Visualization of DNA Replication on Individual Epstein-Barr Virus Episomes

Paolo Norio* and Carl L. Schildkraut*

The duplication of the mammalian genome is an organized event, but there is limited information about the precision of the duplication program at specific genetic loci. We developed an approach that allows DNA replication events to be visualized in individual DNA molecules. Studying the latent replication of Epstein-Barr virus episomes, we show that different initiation sites are used to commence DNA replication from a specific portion of the viral genome (zone), whereas termination does not seem to be genomically defined. We conclude that initiation zones and pausing sites are major organizers of the duplication program, but initiation, fork progression, and termination of replication can vary in each molecule.

The Epstein-Barr virus (EBV) uses two independent systems to replicate its genome during the "lytic" and "latent" stages of infection (*1*, *2*). During latency, EBV replicates as a circular episome, once every cell division, in S phase, using the host-cell replication machinery (*3*). Studies performed by two-dimensional gel electrophoresis at neutral pH (2D gel) demonstrated that initiation of DNA replication occurs at many sites in the EBV genome (supplementary fig. 1) (*4*, *5*); however, it was not possible to exclude initiation throughout most of the viral genome. Because various origins of DNA replication are present, the duplication of the viral genome could follow either a precise order, similar for all the episomes, or initiate and progress in different ways in different molecules. To distinguish among these possibilities, we have developed an approach that allows us to detect, by immunofluorescence, the movement of DNA replication forks across the whole EBV genome in individual episomes stretched on microscope slides. We refer to this approach as single molecule analysis of replicated DNA (SMARD).

SMARD relies on two sequential periods of labeling with different halogenated nucleotides to label the replicating DNA molecules in an asynchronous population of cells (*5*). The panel labeled "Time 0" in Fig. 1A shows the stages of replication for a particular portion of the genome (gray lines) before the first labeling period (G_1, S, G_2, and M indicate the position in the cell cycle of the corresponding cell). When the first nucleotide analog is added, it is incorporated into the replicating DNA (shown in red) until the end of the first labeling period at Time 1. A second long labeling is then performed, with a different nucleotide analog (green color) and, at Time 2, the DNA molecules are either fully substituted with the nucleotide analogs (R, G, and RG), or partially substituted, or not substituted at all. The population of molecules fully substituted with both nucleotide analogs (RG in Fig. 1A) represents those molecules that started their duplication during the first labeling period and completed it during the second labeling period. In these molecules, the red-to-green transitions define the position of the replication forks at the time of the switch from the first labeling period to the second. Provided that the length of each labeling period is longer than the time required to replicate the region analyzed, the RG molecules are representative of all stages of DNA replication and can be used to infer how DNA replication initiated, progressed, and terminated across the genomic region analyzed. Whenever DNA replication initiates, progresses, and terminates in a similar way in all molecules, this order becomes evident by arranging the images of the molecules according to their increasing content of DNA labeled during the first labeling period (red). The staining pattern of the DNA molecules will depend on the position of initiation and termination sites along the molecules (Fig. 1, B to E, shows four different situations). In all these cases, the red portion of the molecules is nested around the region used to initiate DNA replication. On the contrary, the red portion of the molecules does not appear nested if molecules use different ways to initiate and terminate DNA replication (Fig. 1F).

The visualization of the replication patterns on the EBV episomes was performed by using exponentially growing Raji cells sequentially labeled with iododeoxyuridine and chlorodeoxyuridine (*5*). Agarose-embedded total DNA was prepared and the circular EBV episomes were converted to linear molecules by digestion with the restriction endonuclease Pac I (Fig. 2A) (*5*). The viral DNA was recovered from a pulsed-field gel by agarase treatment and stretched on silanized microscope slides by molecular combing (Fig. 2B) (*5*–*8*). The hybridization of specific biotinylated DNA probes (visualized by fluorescence microscopy) was used to distinguish the EBV genomes from the cellular genomic DNA that comigrates during pulsed-field gel electrophoresis and to define the orientations of the molecules (*5*). The patterns of incorporation of the nucleotide analogs along the DNA molecules were also detected by immunofluorescence with primary antibodies raised against the halogenated nucleotides (*5*, *9*).

Among 143 full-length EBV episomes recovered (fully substituted with the nucleotide analogs), 60 were fully red, 63 were fully green, and 20 were both red and green. Knowing the length of the first labeling period and the ratio between the number of molecules fully stained in red and the number of molecules fully stained in both red and green, it is possible to calculate the average time required to duplicate the EBV genome [supplementary fig. 2 (*5*)]. From the data reported above, we calculate that the average time required to duplicate an EBV episome in Raji cells is about 60 min.

To measure the positions of the replication forks along the viral episomes, we compared the immunostaining patterns obtained in various DNA molecules with the map of the EBV genome. We achieved this by using the positions of the DNA probes hybridized to the EBV genomes to digitally adjust the images of the molecules to a uniform size (Fig. 2C) [supplementary methods (*5*); for similar procedures see (*10*, *11*)]. Among the 20 red and green molecules, 15 were stretched sufficiently well to be studied in detail (Fig. 3). The remaining five, although unsuitable for precise measurements, showed similar replication patterns (*12*). The red portion of these molecules is nested around the ends of the linearized episomes. This indicates that, in Raji cells, the EBV genome follows an ordered pattern of replication that is similar for all episomes. The duplication commences in a specific portion of the viral genome approximately centered on the Pac I restriction site; then, replication progresses bidirectionally, to replicate the rest of the genome. However, a comparison of the replication patterns in molecules 1 and 2 shows that the short red patches span different portions of the EBV genome, indicating that these molecules used different initiation sites. This suggests the presence of an initiation zone rather than a discrete origin of replication, a concept originally proposed to explain the results of the 2D gel analysis in the

Department of Cell Biology, Albert Einstein College of Medicine, Bronx, NY 10461, USA.

*To whom correspondence should be addressed. E-mail: norio@aecom.yu.edu; schildkr@aecom.yu.edu

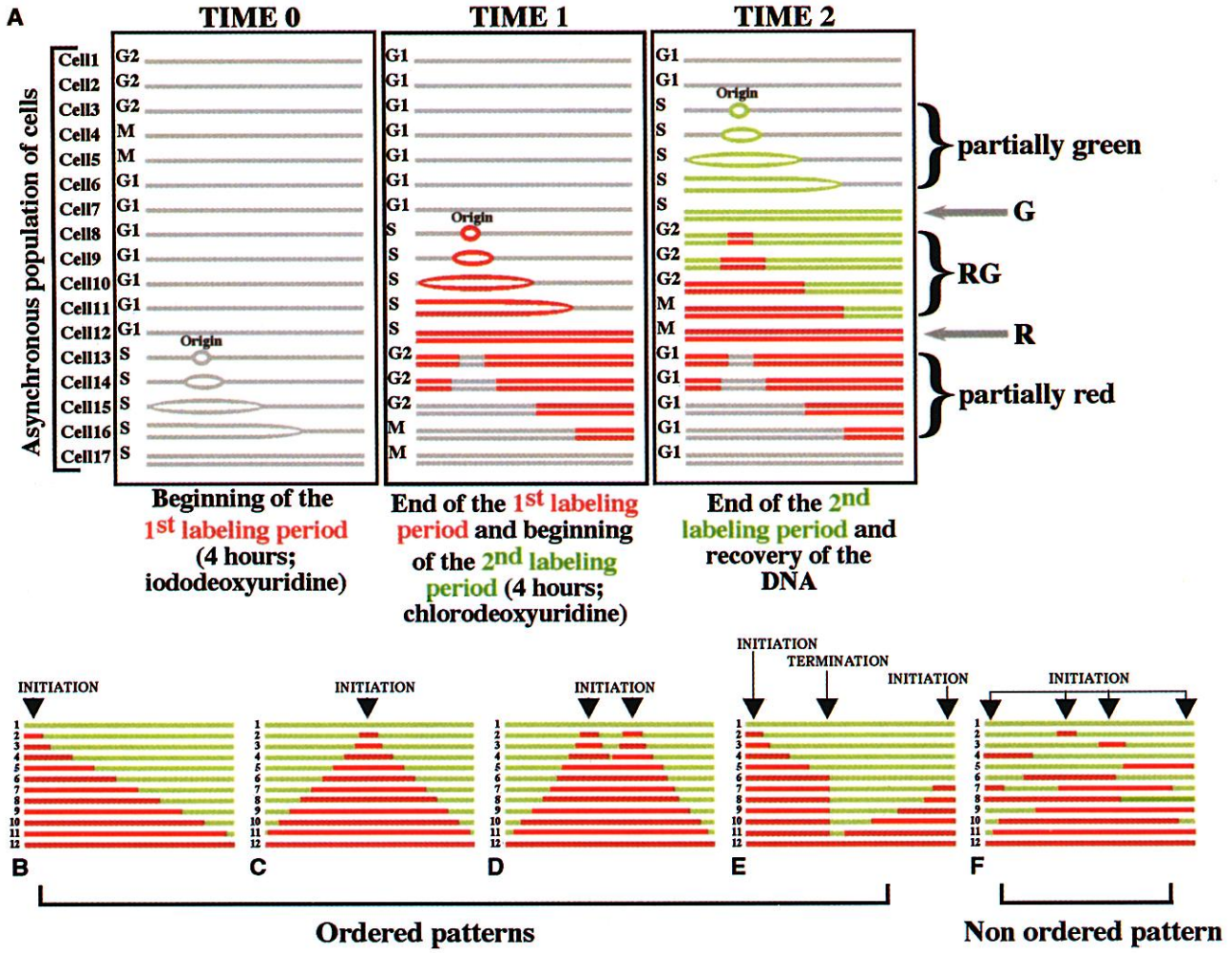

Fig. 1. (**A** to **F**) Double labeling of the replicating DNA. See text for description.

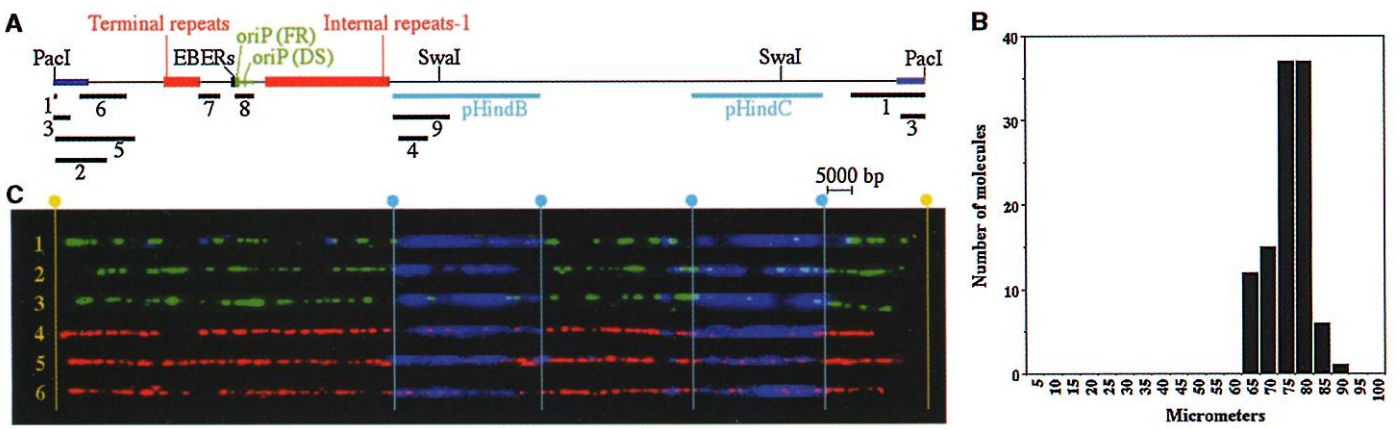

Fig. 2. Fluorescent hybridization immunostaining of individual DNA molecules. (**A**) Schematic of the Pac I–linearized EBV genome from Raji cells with the positions of various genetic elements shown to scale. Black bars indicate the restriction segments analyzed by 2D gel analysis (*4*). Cyan bars show the positions of the hybridization probes used to identify the molecules of interest and their orientation. For further details see supplementary fig. 1 (*5*). (**B**) Length of 109 combed EBV episomes (unbroken) from Raji cells. The most represented size of the molecules is 77 μm, corresponding to about 2.2 kb/μm. (**C**) Images of six Pac I–linearized episomes (representative of the 143 episomes analyzed for the studies described in the text), aligned with the EBV map, after hybridization and immunostaining of the DNA molecules and digital adjustment of length. Vertical lines indicate the positions of the ends of the hybridization probes shown in (A) (blue) and the position of the Pac I site used to linearize the EBV episomes (yellow). The hybridization signals are shown in blue. Cross-hybridization (near pHindC) and background hybridization are also visible (blue spots along the molecules). Immunostaining to detect the halogenated nucleotides is shown by the red and green colors. All the molecules shown in this figure represent EBV episomes duplicated during either the first labeling period (red) or the second labeling period (green). The resolution of analysis is limited to about 5 kb primarily as a result of discontinuities in the fluorescent signals as previously reported for similar assays (*16*, *24*, *25*).

dihydrofolate reductase (DHFR) locus in Chinese hamster cell lines (*13*). This conclusion is fully compatible with data previously obtained by 2D gel analysis on the same portion of the EBV genome [supplementary fig. 1 (*5*)] and was confirmed in an independent SMARD experiment [supplementary fig. 3 (*12*)].

Previous 2D gel analyses revealed the accumulation of replication intermediates (as a consequence of replication fork pausing) at four different sites within a 2.5-kb region including oriP and the EBER genes (*4*, *14*, *15*). By comparing the intensities of the stalled replication intermediates in the 2D gel blots, it was possible to infer that the stronger pausing occurs at the family of repeats (FR) where termination of DNA replication can also take place (*4*). However, 2D gel analysis cannot be reliably used to measure the magnitude of the replication fork pausing in mammalian cells. In the absence of pausing, replication forks detected in the population of molecules shown in Fig. 3 would be homogeneously distributed throughout the EBV genome. Instead, more than half of these molecules, at the time of the switch from the first to the second labeling period, had rightward-moving replication forks within a region of a few kilobases approximately centered on the FR element (see the position of the yellow arrows near the green vertical line in Fig. 3B). This result is in agreement with the direction of fork movement suggested by the 2D gel analysis through oriP and indicates that rightward-moving replication forks pause at these four sites during a large part of the time required to replicate the entire EBV genome. The procedure used to calculate the duplication time for the whole EBV genome can also be used, with minor modifications, to calculate the average time required to replicate specific genomic segments [supplementary fig. 4A (*5*)]. We calculated that the time required to duplicate a region of 10 kb, including the four known pausing sites (Fig. 3B, bottom), is about 36 min [supplementary fig. 4B (*5*)]. As a comparison, a different portion of the EBV genome (55 kb in size; Fig. 3B, bottom), where the accumulation of replication forks is not evident, is replicated in about 42 min [supplementary fig. 4B (*5*)]. Because both regions appear to be replicated by replication forks moving predominantly in a single direction, we estimate that the average speed of the replication forks within the 55-kb region is 1.3 kb/min. The time required to replicate a region of 10 kb at 1.3 kb/min is about 8 min; therefore, on the average, replication forks may pause at the four pausing sites for a total time of about 28 min. This suggests that pausing sites can play a major role in organizing the replication of the viral genome.

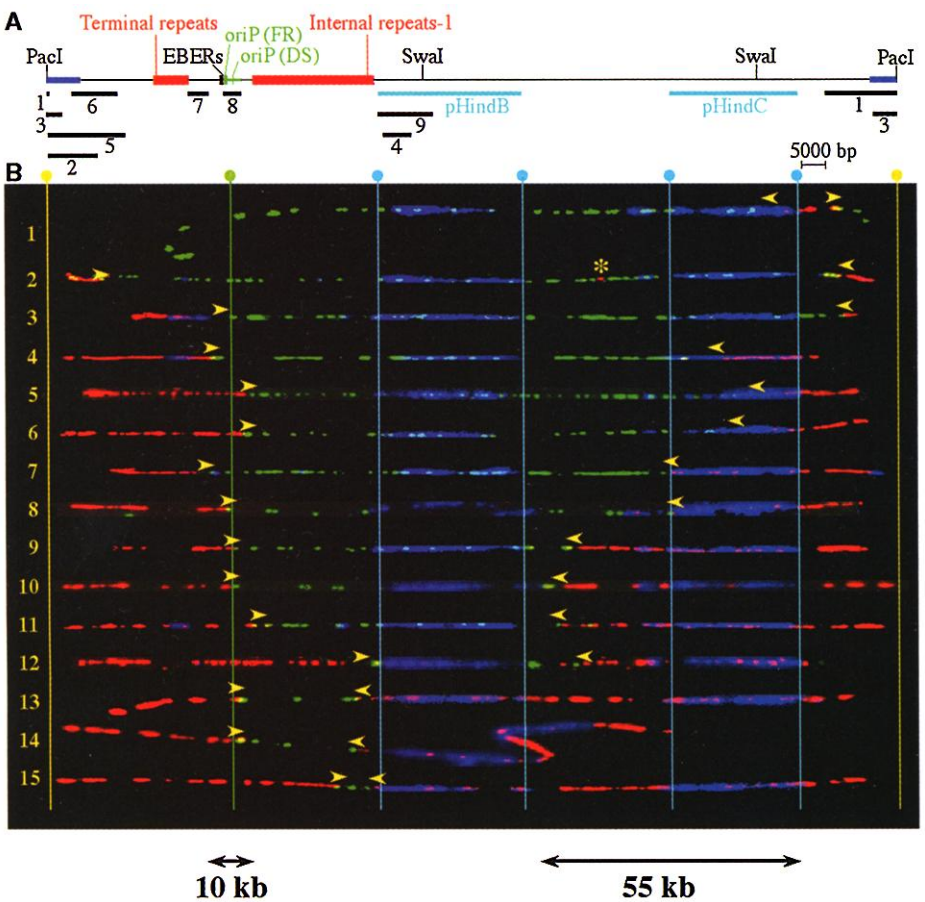

Fig. 3. SMARD of the Pac I–linearized EBV episomes. (**A**) Map of the EBV genome (see Fig. 2A for details). (**B**) Images of 15 EBV episomes containing both halogenated nucleotides and ordered (1 to 15) for increasing content of DNA labeled during the first labeling period. Arrowheads indicate the approximate position of the replication forks at the time of the switch from the first to the second labeling period. Individual dots, such as the one indicated by the yellow asterisk, were not considered in our analysis of the data because they are too short to be unequivocally distinguished from possible background immunostaining. However, initiation of DNA replication at multiple sites, on the same molecule, remains a possibility. The green line indicates the approximate position of oriP. Other vertical lines are as in Fig. 2C.

In each of the viral episomes shown in Fig. 3, a termination event must have occurred within the portion of genome stained in green. In some of the molecules that, at the time of the switch from the first to the second labeling period, were about to complete replication (molecules 12 to 15), we see that the leftward-moving forks appear to have traveled along the whole EBV genome to finally meet the forks stalled in the opposite direction at oriP (molecules 13 and 14). In other molecules, the rightward-moving forks appear to have overcome the block at oriP (or replication initiated at oriP) (*4*) and traveled through part of the central portion of the molecules to meet the leftward-moving forks at different positions within the viral genome (molecules 12 and 15). These data indicate that the duplication of the episomes can terminate throughout a large part of the viral genome, probably with a slight preference for oriP due to the pausing of the replication forks at this site. This excludes the presence of a specialized site for terminating the EBV genome duplication and indicates that even if the replication of the viral genome proceeds according to a precise scheme, this scheme is not rigid.

Studies of DNA fibers from human genomic DNA, after cell synchronization, suggested that groups of replicons can be activated in two subsequent S phases with very high efficiency (~95%) (*16*). However, these results were in marked contrast to the frequently inefficient origin firing suggested by 2D gel studies in mammalian cells (*17*). At least in the case of the EBV genome, our results resolve this apparent contradiction. We show that the viral duplication can start in a specific portion of the viral genome in all the molecules analyzed, even if the frequency of activation of individual initiation sites, as detected by 2D gels, is very low (*4*). Our results also indicate that the replication patterns are dictated mainly by the position of the initiation region and that pausing sites can influence these patterns much more than pre-

viously expected (*18*, *19*). This also suggests that, in mammalian cells, initiation zones rather than individual origins may represent the functional units regulating DNA replication at the locus level.

Compared with other studies on DNA fibers (*16*, *20–23*), SMARD can provide qualitative and quantitative information about many aspects of DNA replication [supplementary table 1 (*5*)] and is potentially applicable to the study of endogenous loci in the mammalian genome.

References and Notes

1. J. L. Yates, in *DNA Replication in Eukaryotic Cells*, M. L. DePamphilis, Ed. (Cold Spring Harbor Laboratory, Plainview, NY, 1996), p. 751.
2. E. Kieff, in *Fields Virology*, B. N. Fields *et al.*, Eds. (Lippincott-Raven, Philadelphia, 1996), p. 2343.
3. J. L. Yates, N. Guan, *J. Virol.* **65**, 483 (1991).
4. R. D. Little, C. L Schildkraut, *Mol. Cell. Biol.* **15**, 2893 (1995).
5. Supplementary text (method), figures, and table are available on *Science* Online at www.sciencemag.org/cgi/content/full/294/5550/2361/DC1.
6. A. Bensimon, A. Simon, A. Chiffaudel, V. Croquette, F. Heslot, D. Bensimon *Science* **265**, 2096 (1994).
7. X. Michalet *et al.*, *Science* **277**, 1518 (1997).
8. J. F. Allemand, D. Bensimon, L. Jullien, A. Bensimon, V. Croquette, *Biophys. J.* **73**, 2064 (1997).
9. J. A. Aten, P. J. Bakker, J. Stap, G. A. Boschman, C. H. Veenhof, *Histochem. J.* **24**, 251 (1992).
10. J. W. Vaandrager *et al.*, *Blood* **88**, 1177 (1996).
11. R. J. Florijn *et al.*, *Hum. Mol. Genet.* **4**, 831 (1995).
12. P. Norio, C. L. Schildkraut, unpublished data.
13. J. P. Vaughn, P. A. Dijkwel, J. L. Hamlin, *Cell* **61**, 1075 (1990).
14. T. A. Gahn, C. L. Schildkraut, *Cell* **58**, 527 (1989).
15. V. Dhar, C. L. Schildkraut, *Mol. Cell. Biol.* **11**, 6268 (1991).
16. D. A. Jackson, A. Pombo, *J. Cell Biol.* **140**, 1285 (1998).
17. R. F. Kalejta, H. B. Lin, P. A. Dijkwel, J. L. Hamlin, *Mol. Cell. Biol.* **16**, 4923 (1996).
18. S. A. Greenfeder, C. S. Newlon, *Mol. Cell. Biol.* **12**, 4056 (1992).
19. A. M. Deshpande, C. S. Newlon, *Science* **272**, 1030 (1996).
20. J. A. Huberman, A. D. Riggs, *Proc. Natl. Acad. Sci. U.S.A.* **55**, 599 (1966).
21. J. A. Huberman, A. D. Riggs, *J. Mol. Biol.* **32**, 327 (1968).
22. J. Herrick, P. Stanislawski, O. Hyrien, A. Bensimon, *J. Mol. Biol.* **300**, 1133 (2000).
23. J. J. Blow, P. J. Gillespie, D. Francis, D. A. Jackson, *J. Cell Biol.* **152**, 15 (2001).
24. F. M. van de Rijke, R. J. Florijn, H. J. Tanke, A. K. Raap, *J. Histochem. Cytochem.* **48**, 743 (2000).
25. I. Parra, B. Windle, *Nature Genet.* **5**, 17 (1993).

26. We thank D. Dimitrova, R. Florijn, D. Schwartz, and B. Windle for advice during the development of SMARD; F. Camia and G. Mazzeo for advice on the analysis of the data; E. Bouhassira, J. Greally, J. A. Huberman, and J. Yates for helpful discussions and critical reading of the manuscript; and S. M. Shenoy and the Albert Einstein College of Medicine analytical imaging facility for technical support. This work was supported by NIH grants GM45751. Support was also provided by Cancer Center Support Grant NIH/NCI P30CA13330.

20 July 2001; accepted 6 November 2001

Systematic Genetic Analysis with Ordered Arrays of Yeast Deletion Mutants

Amy Hin Yan Tong,[1,2] Marie Evangelista,[3] Ainslie B. Parsons,[1,2] Hong Xu,[1,2] Gary D. Bader,[4,5] Nicholas Pagé,[6] Mark Robinson,[1] Sasan Raghibizadeh,[7] Christopher W. V. Hogue,[4,5] Howard Bussey,[6] Brenda Andrews,[2]* Mike Tyers,[2,5]* Charles Boone[1,2,3]*

In *Saccharomyces cerevisiae*, more than 80% of the ~6200 predicted genes are nonessential, implying that the genome is buffered from the phenotypic consequences of genetic perturbation. To evaluate function, we developed a method for systematic construction of double mutants, termed synthetic genetic array (SGA) analysis, in which a query mutation is crossed to an array of ~4700 deletion mutants. Inviable double-mutant meiotic progeny identify functional relationships between genes. SGA analysis of genes with roles in cytoskeletal organization (*BNI1*, *ARP2*, *ARC40*, *BIM1*), DNA synthesis and repair (*SGS1*, *RAD27*), or uncharacterized functions (*BBC1*, *NBP2*) generated a network of 291 interactions among 204 genes. Systematic application of this approach should produce a global map of gene function.

For *S. cerevisiae* deletion mutations have been constructed for all ~6200 known or suspected genes, identifying ~1100 essential yeast genes and resulting in 5100 viable haploid gene-deletion mutants, with over 30% of the genes remaining functionally unclassified (*1*). These findings highlight the capacity of yeast cells to tolerate deletion of a substantial number of individual genes, perhaps reflecting the molecular mechanisms that evolved to buffer the phenotypic consequences of genetic variation (*2*, *3*). Due to the high degree of genetic redundancy in yeast, the functions of thousands of yeast genes remain obscure.

Redundant functions can often be uncovered by synthetic genetic interactions, usually identified when a specific mutant is screened for second-site mutations that either suppress or enhance the original phenotype. In particular, two genes show a "synthetic lethal" interaction if the combination of two mutations, neither by itself lethal, causes cell death (*4*, *5*). Synthetic lethal relationships may occur for genes acting in a single biochemical pathway or for genes within two distinct pathways if one process functionally compensates for or buffers the defects in the other (*2*). Synthetic lethal screens have been used to identify genes involved in cell polarity, secretion, DNA repair, and numerous other processes (*6–9*). Despite the utility of this approach, just one or two different interactions are typically identified in a single screen (*2*). Saturation is rarely achieved presumably because some genes are refractory to forward mutagenesis and because subsequent analysis of synthetic lethal mutations and gene cloning is limited by practical constraints.

To enable high-throughput synthetic lethal analysis, we assembled an ordered array of ~4700 viable yeast gene-deletion mutants (*1*) and developed a series of pinning procedures in which mating and meiotic recombination are used to generate haploid double mutants (Fig. 1). A query mutation is first introduced into a haploid starting strain, of mating type *MAT*α, and then crossed to the array of gene-deletion mutants of the opposite mating type, *MAT***a**. Sporulation of resultant diploid cells leads to the formation of double-mutant meiotic progeny. The *MAT*α starting strain carries a reporter, *MFA1pr-HIS3*, that is only expressed in *MAT***a** cells and allows for germination of *MAT***a** meiotic progeny (*10*), which ensures that carryover of the diploid parental strain and/or conjugation of meiotic progeny does not give rise to false-negative interactions. Both the query mutation and the gene-deletion mutations were linked to dominant selectable markers to allow for selection of double mutants. Final pinning results in an ordered array of

[1]Banting and Best Department of Medical Research, University of Toronto, Toronto ON, Canada M5G 1L6. [2]Department of Medical Genetics and Microbiology, University of Toronto, Toronto ON, Canada M5S 1A8. [3]Biology Department, Queens University, Kingston ON, Canada K7L 3N6. [4]Department of Biochemistry, University of Toronto, Toronto ON, Canada M5S 1A8. [5]Program in Molecular Biology and Cancer, Samuel Lunenfeld Research Institute, Mt. Sinai Hospital, Toronto ON, Canada M5G 1X5. [6]Department of Biology, McGill University, Montreal PQ, Canada H3A 1B1. [7]Virtek Engineering Sciences, Inc. (VESI), 1 Bedford Road, Toronto ON, Canada M5R 2J7.

*To whom correspondence should be addressed. E-mail: brenda.andrews@utoronto.ca, tyers@mshri.on.ca, and charlie.boone@utoronto.ca