Volume 13 Number 23 1985 Nucleic Acids Research

A transcription map of a yeast centromere plasmid: unexpected transcripts and altered gene expression

Gregory T.Marczynski* and Judith A.Jaehning*

Department of Biochemistry, University of Illinois, Urbana, IL 61801, USA

Received 12 July 1985; Revised and Accepted 13 November 1985

ABSTRACT

YCp19 is a yeast centromere plasmid capable of autonomous replication in both yeast and E. coli (J. Mol. Biol., 158: 157-179, 1982). It is stably maintained as a single copy in the yeast cell and is therefore a model yeast "minichromosome" and cloning vector. We have located the positions and measured the abundance of the in vivo yeast transcripts from YCp19. Transcripts from the selectable marker genes TRP1 and URA3 were present at increased levels relative to chromosomal copies of the genes. Unanticipated transcripts from the yeast CEN4 and E. coli pBR322 sequences were also found. Although much of the plasmid vector is actively transcribed in vivo, the regions around the most useful cloning sites (BamHI, EcoRI, SalI) are free of transcripts. We have analyzed transcription of BamHI inserts containing promoter variants of the HIS3 gene and determined that although initiation events are accurate, plasmid context may alter levels of gene expression.

INTRODUCTION

S. cerevisiae (yeast) shuttle vectors have proven to be valuable tools for studying many biological functions (1). Yeast centromere plasmids (YCp) are specialized shuttle vectors containing centromere DNA sequences which essentially transform these plasmids into "minichromosomes"; the centromere DNA sequences confer mitotic and meiotic stability and limit the plasmid copy number to approximately one to two per cell (2). YCp19 is a typical circular YCp containing a centromere (CEN4), yeast selectable marker genes (TRP1 and URA3), yeast autonomously replicating sequences (ARS1), and pBR322 sequences for ampicillin selection and autonomous replication in E. coli (2). The precursor plasmids for YCp19 (YIp5, YRp7, and YRp17) are among the most commonly used yeast shuttle vectors (1). YCp19 has been used to study factors determining mitotic and meiotic stability (3), regulatory gene dosage effects (4) and regulated extrachromosomal gene transcription (5). Derivatives of YCp19 have been used to isolate revertants of RNA polymerase II promoter mutants (6) and to study the mechanism of Ty element transposition (7).

To extend the usefulness of YCp19 (and other related yeast vectors) we

have used Northern and S1 hybridization techniques to compile the first transcription map of a yeast shuttle vector. This work has revealed several unexpected features of gene expression from a single copy plasmid.
Transcripts were found originating within the bacterial vector sequences and the expression of yeast genes on the plasmid (TRP1, URA3 and a HIS3 gene insert) was elevated three to five fold relative to levels of the same genes within the chromosome. In addition to guiding future in vitro transcription experiments, the data presented in this paper will aid investigators concerned about unexpected transcription from yeast plasmid vector sequences and the effects of plasmid context on gene expression.

MATERIALS AND METHODS

S. cerevisiae Strains and Plasmids

S. cerevisiae strains SC3 (MAT α, his3 Δ1, ura3-52, trp1-289, gal2, gal10) and SX1-2, its isogenic HIS3$^+$ parent, were obtained from R. Davis (Stanford University, Stanford, California) (8), YNN218 (MATα , his3 Δ 200, ura3-52, ade2-101, lys2-801) was obtained from M. Fasullo (Stanford University, Stanford, California), and YNN22 (20B-12, MAT α , trp1, pep 4-3) was from E. Jones (Carnegie-Mellon University, Pittsburgh, Pennsylvania) (9). S. cerevisiae plasmids YCp19 and YRp17 were obtained from R. Davis (Stanford University, Stanford, California) (2), YARp1 (TRP1 RI circle) was obtained from J. Scott (University of Illinois, Urbana, Illinois) (10). Plasmids were prepared in large quantities by two bandings in CsCl gradients (11) or in small quantities by alkaline lysis minipreps (12). YCp19-HIS3 gene recombinant plasmids, YCp19-HIS3Δ 26 and YCp19-his3 Δ38 were constructed using sequences derived from plasmids YRp14-Sc3121 (HIS3 Δ26) and YRp14-Sc2884 (his3 Δ 38) which were obtained from K. Struhl (Harvard University, Boston, Massachusetts) (13). The 1.7 kb BamHI fragments containing the HIS3 gene promoter variants were ligated into the BamHI site of YCp19 (11).

S. cerevisiae Transformation and Growth Conditions

S. cerevisiae cells were transformed by the lithium acetate protocol (14) using plasmids isolated as above. Media and plates for routine maintenance and selective growth have been described (5). Yeast transformants were selected and grown on synthetic complete (SC) media lacking uracil (SC-Ura). Control untransformed cultures were grown in SC medium, but under otherwise identical conditions. Typically 1 liter log phase cultures were grown to cell densities of 1-3 x 10^7 cells/ml and harvested by centrifugation. Cells were washed once in cold distilled water, and cell pellets were frozen for future

nucleic acid isolation. At harvest the percentage of plasmid containing cells and plasmid mitotic stability (CEN4 function) were determined by comparing selective versus nonselective cell growth as described previously (5).

S. cerevisiae Enzyme Assays

Orotidine-5'-phosphate decarboxylase (OMP decase) activity was determined by a spectrophotometric assay as described by Chevallier et al. (15) from the same cells used to prepare nucleic acids. ß -lactamase assays, using the iodine-starch overlay plate technique modified for S. cerevisiae, were done exactly as described by Chevallier and Aigle (16).

Preparation of Nucleic Acids

Total nucleic acid was prepared from S. cerevisiae cells by grinding with glass beads as described (17) except that the process was sometimes scaled up 4 to 10 fold. The yields were 20 to 40 µg of nucleic acid per 10^8 cells. Poly (A^+) RNA was selected by batchwise adsorption of total nucleic acids to oligo (dT)-cellulose Type 2 resin (Collaborative Research, Inc.) using a 5-fold excess binding capacity as specified by the manufacturer and assuming that total yeast nucleic acid contains 2% poly(A^+) mRNA.

Hybridization of DNA and RNA

Methods for treating the nucleic acid samples, separation by denaturing agarose gel electrophoresis, transfer to membrane supports, [^{32}P] labeling DNA fragments by nick-translation, and hydridization with nick-translated double strand DNA probes have previously been described in detail (5,18). Northern hybridization with single strand DNA probes and the preparation of [^{32}P] labeled single strand DNA homologous to the (+) strand of M13 mp8 was as described by Church and Gilbert (19). Individual membrane panels were probed by a two-step hybridization protocol; 10 µg of unlabeled single strand recombinant M13 phage were first hybridized to the unlabeled RNA bound to individual membrane panels (Zeta-Probe,Bio-Rad), then all panels were hybridized together with [^{32}P] labeled single strand DNA (10^8cpm/µ g) complementary to only the M13 phage sequences. To ensure that equivalent amounts of nucleic acid were analyzed, agarose gels were stained and photographed with acridine orange (20) following electrophoresis in addition to measuring absorbance at 260 nm and 280 nm. To accurately size transcripts, several types of size standards were used including: the endogenous yeast rRNAs (21); endogenous yeast TY1 (17), HIS3 (13), GAL7 (5), and GAL10 (5) mRNA's; exogenous E. coli and rat liver rRNA's (22); and DNA from plasmid YRp17 cut with EcoRI and PstI (2). These size standards were electrophoresed in parallel lanes on the agarose gels, and visualized by acridine orange

staining or by hybridization with nick-translated probes (as above). The sizing errors with these standards are 5 to 10%.

The copy numbers of YCp19 and YCp19-HIS3 recombinant plasmids were determined by quantitative genomic DNA transfer and hybridization experiments (Southern blots) as described (5) with the following specific changes: All DNA samples were prepared from total yeast nucleic acid (50 µg per lane) and digested with excess EcoRI; the membranes were probed with a [^{32}P] labeled nick-translated CEN4 DNA fragment (Sc4137) (2) prepared and isolated from YCp19 as described (5).

M13 Recombinant Phage Construction and [^{32}P] Labeling

M13 recombinant phage were constructed, screened, and maintained using standard protocols developed for dideoxy sequencing (23). Replicative double strand DNA forms of M13 cloning vectors mp8, mp9, mp10, and mp11 (23) were from BRL. The appropriate subcloning strategy for each case can be deduced from Figs. 1 and 4; the indicated restriction sites were ligated into the complementary sites in the M13 polylinker. In each case CsCl gradient purified plasmids were digested with the appropriate restriction enzymes and ligated directly from low melting temperature agarose (BRL, 24). Plasmids passaged through JM103 were used as the DNA source for subclones containing pBR322 sequences to avoid this host's restriction endonuclease system (25).

Recombinant M13 phage were uniformly labeled with [^{32}P] orthophosphate during infection of log phase E. coli JM103 cells grown in low phosphate (0.2mM) synthetic medium (26) supplemented with 10 µg/ml thiamine. A fresh overnight low phosphate JM103 culture was diluted 1:100 into fresh medium, allowed to enter log phase by shaking at 37°C for 1 hour (2 x 10^7 cells/ml) and 0.1 - 0.5 mCi/ml of [^{32}P] carrier free orthophosphate (NEN) was added simultaneously with 3 x 10^8 to 1 x 10^9 plaque forming units of recombinant phage per ml of culture. After five hours incubation, labeled cells were spun down in a microfuge and discarded. The phage containing supernatant was heated to 65°C for 10 min and was stored at 4°C for later use. Under these labeling conditions, 5 to 10% of the input [^{32}P] orthophosphate is incorporated into phage DNA, and 30 to 40 µg/ml of phage DNA is produced with a specific activity ranging from 0.5 to 5.0 x 10^6 cpm/µg depending on the amount of input [^{32}P].

S1 Nuclease Mapping with Labeled Phage

Phage particles were prepared for use as single strand DNA probes by two phenol/$CHCl_3$ extractions followed by two precipitations with 2M ammonium acetate and 2 volumes of isopropanol. Typically, 100 to 200 ng of [^{32}P]

labeled phage were coprecipitated with 10 to 100 µg of yeast total nucleic acid. The mixture was then hybridized as described by Gaynor and Berk (27) and digested with S1 nuclease (Boehringer Mannheim) as described by Gilman and Chamberlin (28). The digestion products were precipitated as described above with an extra 20 µg E. coli carrier tRNA, resuspended in 10 µl gel loading buffer (96% formamide, 10 mM EDTA, 0.1% xylene cyanol, 0.05% bromophenol blue), heat denatured at 90°C for 2 min, and loaded onto a 4% polyacrylamide/7M urea gel (29) cast on Gel Bond backing (FMC Corp.). After electrophoresis the gels were soaked in 5% glycerol, air dryed, and exposed to X-ray film (Kodak, XAR-5) with an intensifying screen (Dupont, Cronex Lightning Plus). To generate uniformly labeled single strand DNA size standards, unlabeled phage with defined strands complementary to the [^{32}P] labeled phage were hybridized and processed as above. The sizing errors using these standards were approximately 5%.

RESULTS

Strategies for Deriving a Transcription Map of YCp19

An important parameter to establish before measuring plasmid transcript position and abundance was the copy number of the YCp19 plasmids in yeast strains SC3 and YNN218. Using the procedures described in Materials and Methods, we found that all the centromere plasmids used in our studies were present at 0.8 to 1.0 copies per cell and that they exhibited mitotic stability characteristic of yeast centromere plasmids (2). To generate a collection of transcript probes, we subcloned DNA fragments homologous to YCp19 into the single strand bacteriophage M13 as shown in Figs. 1 and 4. Each probe was given a letter designation; probes A to Z refer to the upper strand (5' to 3') of the indicated fragment, A' to Z' correspond to the lower strand (3' to 5'). These probes cover all of the DNA contained in YCp19 and overlap predicted transcript regions (TRP1 and URA3) and other functional portions of the molecule (ARS1 and CEN4, see also Fig. 4).

To determine the distribution and relative abundance of all the transcripts around YCp19, RNA was prepared from yeast cells transformed with YCp19 and analyzed in parallel with RNA prepared from untransformed cells. An RNA molecule was assumed to originate from the plasmid if it was unique to, or enhanced in, RNA prepared from transformed cells versus the control untransformed cells grown under similar conditions. Transcripts were detected by Northern analysis as described in Materials and Methods. Fig. 2 is an illustration of this type of analysis. The majority of the bands are only

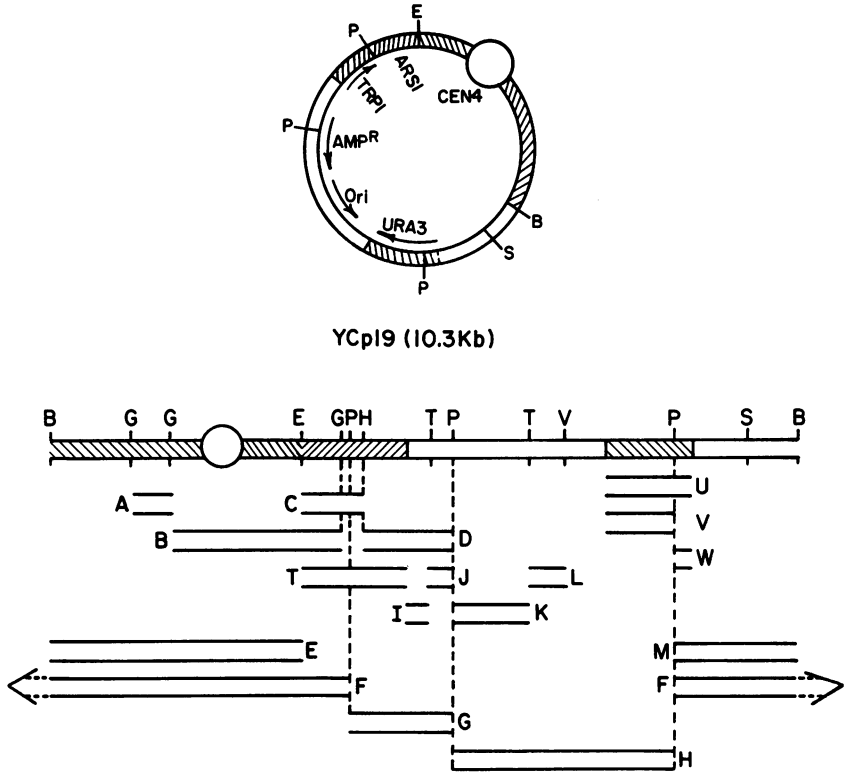

Figure 1: Structure of YCp19 and M13 recombinant phage probes. The upper portion of the figure contains a circular map of YCp19. The hatched areas indicate yeast DNA and open areas indicate bacterial (pBR322) DNA. The small, open circle represents the centromere (CEN4) (38). The arrows indicate the anticipated transcripts from yeast and E. coli genes based on sequence data for TRP1 (34), URA3 (31), Amp^R and Ori (33). The lower portion of the figure contains a linearized drawing of YCp19. The letters above the line drawing refer to selected restriction enzyme sites: B = BamHI; G = BglII; E = EcoRI; P = PstI; H = HindIII; V = PvuII; T = TaqI; S = SalI. The restriction fragment subclones homologous to YCp19 are indicated below the line drawing; the capital letters refer to the M13 recombinant phage containing the top strand of each fragment. Primed capital letters refer to the bottom (complementary) strand. Phage series A-H are subclones of YCp19; I-L are subclones of pBR322; T derives from the TRP1 ARS1 EcoRI fragment isolated as the circular yeast plasmid YARp1 (provided by C. Long) (10); phage series U-W derive from the URA3 HindIII fragment from strain D4 (31) isolated from a YEp24 plasmid (provided by S. Baker); M derives from the PstI to BamHI fragment of YRp17 (2).

seen in the lanes that contain RNA from YCp19 transformed cells indicating that they are transcribed from YCp19 DNA sequences. These transcripts are strand specific; each transcript hybridizes to only one orientation of the recombinant M13 probes.

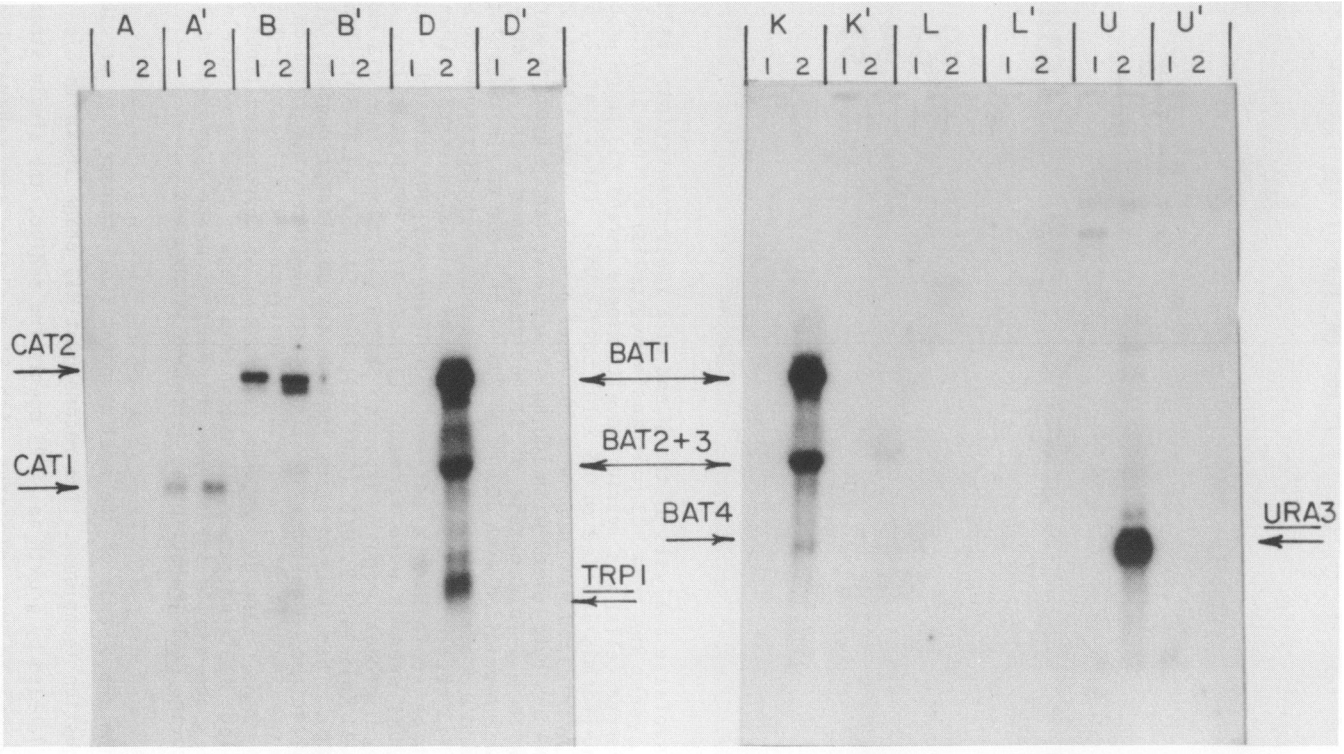

Figure 2: Northern analysis of YCp19 transcripts. Yeast RNA from untransformed and from YCp19 transformed cells were anayzed in parallel lanes. Lane 1, SC3 untransformed cells; lane 2, SC3/YCp19 plasmid transformed cells. Total yeast nucleic acid preparations (10 µg per lane) were separated by denaturing agarose gel electrophoresis, and transferred to cationized nylon membranes. The letters above each panel correspond to the recombinant phage probes shown in Fig. 1. Additional details are provided in Materials and Methods.

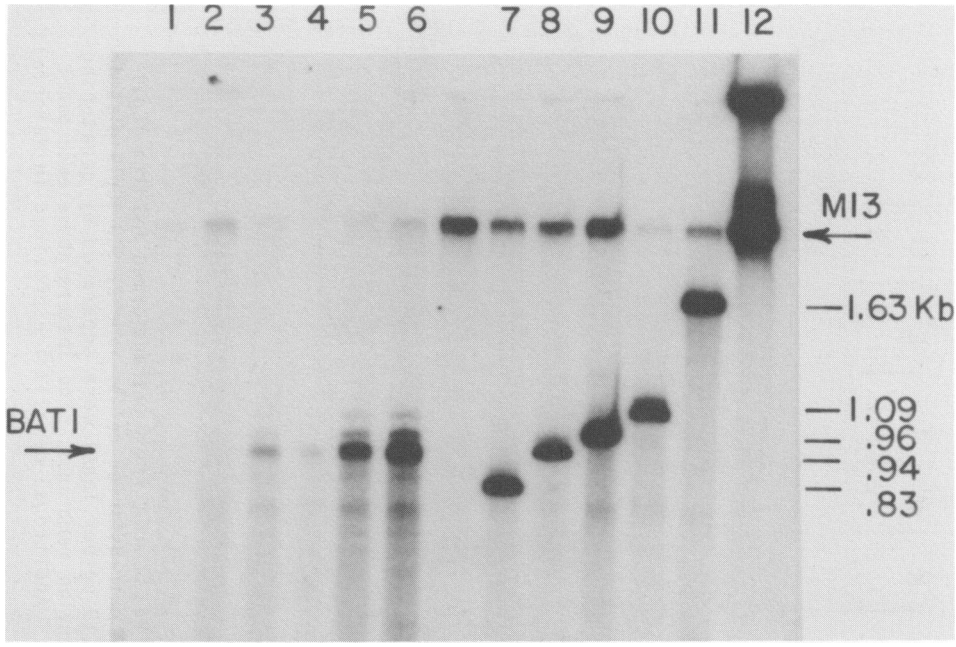

Figure 3: S1 nuclease analysis of the 5' end of a YCp19 transcript: BAT1. In lanes 1 to 6, yeast RNA from SC3 (untransformed cells) or from SC3/YCp19 (plasmid transformed cells) was hybridized with single strand M13 phage probe K (see Fig. 1) uniformly labeled to 0.5 x 10^6 cpm/µg DNA: Lanes 1 and 2, 60 µg SC3 RNA was hybridized with 100 ng or 300 ng of probe K; lanes 3 and 4, 20 µg SC3/YCp19 RNA was hybridized with 100 ng of probe K, lane 4 hybrids were overdigested with a 5-fold excess of S1 nuclease; lanes 5 and 6, 60 µg SC3/YCp19 N.A. was hybridized with 100 ng or 300 ng of probe K. In lanes 7 to 11, single strand DNA size standards were generated by hybridizing 60 ng of [^{32}P] labeled M13 recombinant phage with 2 µg of unlabeled single strand recombinant phage containing a defined complementary DNA strand. The following DNA/DNA hybrids were formed: Lane 7, G/T' (0.83 Kb); lane 8, GAL1/GAL10 gene divergent promoter fragment Sc4816 (49) cloned into M13 mp8 and M13 mp9 (0.94 Kb); lane 9, U/U' (0.96 Kb); lane 10, K/K' (1.09 Kb); lane 11, G/G' (1.63 Kb). These hybrid sizes include the polylinker region. All samples were hybridized, digested with S1 nuclease, denatured, separated on a 4% polyacrylamide 7M urea gel, and otherwise processed as described in Materials and Methods. Lane 12 contains about 20 ng of unprocessed probe K. The arrow labeled "M13" indicates the undigested phage. The arrow labeled "BAT1" indicates the single strand DNA fragment of probe K protected by hybridization to the BAT1 transcript as described in the text.

To determine the 5' ends of the most abundant YCp19 transcripts, we used S1 nuclease digestion techniques. The recombinant M13 phage shown in Fig. 1 were uniformly labeled by growth in [^{32}P] containing medium, the single strand DNA was hybridized to RNA, digested with S1 nuclease and separated on

denaturing polyacrylamide gels as described in Materials and Methods. Fig. 3 illustrates how this technique was used to map the 5' end of a major YCp19 transcript which initiates within bacterial sequences on the plasmid. This transcript is described further below.

Abundant Transcripts from YCp19

The URA3 gene produced the most abundant transcript from YCp19. Fig. 2 shows that the URA3 gene probe U, but not the complementary DNA strand probe U', hybridizes to a 1.1 Kb RNA, thus assigning its approximate position and orientation. This assignment is consistent with previous studies by Buckholz and Cooper (30) and Rose and Botstein (31). No URA3 mRNA of normal size is detected in untransformed cells which contain the ura3-52 mutant allele (SC3 and YNN218) but we have detected a short, 0.55 Kb ura3-52 transcript with probe U. This transcript is about 20 fold less intense than the 1.1Kb URA3 mRNA from YCp19 (data not shown). To locate the 5' end of the YCp19 URA3 transcript, we performed S1 nuclease mapping experiments similar to the example in Fig. 3, using M13 recombinant probes F and M (see Fig. 1). These probes end with a common PstI site positioned just 17 bp upstream of the URA3 gene open reading frame (32). We observed that a plasmid specific RNA protected a short band (less than 150 bp) from S1 nuclease digestion when we used either probe F or probe M; neither of the opposite strand probes yielded a protected fragment (data not shown). This positioned the YCp19 URA3 transcript 5' end within the yeast DNA sequences less than 150 bp upstream of the URA3 gene PstI site.

The next most abundant transcripts initiate from pBR322 DNA sequences. These are called BAT for Bacterial DNA Associated Transcripts. BAT1, the most abundant bacterial transcript, is approximately 2.5 Kb long; it is shown to hybridize strongly to probes K and D, and also weakly to probe B in Fig. 2. This implies that the 3' sequences of BAT1 lie within probe B, and its 5' sequences lie within probe K. The S1 nuclease analysis with probe K in Fig. 3 is consistent with this assignment for BAT1. Only RNA prepared from YCp19 transformed cells protected a subfragment of probe K which migrates as a major band 0.92 Kb long, as measured from single strand DNA size standards in the adjacent lanes. This fixes the major 5' end of BAT1 to approximately 920 ± 15 bp upstream of the pBR322 PstI site on YCp19 which corresponds to pBR322 nucleotide number 2690 ± 15 (33).

There are at least three other less abundant bacterial transcripts (BAT2, 3, and 4) which overlap BAT1 and have the same orientation (Fig. 2). In preparations of total yeast RNA BAT2 and BAT3 comigrate; their migration

suggests that they are compressed by the 18S rRNA band. However, they separate as two distinct 1.9 Kb and 1.65 Kb bands when poly(A^+) selected RNA is hybridized with probe G (data not shown). S1 nuclease techniques have not permitted precise mapping of these less abundant transcripts because BAT1 dominates transcription through this region. For example, we cannot determine whether the minor bands in Fig. 3 are extra minor start sites for BAT1 or start sites for the other minor bacterial transcripts. We observed that BAT2 and BAT3 hybridize strongly to probe D but only very weakly to probe T. This suggests that both transcripts end within about 200 bp of the TRP1/pBR322 border. Based on their lengths, this positions their 5' start sites within about 200 bp of the start site for BAT1.

TRP1 Transcripts from YCp19

The mapping of the TRP1 gene transcript from YCp19 was complicated by low level transcription, interference by transcription from upstream bacterial sequences, and the presence of chromosomal TRP1 mRNA. We have assigned the band marked TRP1 in Fig. 2, panel D to the YCp19 TRP1 gene based on the following observations: it is typically enhanced in RNA prepared from YCp19 transformed cells; it has a size (about 0.9 Kb), orientation and position consistent with the location of the TRP1 coding sequences on YCp19. This band hybridizes to probes D, T and faintly to B. This indicates that its 3' end is adjacent to the BglII site on the right side of Probe B (as oriented in Fig. 1). This site is 77 bp downstream from the end of the TRP1 open reading frame (34). The measured length of 0.9 Kb would allow this transcript to span the TRP1 open reading frame, starting within 100 bp of the TRP1/pBR322 border and ending near the BglII site.

In hybridization experiments with higher specific activity TRP1 probes, and in longer exposures of the Northern experiment in Fig. 2, the YCp19 TRP1 band and the faint upper band (about 1.1 Kb), were seen to comigrate with two chromosomal TRP1 transcripts in the adjacent lane (data not shown). The larger of the two transcripts may initiate outside of the TRP1 EcoRI fragment in accordance with previous S1 mapping experiments by Zalkin and Yanofsky (35). The smaller transcript initiates within the TRP1 EcoRI fragment on both YCp19 and the yeast chromosome, based on the following experimental observations. YARp1 (TRP1 RI Circle) consists of the circularized TRP1 ARS1 sequences present on YCp19; it is typically maintained in transformants at 100 to 200 copies per yeast cell (10). We have observed that the 0.9 Kb TRP1 RNA is overproduced in YARp1 transformants 200 to 300 fold relative to the chromosomal abundance (data not shown). S1 nuclease experiments with RNA from

YARp1 transformants have revealed that the 0.9 Kb TRP1 RNA has two start sites approximately 30 ± 10 bp and 60 ± 10 bp upstream of the TRP1 AUG start codon (data not shown). This implies that the YCp19 0.9 Kb TRP1 RNA also has two start sites, as indicated in Fig. 5.

Transcripts Associated with CEN Regions

Two transcripts that hybridize to yeast centromere DNA sequences are shown in Fig. 2 panels A' and B. The band labeled CAT1 (Centromere DNA Associated Transcript) corresponds to a transcript of 1.2 Kb and it is observed in RNA prepared from both transformed and untransformed cells. CAT1 is transcribed from the centromere sequences on YCp19 as well as from those on chromosome IV, because it is more than two fold more abundant in RNA prepared from YCp19 transformed cells as shown in Fig. 2, panel A', and as much as 10-fold more abundant in other similar hybridization experiments (data not shown). The 5' end of CAT1 maps approximately halfway between the BamHI site and the BglII site on the left side of probe A (see Fig. 1). This assignment was based on the 1.2 Kb size of CAT1, and on the observation that its 3' end hybridizes weakly to probe B' (CAT1 was visible in an overexposure of the autoradiogram in Fig. 2, panel B', data not shown). The transcript labeled CAT2 (Fig. 2, panel B) migrates just above the 2.5 Kb BAT1 transcript. CAT2 is not a YCp19 transcript since it hybridizes only to probe B, but it is too long to be fully contained by it (see also below). In addition, we have never observed enhanced levels of CAT2 RNA prepared from transformed cells. It is therefore probable that the promoter for CAT2 lies outsides of the centromere DNA sequences present on YCp19.

To further characterize transcription of the CEN4 region additional single strand probes were constructed as outlined in Fig. 4. These probes were used in combination with probes A and A' (Fig. 1) to establish whether CAT1 and CAT2 transcription included the conserved centromere region ([36], indicated in Fig. 4 as I, II, III). The Northern analysis shown in the upper part of Fig. 4 demonstrates that CAT1 hybridizes to probes A' and X' (but not to A and X) and that CAT2 hybridizes to probe Z (but not to Z'). Neither transcript hybridizes to the 477 bp Y or Y' probes which includes the CEN4 consensus region (38). The CAT1 and CAT2 transcripts correspond to two open reading frames (Fig. 4, ORF) defined by DNA sequence analysis of the region (38). Both RNAs are enriched by polyA$^+$ selection and CAT2 synthesis in isolated yeast nuclei (J. F. Jerome and J. A. Jaehning, submitted for publication) is sensitive to low levels of α-amanitin (data not shown) implying that both are mRNAs transcribed by RNA polymerase II.

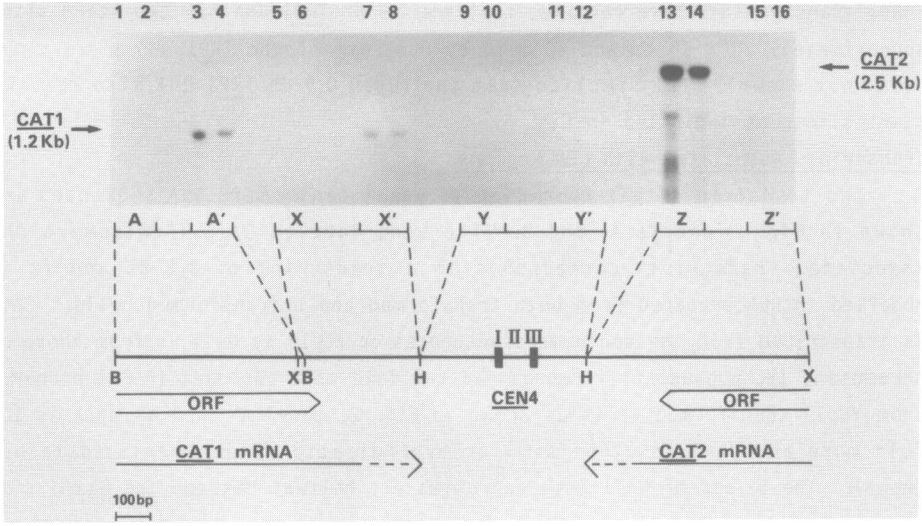

Figure 4: Transcription in the centromere (CEN4) region. Northern analysis was performed as described in the legend to Figure 2. All lanes contain 10 µg of total yeast nucleic acid prepared from SC3/YCp19 transformed cells (odd numbered lanes) or SC3 untransformed cells (even numbered lanes). Single strand M13 recombinant phage probes were constructed from the indicated restriction fragments (B = BglII, X = XhoI, H = HpaII) obtained form the CEN4 region of YCp19, and conform to the strand conventions used in all previous figures. The DNA landmarks were determined from the DNA sequence of Mann and Davis (38): Two long open reading frames (ORF) and the consensus centromere regions I, II, III (36,37) are all drawn to scale.

The YCp19 transcription map shown in Fig. 5 was compiled from the data presented above. Orientation of a transcript (5' to 3') is indicated by the direction of the arrow and abundance is reflected by the arrows thickness. Dashed lines indicate uncertainty in the position of 5' or 3' ends of the transcripts. Table I lists all of the transcripts that we have assigned to YCp19, together with their sizes, and an estimate of their relative abundance. All of these YCp19 transcripts are polyadenylated, because they are enriched by adsorption to oligo (dT) resin (data not shown).

URA3 Gene Expression from YCp19

To test whether URA3 gene expression from the YCp19 plasmid was similar to URA3 gene expression from its normal chromosomal locus, we compared URA3 enzyme activities and mRNA levels from YCp19 in a ura3-52 background to those from a URA3$^+$ untransformed strain, YNN22. The URA3 mRNA from YCp19 transformed cells comigrated on denaturing agarose gels with the URA3 mRNA from YNN22 in experiments analogous to the one shown in Fig. 2, panel U (data

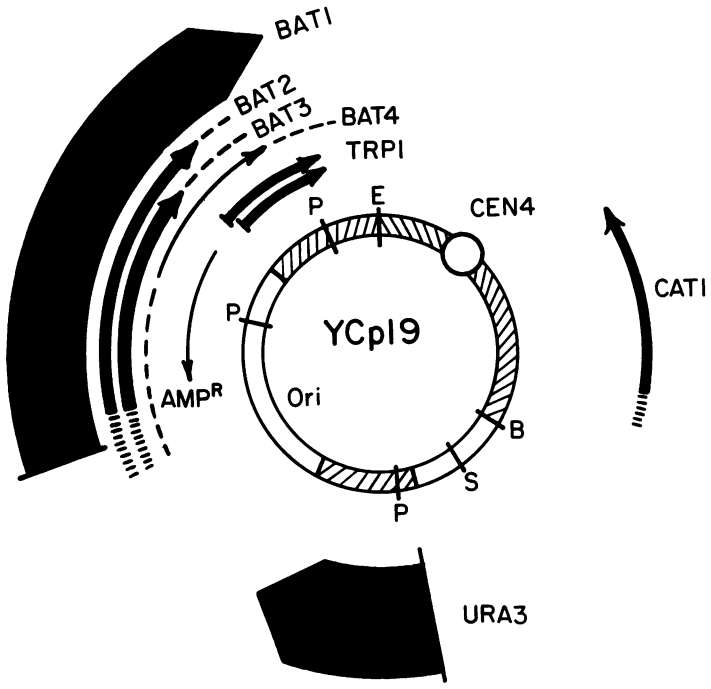

Figure 5: YCp19 yeast in vivo transcription map. This figure summarizes the transcription data obtained by Northern blotting and S1 nuclease analysis. Each of the transcripts is described in the text. The transcripts are drawn as arrows indicate their directions of transcription (5' to 3'). Their thicknesses reflect relative abundance. Dashed lines (---and ||||) refer to uncertainty in the 5' and 3' ends of a transcript.

not shown). This suggests that correct transcription initiation, termination and polyadenylation are taking place from YCp19. The URA3 gene encodes the enzyme orotidine-5'-phosphate decarboxylase (OMP decase) (39). The specific activity of OMP decase is a good measure of the fidelity of URA3 transcription. This enzyme activity is strictly under transcriptional control, since enzyme activity is closely correlated with transcription rate and mRNA abundance (15,40). The data in Table II is consistent with this correlation, and indicates that SC3, ura3-52 cells containing YCp19 coordinately produce about 3 to 4 fold more URA3 mRNA and OMP decase activity than YNN22, URA3$^+$ cells. Previous studies have shown that OMP decase activity is induced several fold by compounds or mutations which cause the accumulation of pyrimidine intermediates but it is unaffected by starvation for uracil (39). Table II shows that uracil starvation has no significant affect on the

TABLE I
YCp19 Specific Transcripts, Sizes, and Relative Abundance

Transcript Name	Size (Kb)[a]	Relative Abundance[b]
URA3	1.1	(1.0)
BAT1	2.5	0.7
BAT2	1.9	0.1
BAT3	1.65	0.1
BAT4	1.1	0.02
Amp^R	1.1	<0.01
TRP1	0.9	0.05 - 0.25
CAT1	1.2	0.05 - 0.25

[a]RNA's were fractionated by denaturing agarose gel electrophoresis, their sizes include a poly(A) tail of unknown length (see text). The size standards and protocols are detailed in Materials and Methods.

[b]Abundance was determined by densitometry of autoradiograms such as shown in Fig. 2. Densities were normalized so that the most abundant transcript, URA3, equals 1.0.

TABLE II
Coordinate Expression of URA3 Enzyme and mRNA

Strain (genotype)	Growth Media	OMP Decase Specific Activity[a]	Relative Abundance of URA3 mRNA[b]
SC3 (ura3-52)	SC	.0 ± .2	---
SC3/YCp19 (ura3-52/$URA3^+$)	SC-uracil	3.2 ± .2	4.2
YNN22 ($URA3^+$)	SC	1.0 ± .2	1.0
YNN22 ($URA3^+$)	SC-uracil	---	0.95

[a]OMP decase specific activity is expressed as nmoles OMP decarboxylated per min per mg protein according to Beckwith et al. (48).

[b]RNA was isolated from aliquots of the same cells assayed for OMP decase activity. Serial dilutions of RNA over an 8-fold range were fractionated and probed as in Fig. 2 panel U. Densitometry was performed on the 1.1 Kb URA3 band in the linear range, and the slopes of the lines (arbitrary density units versus µg RNA) were calculated by linear regression from 4 points with a correlation coefficient of 0.99. The slopes were normalized so that YNN22 grown on SC has a slope equal to 1.0.

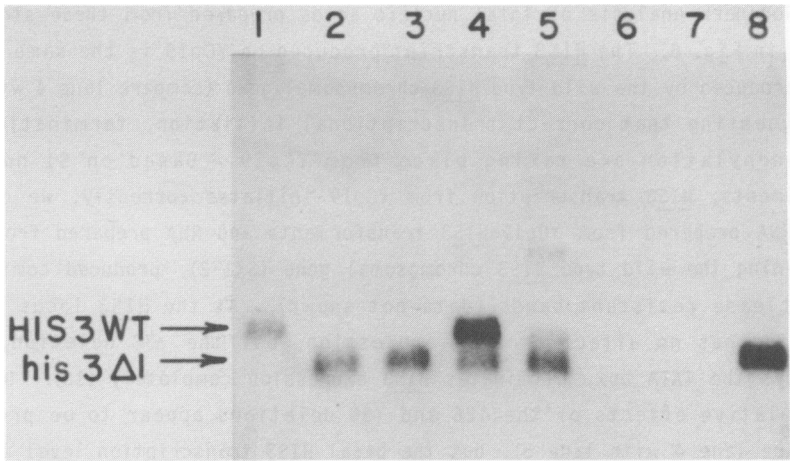

Figure 6: Northern analysis of HIS3 mRNA from chromosomal and YCp19 genes. Total nucleic acid preparations, (10 μ g per lane) from the yeast strains indicated below, were separated by denaturing agarose gel electrophoresis, transferred to a nitrocellulose membrane, and hybridized with a nick-translated 0.9 Kb DNA probe from the EcoRI linker to the XhoI sites of YRp14-Sc3121 (18) which spans the entire HIS3 gene. Lane 1, SX1-2 (HIS3^{+}); lane 2, SC3 (his3 Δ 1); lane 3, SC3/YCp19; lane 4, SC3/YCp19-HIS3Δ 26; lane 5, SC3/YCp19-his3Δ 38; lane 6, YNN218 (his3Δ 200); lane 7, of 50 μg YNN218; lane 8, YNN218/YCp19-HIS3 Δ26. Note that overloading in lane 7 distorted the RNA migration in lane 8.

mRNA levels of YNN22. This suggests that selective pressure for maintaining the YCp19 plasmid can not, by itself, account for the elevated expression from the YCp19 URA3 gene.

HIS3 Gene Expression from YCp19

We asked whether the transcription of other yeast genes would be altered when moved out of their normal chromosomal context and placed on YCp19. Struhl and coworkers have extensively studied the HIS3 gene at its normal chromosomal locus and they have constructed a number of promoter deletions (13). YCp19-HIS3 recombinant plasmids containing different HIS3 promoter deletions were constructed as outlined in Materials and Methods. We chose yeast strains SC3 and YNN218 as the recipients because they would allow us to distinguish plasmid HIS3 transcription from chromosomal HIS3 transcription. SC3 contains the his3 Δ1 structural gene deletion which produces a transcript 0.15 Kb shorter than that from the wild type gene (8, see also below); YNN218 contains the his3 Δ200 deletion which eliminates the entire HIS3 gene and transcript (6, see also below).

Northern analysis of total nucleic acids prepared from these strains is shown in Fig. 6. The HIS3 transcript produced on YCp19 is the same size as that produced by the wild type HIS3 chromosomal gene (compare lane 4 with lane 1) suggesting that correct transcriptional initiation, termination, and polyadenylation are taking place from YCp19. Based on S1 nuclease experiments, HIS3 transcription from YCp19 initiates correctly; we observed that RNA prepared from YCp19-HIS3 transformants and RNA prepared from cells containing the wild type HIS3 chromosomal gene (SX1-2), produced comigrating S1 nuclease resistant bands (data not shown). At the HIS3 locus the Δ 26 deletion has no affect on HIS3 expression, but the Δ38 deletion, which destroys the TATA box, eliminates HIS3 expression completely (13). On YCp19 the relative effects of the Δ26 and Δ38 deletions appear to be preserved (compare lane 4 with lane 5), but the basal HIS3 transcription level is about 5-fold higher as judged by densitometric scanning of the plasmid and chromosomal HIS3 bands in Fig. 6, lanes 1 and 4. Thus, it appears that RNA polymerase II uses the same HIS3 promoter sequences on YCp19, but that it uses them more efficiently.

DISCUSSION

We have found unexpected complexity in the transcription map of the yeast shuttle vector YCp19. Although yeast genes are transcribed accurately from the plasmid vector, the abundance of the transcripts is elevated significantly relative to controls where the same genes are analyzed in their normal chromosomal context. URA3 and HIS3 mRNA levels were consistently four to five fold higher from YCp19 even though copy number experiments confirmed that the plasmid was present as a single copy per cell. TRP1 mRNA and a transcript from the CEN4 region of the plasmid, CAT1, were also overproduced in a more variable fashion, in some cases RNA levels were elevated as much as ten fold. These results are in contrast with a previous study of the regulated expression of the GAL7 gene from plasmid vectors (5). Transcription of the GAL7 gene from YCp19 was found to be identical in every way to a chromosomal copy of the gene.

These results may indicate that promoter strength is a factor in determining how a gene is expressed from a plasmid. Low expression promoters such as TRP1 and CAT1 (each representing approximately 0.001% of total mRNA) may be the most sensitive to alterations in normal template structure while a high expression promoter such as GAL7 (approximately 1% of total mRNA) is insensitive to these context effects (5). Moderate promoters such as URA3 and

HIS3 (approximately 0.01% to 0.1% of total mRNA) show an intermediate effect. These changes in transcript abundance could be caused by alterations in DNA topology on the small circular plasmid molecule. There is increasing evidence that negative supercoiling of DNA is critical for transcriptional activation of eukaryotic and prokaryotic genes (41,42). If the small size of the plasmid vectors is, in fact, constraining the DNA into a superhelical density different than that found in the yeast chromosomes then this would predict that larger plasmids would more closely mimick cellular transcription patterns. It is interesting in this regard that centromere function itself has been shown to be critically dependent on plasmid size (43); a high level of mitotic stability is only observed on plasmids greater than 100 Kb in size.

In addition to identifying the expected transcripts from the known plasmid genes, we also discovered unexpected transcripts from bacterial DNA sequences. These transcripts, as well as those described above, are polyadenylated and are presumably RNA polymerase II products. The most abundant bacterial transcript, BAT1, initiates close to the pBR322 origin of replication and appears to use the same terminator as the TRP1 transcript. If its abundance is an accurate reflection of promoter strength, then the BAT1 promoter may be very active in an *in vitro* transcription assay. The transcription of pBR322 DNA sequences has also been described in human Hela cells (44). In a crude *in vitro* transcription system, Hela cell RNA polymerase II selectively initiates a transcript that appears to be analogous to the yeast BAT1 transcript. The Hela cell transcript initiates approximately 30 bp downstream from a conserved eukaryotic promoter element, the Goldberg-Hogness or "TATA" box (44), positioned at pBR322 nucleotide number 2608 (33). Our BAT1 transcript initiates approximatley 80 bp downstream from this same "TATA" box. The "TATA" box to transcription start site distance for BAT1 is more typical of yeast, where distances of 42-120 bp have been reported, than for higher eukaryotes where distances of 30-40 bp are more common (44,45).

We searched for a transcript corresponding to the pBR322 Amp^R gene on YCp19; it has been reported that 2-μm DNA recombinant plasmids in yeast produced a 1.1 Kb Amp^R transcript, and that yeast cells containing these plasmids secrete an active β-lactamase enzyme (46). None of the bacterial transcripts described above correspond to the Amp^R transcript, however, a 1.1 Kb transcript was in fact observed with the predicted location and orientation in an extensive overexposure of the appropriate panels in Fig. 2 (data not shown, see Table I). This 1.1 Kb RNA is one of 10 very minor transcripts

scattered over YCp19 which we estimate are less than 1% as abundant as the YCp19 URA3 transcript and are not included in the transcription map. We tested cells transformed with YCp19 and YRp17 for β-lactamase activity, but only YRp17 transformants contained detectable activity. Since both plasmids share the same pBR322 sequences, but differ in copy number, it seems likely that YCp19 also expresses the Amp^R gene to some extent.

The other unexpected transcripts that we detected were associated with the centromere. Constructions which induce active transcription through a centromere region have been shown to disrupt its activity (36). The yeast centromere sequences on YCp19 hybridized to two low abundance transcripts, CAT1 and CAT2. CAT1 is transcribed both on YCp19 and from its chromosomal locus. CAT2 is transcribed only at its chromosomal locus, presumably because its promoter is located outside of the CEN4 DNA fragment on YCp19. CAT2 may be the 2.5 Kb transcript that was detected near the centromere in previous studies (2). The two transcripts correspond to open reading frames (38) and converge towards the functional centromere region but do not include sequences in the conserved centromere consensus (36,37). Apparently although high level transcription through a centromere disrupts its function (36), low level transcription near the centromere region is not harmful and occurs normally.

Our transcription data on YCp19 will aid us in interpreting yeast RNA polymerase II *in vitro* transcription experiments. It may also aid other investigators interested in using centromere plasmids as single copy vectors. Transcription of a cloned gene will most likely be driven by its own promoter because none of the YCp19 transcripts that we have detected run through the unique EcoRI, BamHI, and SalI cloning sites. However, one should also consider that yeast promoters may behave unexpectedly on centromere plasmids and on plasmid vectors in general. The use of multi-copy cloning vectors in yeast can lead to the artifactual isolation of DNA sequences that complement mutations in nonhomologus DNA when expressed at high levels in the cell (47). Single copy vectors have been thought to avoid these problems but our studies suggest that caution must be used in the interpretation of these experiments as well.

Acknowledgements

We thank R. Davis, M. Fasullo, and E. Jones for strains, K. Struhl for the HIS3 gene and promoter variants, and D. Stinchcomb for YCp19, P. Hieter, S. Baker, C. Long, and M. Fagan for donated DNA fragments and much practical advice, S. Baker, M. Woontner, D. McKearin, and D. Shapiro for critically reading this manuscript, and J. Lewis for expert typing. This work was

supported by a grant from the National Science Foundation (PCM 83-15903) to J.A.J.. G.T.M. was a trainee of the National Institutes of Health (PHS 5 R01 GM07283).

*Present address: Department of Biology, Indiana University, Bloomington, IN 47405, USA

REFERENCES

1. Botstein, D., and Davis, R.W. (1982) in The Molecular Biology of the Yeast Saccharomyces, Metabolism and Gene Expression, (J. N. Strathern, E. W. Jones, and J. R. Broach, Eds.) Cold Spring Harbor Press, Cold Spring Harbor, New York.
2. Stinchcomb, D.T., Mann, C., and Davis, R.W. (1982) J. Mol. Biol., 158, 157-179.
3. Dani, G.M., and Zakian, V.A. (1983) Proc. Natl. Acad. Sci. USA, 80 3406-3410.
4. Hashimoto, H., Kikuchi, Y., Nogi, Y., and Fukasawa, T. (1983) Mol. Gen. Genet., 191, 31-38.
5. Baker, S.M., Okkema, P.G., and Jaehning, J.A. (1984) Molecular and Cellular Bio., 4, 2062-2071.
6. Scherer, S., Mann, C., and Davis, R.W. (1982) Nature, 298, 815-819.
7. Boeke, J.D., Garfinkel, D.J., Styles, C.A., and Fink, G.R. (1985) Cell, 40, 491-500.
8. Scherer, S., and Davis, R. (1979) Proc. Natl. Acad. Sci. USA, 76, 4951-4955.
9. Jones, E. (1977) Genetics, 85, 23-33.
10 Zakian, V.A., and Scott, J.F. (1982) Mol. Cell. Bio. 2, 221-232.
11. Davis, R.W., Botstein, D., and Roth, J.R. (1980) Advanced Bacterial Genetics, Cold Spring Harbor Press, Cold Spring Harbor, New York.
12. Birnboim, H.C., and Doly, J. (1979) Nucleic Acids Res., 7, 1513-1523.
13. Struhl, K. (1982) Proc. Natl. Acad. Sci. USA, 79, 7385-7389.
14. Ito, H., Fukuda, Y., Murata, K., and Kimura, A. (1983) J. Bacteriology, 153. 163-168.
15. Chevallier, M.R., Bloch, J.C., and Lacroute, F. (1980) Gene, 11, 11-19.
16. Chevallier, M.R., and Aigle, M. (1979) FEBS Lett., 180, 179-180.
17. Elder, R.T., Loh, E.Y., and Davis, R.W. (1983) Proc. Natl. Acad. Sci. USA, 80, 2432-2436.
18. Maniatis, T., Fritsch, E.F., and Sambrook, J. (1982) Molecular Cloning, Cold Spring Harbor Press, Cold Spring Harbor, New York.
19. Church, G., and Gilbert, W. (1984) Proc. Natl. Acad. Sci. U.S.A., 81, 1991-1995.
20. McMaster, G.K., and Carmichael, G.G. (1977) Proc. Natl. Acad. Sci. USA, 74, 4835-4838.
21. Warner, J.R. (1982) In "The Molecular Biology of the Yeat Saccharomyces, Metabolism and Gene Expression," (J. N. Strathern, E.W. Jones, and J.R. Broach, Eds.) pp. 529-560, Cold Spring Harbor Press, Cold Spring Harbor, New York.
22. Loening, U.E. (1968) J. Mol. Biol., 38, 355.
23. Messing, J. (1983) Methods in Enzymology, 11, 20-78.
24. Frischauf, A.M. Gasoff, J., and Lehrach, H. (1980) Nuc. Acids. Res., 8, 5541-5549.
25. Felton, J. (1983) BioTechnqiues, 42-43.
26. Nomura, N., and Ray, D.S. (1980) Proc. Natl. Acad. Sci. USA, 77, 6566-6570.
27. Gaynor, R.B., and Berk, A.J., (1983) Cell, 33, 683-693.

28. Gilman, M.Z., and Chamberlin, M.J. (1983) Cell, 35, 285-293.
29. Donis-Keller, M., Maxam, A.M., and Gilbert, W. (1977) Nuc. Acids Res., 4, 2527-2538.
30. Buckholz, R.G., and Cooper, T.G. (1983) Mol. Cell Biol., 3, 1889-1897.
31. Rose, M. and Botstein, D.(1983) J. Mol Biol., 170, 883-904.
32. Rose, M., Grisafi, P., and Botstein, D. (1984) Gene, 29, 113-124.
33. Sutcliffe, J.G. (1979) Cold Spring Harbor Sympo. Quant. Biol., 43, 77-90.
34. Tschumper, G., and Carbon, J. (1980) Gene, 10, 157-166.
35. Zalkin, H., and Yanofsky, C. (1982) J. Biol. Chem., 257, 1491-1500.
36. Blackburn, E.H., and Szostak, J.W. (1984) Ann. Rev. Biochem., 53, 163-194.
37. Carbon, J. (1984) Cell, 37, 351-353.
38. Mann, C., and Davis, R.W. (1985) Mol. Cell. Biol. (in press).
39. Lacroute, F. (1968) J. Bacteriology, 65, 824-832.
40. Bach, M.L., Lacroute, F., and Botstein, D. (1979) Proc. Natl. Acad. Sci. USA, 76, 386-390.
41. Ryoji, M., and Worcel, A. (1984) Cell, 37, 21-32.
42. Brahms, J.G., Dargouge, O., Brahms, S., Ohara, Y., and Vagner, V. (1985) J. Mol. Biol., 181, 455-465.
43. Hieter, P., Mann, C., Snyder, M., and Davis, R.W. (1985) Cell, 40, 381-392.
44. Sassone-Corsi, P., Kedinger, C., and Chambon, P. (1981) Nuc. Acids Res., 9, 3941-3957.
45. Guarente, L. (1984) Cell, 36, 799-800.
46. Breunig, K.D., Mackedonski, V., and Hollenberg, C.P. (1982) Gene, 20, 1-10.
47. Kuo, Chia-Lam, and Campbell, J. (1983) Mol. Cell. Biol., 3, 1730-1737.
48. Beckwith, J.R., Pardee, A.B., Austrian, R., and Jacob, F. (1962) J. Mol. Biol., 5, 618-634.
49. St. John, T.P., and Davis, R.W. (1981) J. Mol. Biol., 152, 285-315.