.* **21,** 2541–2547.
14. Kubota, Y., Mimura, S., Nishimoto, S., Takisawa, H. & Nojima, H. (1995) *Cell* **81,** 601–609.
15. Madine, M., Khoo, C.-Y., Mills, A. D. & Laskey, R. A. (1995) *Nature (London)* **375,** 421–424.
16. Chong, J. P. J., Mahbubani, H. M., Khoo, C.-Y. & Blow, J. J. (1995) *Nature (London)* **375,** 418–421.
17. Thömmes, P., Fett, R., Schray, B., Burkhardt, R., Barnes, M., Kennedy, C., Brown, N. C. & Knippers, R. (1992) *Nucleic Acids Res.* **20,** 1069–1074.
18. Kimura, H., Wozaki, N. & Sugimoto, K. (1994) *EMBO J.* **13,** 4311–4320.
19. Todorov, I. T., Peppercock, R., Philipova, R. N., Kearsey, S. E., Ansorge, W. & Werner, D. (1994) *J. Cell Sci.* **107,** 253–265.
20. Romanowski, P., Madine, M. A. & Laskey, R. A. (1996) *Proc. Natl. Acad. Sci. USA* **93,** 10189–10194.
21. Madine, M. A., Khoo, C.-Y., Mills, A.d., Musahl, C. & Laskey, R. A. (1995) *Curr. Biol.* **5,** 1270–1279.
22. Bell, S. P. & Stillman, B. (1992) *Nature (London)* **357,** 128–134.
23. Fox, C. A., Loo, S., Dillin, A. & Rine, J. (1995) *Genes Dev.* **9,** 911–924.
24. Liang, C., Weinrich, M. & Stillman, B. (1995) *Cell* **81,** 667–676.
25. Gossen, M., Pak, D. T., Hansen, S. K., Acharya, J. K. & Botchan, M. R. (1995) *Science* **270,** 1674–1677.
26. Carpenter, P. B., Mueller, P. R. & Dunphy, W. G. (1996) *Nature (London)* **379,** 357–360.
27. Dowell, S. J., Romanowski, P. & Diffley, J. F. (1994) *Science* **265,** 1243–1246.
28. Jackson, A. L., Pahl, P. M. B., Harrison, K., Rosamond, J. & Sclafani, R. A. (1993) *Mol. Cell. Biol.* **13,** 2899–2908.
29. Piatti, S., Lengaves, C. & Nasmyth, K. (1995) *EMBO J.* **14,** 3788–3799.
30. Diffley, J. F., Cocker, J. H., Dowell, S. J. & Rowley, A. (1994) *Cell* **78,** 303–316.
31. Cocker, J. H., Piatti, S., Santocanale, C., Nasmyth, K. & Diffley, J. F. (1996) *Nature (London)* **379,** 180–182.
32. Nishitani, H. & Nurse, P. (1995) *Cell* **83,** 397–405.
33. Muzi-Falconi, M., Brown, G. W. & Kelly, T. J. (1996) *Proc. Natl. Acad. Sci. USA* **93,** 1666–1670.
34. Su, T. T., Follette, P. J. & O'Farrell, P. H. (1995) *Cell* **81,** 825–828.
35. Wuarin, J. & Nurse, P. (1996) *Cell* **85,** 785–787.
36. Hartwell, L. H. & Weinert, T. A. (1989) *Science* **246,** 629–634.
37. Morgan, D. O. (1995) *Nature (London)* **374,** 131–134.
38. Kelly, T. J., Martin, G. S., Forsburg, S. J., Stephen, R. J., Russo, A. & Nurse, P. (1993) *Cell* **74,** 371–382.
39. Navas, T. A., Zhou, Z. & Elledge, S. J. (1995) *Cell* **80,** 29–39.
40. Jallepalli, P. V. & Kelly, T. J. (1996) *Genes Dev.* **10,** 541–552.
41. Murray, A. W. & Kirschner, M. (1989) *Nature (London)* **339,** 275–280.
42. Treisman, J. E., Follette, P. J., O'Farrell, P. H. & Rubin, G. M. (1995) *Genes Dev.* **9,** 1705–1715.