

Defining the minimal length of sequence homology required for selective gene isolation by TAR cloning

V. N. Noskov, M. Koriabine, G. Solomon¹, M. Randolph¹, J. C. Barrett¹, S.-H. Leem², L. Stubbs³, N. Kouprina and V. Larionov*

Laboratory of Molecular Genetics and ¹Laboratory of Molecular Carcinogenesis, National Institute of Environmental Health Sciences, Research Triangle Park, NC 27709, USA, ²Department of Biology, Faculty of Natural Sciences, Dong-A University, Pusan 604-714, Korea and ³Genome Division, Lawrence Livermore National Laboratory, Livermore, CA 94550, USA

Received November 20, 2000; Revised January 20, 2001; Accepted January 26, 2001

ABSTRACT

The transformation-associated recombination (TAR) cloning technique allows selective and accurate isolation of chromosomal regions and genes from complex genomes. The technique is based on *in vivo* recombination between genomic DNA and a linearized vector containing homologous sequences, or hooks, to the gene of interest. The recombination occurs during transformation of yeast spheroplasts that results in the generation of a yeast artificial chromosome (YAC) containing the gene of interest. To further enhance and refine the TAR cloning technology, we determined the minimal size of a specific hook required for gene isolation utilizing the Tg.AC mouse transgene as a targeted region. For this purpose a set of vectors containing a B1 repeat hook and a Tg.AC-specific hook of variable sizes (from 20 to 800 bp) was constructed and checked for efficiency of transgene isolation by a radial TAR cloning. When vectors with a specific hook that was ≥ 60 bp were utilized, $\sim 2\%$ of transformants contained circular YACs with the Tg.AC transgene sequences. Efficiency of cloning dramatically decreased when the TAR vector contained a hook of 40 bp or less. Thus, the minimal length of a unique sequence required for gene isolation by TAR is ~ 60 bp. No transgene-positive YAC clones were detected when an *ARS* element was incorporated into a vector, demonstrating that the absence of a yeast origin of replication in a vector is a prerequisite for efficient gene isolation by TAR cloning.

INTRODUCTION

Sequencing the genomes of numerous organisms offers great promise for basic scientific research. Soon, emphasis will be

placed upon the identification, isolation and examination of the genes within the sequence(s), with the aim of understanding their function. Isolation of entire genes or specific chromosomal regions from mammalian genomes has typically involved a long and laborious process of identifying the clone containing the region of interest among thousands of random yeast artificial chromosome (YAC) or bacterial artificial chromosome (BAC) clones. The transformation-associated recombination (TAR) cloning technique greatly simplifies this cloning procedure, allowing entire genes and large chromosomal regions to be specifically, accurately and quickly isolated from total genomic DNA as large as YACs or YAC/BACs (1,2). This technique uses the *Saccharomyces cerevisiae* yeast as a tool to isolate the desired gene. The TAR cloning technique is based on *in vivo* recombination between genomic DNA and a vector containing sequences homologous to a gene of interest (hooks). During transformation of yeast spheroplasts, recombination results in the generation of a YAC containing the gene. The high selectivity of TAR gene isolation is provided by the absence of a yeast origin of replication (*ARS* element) in the TAR cloning vector. Therefore, propagation of TAR-generated YACs in yeast cells absolutely depends on acquisition of genomic DNA with *ARS*-like sequences that can function as origins of replication in yeast. Because these sequences are common in mammalian DNA and are present on average once every 20–40 kb (3), most mammalian genes can be isolated by TAR cloning using vectors containing two specific hooks that bracket the targeted gene. To isolate genomic regions lacking *ARS*-like sequences, a modified protocol of TAR cloning was developed (4). In this case a vector that has one unique and one repeated sequence hook, *Alu* or *B1* repeats for human or mouse DNA, respectively, is used. The presence of the repeated element makes it possible to isolate a gene as a set of nested overlapping fragments that extends from the specific hook to different upstream or downstream *Alu* or *B1* positions. Such an approach increases the likelihood of isolation of a genomic region (a gene) because at least one of the nested fragments is likely to contain an *ARS*-like sequence. Since only one of the ends is fixed, this approach is named radial TAR cloning.

*To whom correspondence should be addressed at: Laboratory of Biosystems and Cancer, National Cancer Institute, National Institutes of Health, Building 49, Room 4-A56, 49 Convent Drive, Bethesda, MD 20892, USA. Tel: +1 301 496 7941; Fax: +1 301 496 0332; Email: larionov@mail.nih.gov
Permanent address:

V.N. Noskov, St Petersburg State University, St Petersburg, Russia

TAR cloning is highly selective and typically produces libraries in which 0.05–1% of the transformants contain the desired gene. (Variation in yield of positive transformants is apparently caused by selection of targeting sequences; the presence of even short repetitive sequences in the hooks greatly decreases selectivity of TAR cloning.) It is also important to note that, in contrast to YACs constructed by a standard ligation method, the YACs generated by a TAR cloning technique exhibit a low level (if any) of chimeras. Over the last two years this new method was successfully applied for isolation of 10 genes and specific regions from human and mouse genomes as well as for closing the gaps on physical maps (2). For three genes (*BRCA1*, *HPRT1* and *BRCA2*), enriched by repeats, isolation of functional copies has been demonstrated, suggesting a high fidelity of TAR gene isolation.

A prerequisite for TAR cloning of a single copy gene from complex genomes is the selection of a unique targeting sequence called a hook. In all previous experiments on gene isolation, the size of a specific hook in a vector varied from 0.2–1.5 kb (2). In the current study, to further enhance and refine TAR cloning technology, we investigated the minimal size of a specific hook required for gene isolation using the Tg.AC mouse transgene as the targeted sequence (5,6). Based on our results on TAR cloning of this transgene, the minimal length of a unique sequence that is required for gene isolation from genomic DNA is ~60 bp.

MATERIALS AND METHODS

Host strain and yeast transformation

A highly transformable *S.cerevisiae* strain VL6-48 (*MAT α* , *his3- Δ 200*, *trp1- Δ 1*, *ura3-52*, *lys2*, *ade2-101*, *met14*), which has *HIS3* deleted (2), was used as a host for TAR cloning experiments. Yeast spheroplasts that enable efficient transformation were generated using a previously described protocol (2). Agarose plugs (100 μ l) containing ~5 μ g of mouse genomic DNA of high molecular weight from the liver tissue of the Tg.AC mouse were prepared for transformation as previously described (2). Linearized TAR cloning vectors (1 μ g) were added to the DNA-containing plugs before treating with agarase and presented to yeast spheroplasts.

Construction of TAR vectors for cloning of the Tg.AC transgene

All plasmid constructs were created using a universal TAR cloning vector pVC604. pVC604 (*HIS3/CEN6*) is a derivative of the Bluescript-based yeast–*Escherichia coli* shuttle vector pRS313 (7). The vector was generated by deletion of a ~295 bp fragment containing a yeast origin of replication (*ARSH4*) from pRS313. A set of radial TAR cloning vectors containing a *B1* repeat and different lengths of the simian virus 40 (SV40)-specific targeting sequence [3'-terminal polyadenylation signal sequence in the Tg.AC transgene (Fig. 1)] were constructed as follows. A 130 bp mouse *B1* element from WJ522 (8) was cloned into *SalI*–*SmaI* sites of the pVC604 giving a pVC604-B1 plasmid. An 800 bp SV40-specific fragment was PCR amplified from the pSV2 neo plasmid (Clontech, Palo Alto) using a pair of primers described in Table 1 and inserted into *ApaI*–*SalI* sites of pVC604-B1 to generate the SV800 TAR cloning vector with the targeting sequence (hook) of the largest size.

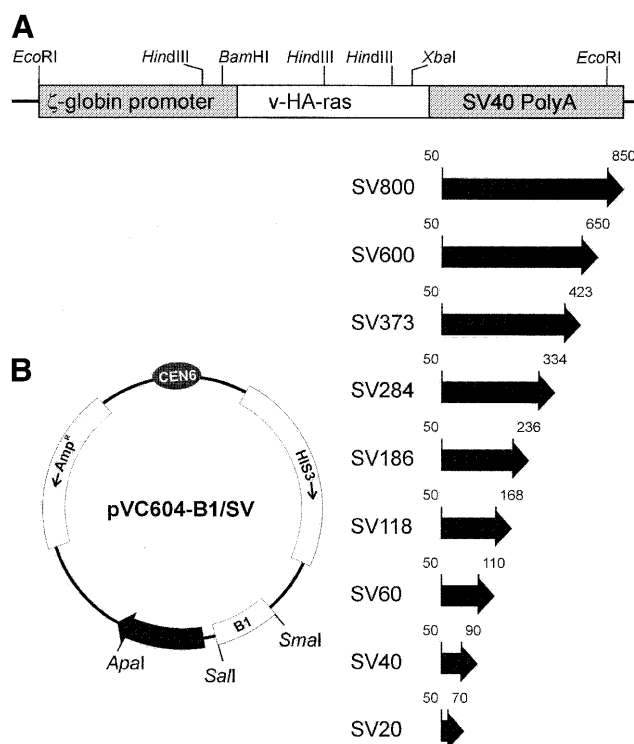


Figure 1. Construction of TAR cloning vectors containing a different length of homology to the Tg.AC transgene. (A) The structure of the Tg.AC transgene. The transgene unit consists of a ζ -globin promoter fused to the *v-Ha-ras* structural gene with a terminal SV40 polyadenylation signal sequence (5,20). Nine different sizes of targeting sequence (from 800 to 20 bp), shown by filled bars, were selected from the 3'-terminal region of the transgene containing an SV40 polyadenylation signal sequence. (B) A scheme of a TAR vector. A set of TAR vectors with targeting sequences of different sizes (800, 600, 373, 284, 186, 118, 60, 40 and 20 bp) homologous to the 3'-end of the transgene was constructed. Each vector contains a 130 bp *B1* repeat as a second targeting sequence. Vectors were linearized by *SalI* restriction endonuclease before transformation.

This targeting sequence has no ARS activity based on a low transformation efficiency of SV800. To construct the vectors with smaller targeting sequences, eight different SV-B1 cassettes were PCR amplified from the SV800 plasmid using the *B1*-specific 1-B1-R primer and one of the following primers: SV600-F, SV373-F, SV284-F, SV186-F, SV118-F, SV60-F, SV40-F and SV20-F (Table 1). Each PCR product, containing different sizes of the SV40 sequence (600, 373, 284, 186, 118, 60, 40 and 20 bp) and a 130 bp *B1* repeat, were inserted into *ApaI*–*SmaI* sites of pVC604. The TAR constructions were designated based on the length of a targeting SV40 sequence homologous to the Tg.AC transgene (Fig. 1). The constructed TAR cloning vectors were cut with *SalI* (the site is located between hooks) before transformation to yield linear molecules, bounded by the SV40 and *B1* hooks. To investigate how the presence of an ARS sequence within the vector affects the efficiency of gene isolation, two additional vectors, SV286-ARS⁻ and SV286-ARS⁺, were constructed based on the pVC604 plasmid. Both vectors contain a 286 bp SV40 targeting sequence inserted into *BamHI*–*XbaI* sites and a

Table 1. Primers used in this study for (A) construction of TAR-cloning vectors and (B) detection of Tg.AC positive YAC clones

	Name	Sequence	Restriction site
A	SV800-F	5'-atgc gggccc CATGATAAGATACATTGATG-3'	<i>ApaI</i>
	SV800-R	5'-atgc gtcgac GACAACTACCTACAGACAT-3'	<i>SalI</i>
	SV600-F	5'-atgc gggccc ACCTCTACAAATGTGGTATG-3'	<i>ApaI</i>
	SV373-F	5'-atgc gggccc TCTGTTATAGCAGTGCAGCT-3'	<i>ApaI</i>
	SV284-F	5'-atgc gggccc CCTTAGCAATTCTGAAGGAAAG-3'	<i>ApaI</i>
	SV184-F	5'-atgc gggccc AGATGGCATTCTTCTGAGCA-3'	<i>ApaI</i>
	SV118-F	5'-atgc gggccc CAGTTCATAGGTTGGAATCT-3'	<i>ApaI</i>
	SV60-F	5'-atgc gggccc TTATACACTTAAAAATTTTA-3'	<i>ApaI</i>
	SV40-F	5'-atgc gggccc TATTTACCTTAGAGCTTAA-3'	<i>ApaI</i>
	SV20-F	5'-atgc gggccc ATCTCTGTAGGTAGTTTGTTC-3'	<i>ApaI</i>
	SV286-F	5'-atgc ggatcc AACTTAGCAATTCTGAAGGA-3'	<i>BamHI</i>
	SV286-R	5'-atgc ctctaga GACAACTACCTACAGAGAT-3'	<i>XbaI</i>
	1-B1-F	5'-atgc gtcgac GGGCATGGTGGCGCACGCCTT-3'	<i>SalI</i>
	1-B1-R	5'-atgc ccggg AACAGGGTTTCTCTGTGTAGCC-3'	<i>SmaI</i>
	B	2-B1-F	5'-cctt gagctc GGGCATGGTGGCGCACGCCTT-3'
2-B1-R		5'-tata gagctc ACAGGGTTTCTCTGTGTAGCC-3'	<i>SacI</i>
ZG-F		5'-GTGAGAGGAATTACTGCTTCC-3'	
ZG-R		5'-AGGCTGCGCTGGAGTTGAGT-3'	

Upper case letters indicate a region of homology to the SV40 promoter. Lower case letters indicate sequences non-homologous to the SV40 promoter; bold letters indicate endonuclease restriction sites used for cloning of PCR products.

130 bp mouse *B1* element cloned into the *SacI* site of the polylinker. In addition, the vector SV286-ARS⁺ contains ~300 bp *ARSH4* sequence (7) cloned into a unique *NsiI* site located outside of a polylinker. These vectors were linearized by *XbaI* and *NotI* digestion before use for yeast transformation.

Detection of positive clones among yeast transformants

A pair of primers, ZG-F and ZG-R, specific to a ζ -globin promoter region was used for PCR screening of transformants for the presence of Tg.AC transgene sequences (Table 1). These primers generate a 419 bp PCR product that is diagnostic for recombination between a TAR vector and genomic Tg.AC transgene sequences (Fig. 2). Yeast genomic DNA isolated from the transformants was amplified under the following standard PCR conditions: 50 mM KCl, 10 mM Tris-HCl pH 9.0, 3.0 mM MgCl₂, 0.2 mM dTTP, dCTP, dGTP and dATP in a final volume of 50 μ l. Thermocycling conditions consisted of 35 cycles of 1 min at 94°C, 45 s at 55°C and 2 min at 68°C, followed by one extension cycle of 10 min at 72°C in a 9600 Thermocycler (Perkin-Elmer). In addition, individual positive clones were identified by yeast colony hybridization using a ³²P-labeled ζ -globin-specific fragment as a probe.

Characterization of YAC clones

Chromosome-sized DNAs from yeast transformants carrying YACs were separated by transverse alternating field electrophoresis (TAFE), blotted and hybridized with either an *Alu* or a ζ -globin promoter probe. To estimate the size of circular YACs, agarose DNA plugs were prepared from yeast

transformants and exposed to a low dose of γ -rays (5 krad) before TAFE analysis. At this dose ~10% of 100 kb circular DNA molecules are linearized (2).

RESULTS

TAR cloning of the Tg.AC transgene sequence with vectors containing different sizes of targeting sequence

In previous experiments on isolation of single copy genes by TAR cloning, the size of a specific targeting sequence in a vector varied from 200 to 1400 bp [summarized by Kouprina and Larionov (2)]. Experiments on isolation of human DNA from hybrid cell lines suggested that the size of the targeting sequences could be significantly shorter. To test this hypothesis, we chose the mouse Tg.AC transgene cassette as a model targeted region. The Tg.AC transgenic mouse contains approximately 40 copies of the transgene integrated into a unique site on chromosome 11 (5). A multicopy gene array was chosen as a target for this study to obtain statistically significant numbers of independent gene isolates. Each transgene unit consists of a ζ -globin promoter fused to the *v-Ha-ras* structural gene with a terminal SV40 polyadenylation signal and spans ~4.0 kb (Fig. 1A). To study regulation and expression of the transgene, the transgene and flanking genomic sequences were recently isolated using radial TAR cloning. The vector used contained a complete sequence of the *B1* repeat (130 bp) and ~400 bp of the SV40 sequence homologous to the transgene. Isolation of the transgene region from

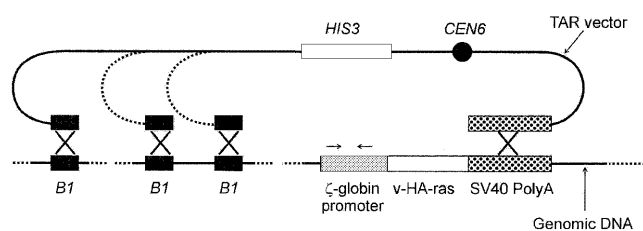


Figure 2. A scheme of isolation of the Tg.AC transgene as a series of circular YACs using a TAR cloning vector containing a 3' Tg.AC sequence and a *BI* repeat. Yeast spheroplasts were transformed with genomic mouse DNA along with a TAR cloning vector containing a 3' sequence (dotted box) and a *BI* at the ends of the linearized plasmid. Recombination between sequences in the vector and genomic DNA containing the transgene leads to the establishment of circular YACs that extend from the 3' sequence to various *BI* positions. Because the vector lacks *ARS*, the only YACs that will be stably maintained are those that included mouse DNA fragments containing a yeast *ARS*-like sequence. *CEN6* corresponds to the yeast chromosome VI centromere and *HIS3* is a selectable marker. Arrows indicate positions of primers (specific for a ζ -globin promoter region) used for detection of positive clones among primary transformants.

total genomic DNA was highly selective when utilizing this vector: roughly one out of every 100 primary transformants contained the targeted region (9).

To determine what would be the minimum size of a hook sufficient for specific gene isolation, we constructed a set of radial TAR cloning vectors with larger and smaller sizes of SV40 specific hooks (from 20 to 800 bp) (Fig. 1). When the vectors are linearized, one end contains the SV40 sequence and the other end contains 130 bp of the *BI* repeat (see Materials and Methods). Homologous recombination between Tg.AC mouse DNA and the linearized vector results in the formation of circular YACs containing the Tg.AC transgene and flanking mouse chromosome 11 regions (Fig. 2). The efficiency of these vectors in isolation of the transgene from total genomic DNA, prepared from the liver tissue of a transgenic animal, was tested. Five yeast spheroplast transformation experiments were carried out as previously described (2). Using 5 μ g of mouse DNA, 1 μ g of the linearized vector and 2×10^8 spheroplasts, approximately 500–800 *His*⁺ transformants per experiment were obtained.

Clones containing the Tg.AC transgene and flanking regions were identified among primary yeast transformants by using a pair of diagnostic primers specific for the transgene ζ -globin promoter sequence (Table 1). The results on efficiency of isolation of the Tg.AC transgene region using TAR vectors containing targeting sequences of different sizes is summarized in Tables 2 and 3. With a vector containing the largest hook (800 bp), cloning of the Tg.AC transgene sequence from mouse genome was highly specific: one or two out of 100 yeast transformants analyzed contained a YAC with the Tg.AC transgene. Surprisingly, the same yield of positive clones was observed with the vectors containing a 600, 373, 284, 186, 118 and 60 bp SV40-specific hook. The selectivity of cloning decreased significantly when the vector contained a 40 bp SV40-specific sequence (Table 3). TAFE analysis of irradiated DNA isolated from the transgene-positive transformants demonstrated the presence of circular YACs with sizes ranging from 50 to >200 kb (data not shown). Such a broad distribution in the size of YAC isolates is characteristic for clones generated by a radial TAR cloning when one of the hooks has homology

to multiple regions (4). No positive clones were obtained when only a 20 bp SV40-specific sequence was included in the TAR vector, indicating that the minimal length of a unique sequence sufficient for selective gene isolation is close to that required for human DNA isolation with the *Alu*-containing vectors.

Table 2. TAR cloning with *ARS*-containing vectors^a

Vector	Transformation efficiency	Yield of Tg.AC YACs
SV286- <i>ARS</i> ⁻	335 (280–440) ^b	10/500 (2.0%)
SV286- <i>ARS</i> ⁺	1683 (1280–21 600)	0/3000 (0.0%)

^aThe results are a summary of two independent TAR cloning experiments; i.e. for experiments with separately prepared yeast spheroplasts.

^bNumbers in parentheses show variation of transformation efficiency observed in parallel samples with 1 μ g of *ARS*less vector and 0.1 μ g of SV286-*ARS*⁺ vector.

Table 3. TAR cloning of Tg.AC transgene by vectors containing different sizes of targeting sequences

TAR vector	No. of transformants analyzed ^a	No. of Tg.AC positive clones	Yield (%)
SV800	200	3	1.50
SV600	400	10	2.50
SV373	1000	27	2.70
SV284	200	4	2.00
SV184	200	5	2.50
SV118	800	19	2.38
SV60	800	18	2.25
SV40	800	4	0.50
SV20	400	0	0

^aTransformants were selected from several independent transformation experiments. The frequency of transformation varied from 296 to 976 *His*⁺ colonies in different TAR cloning experiments.

Cloning of the Tg.AC transgene cassette with an *ARS*-containing TAR vector

We have investigated the effect that the presence of a yeast origin of replication (an *ARS* element) in a vector has on the efficiency of TAR gene isolation. We compared the efficiency of isolation of a transgene using two vectors, SV286-*ARS*⁻ and SV286-*ARS*⁺. The vector SV286-*ARS*⁺ is identical to SV286-*ARS*⁻ except it contains a functional *ARS* element (*ARSH4*) (see Materials and Methods). After linearization with two enzymes producing non-compatible ends, the vectors were transformed into yeast spheroplasts along with genomic DNA from the transgenic mouse. In contrast to results obtained with TAR vectors lacking a yeast origin of replication, SV286-*ARS*⁺ transformed yeast spheroplasts with a high efficiency. When 1 μ g of the *Xba*I/*Not*I-linearized vector and 5 μ g of genomic DNA were used in transformation experiments, the yield of *His*⁺ transformation was ~100-fold greater than that observed with the *ARS*less vector (Table 2). This result suggests a high

level of vector circularization occurred during yeast transformation despite the fact that the vector contains non-compatible protuberant ends. Indeed, analysis of 100 randomly selected transformants illustrated the presence of circularized SV286-ARS⁺ vector molecules without any mouse DNA inserts in all clones analyzed (data not shown). When the ARSless vector SV286-ARS⁻ was used >99% of the transformants contained the vector with genomic DNA inserts.

When 3000 transformants, obtained with the SV286-ARS⁺ vector in two independent experiments, were analyzed by PCR for the presence of the Tg.AC transgene sequence, no positive clones were found. In contrast, 10 positive clones were identified among 500 transformants obtained in the similar experiments when the plasmid SV286-ARS⁻, lacking an origin of replication, was used (Table 2). Thus, ARS-containing TAR vectors are not efficient for isolating specific regions from complex genomes.

DISCUSSION

The TAR cloning strategy that allows for selective isolation of large genomic regions promises to greatly simplify the physical and functional analysis of mammalian and other complex genomes (2). The analysis of polymorphic regions, separation of haplotypes and cloning of regions that are not present in existing BAC (or YAC) libraries to close gaps on physical maps are among the many utilities of this technique. With the completion of sequencing of the human genome, the TAR cloning technique will also be very useful for isolating regions containing intact genes, including all regulatory elements required for correct gene expression, for functional studies.

In this work we demonstrated that the minimal length of a unique targeting sequence required for gene isolation by TAR is only 60 bp. A further increase in length of a targeting sequence had no effect on selectivity of gene isolation. Therefore, the minimal length of homology sufficient for gene isolation is smaller than that required for spontaneous mitotic recombination in yeast (10). It is worth noting that the same length of homology is required during gene disruption experiments using a PCR product (11,12) and during recombination-mediated plasmid construction in yeast (13,14). Such a small size greatly facilitates selection of a targeting sequence(s) for future gene isolation experiments. Firstly, repeated elements are so frequent in mammalian genomes that it may not be possible to identify large, unique hook sequences in all genomic regions. Secondly, short sequences can be easily synthesized; this fact should streamline the design of TAR vectors significantly. In addition, the data we present supports our previous conclusion that the absence of a yeast origin of replication (an ARS sequence) in a vector is a prerequisite for efficient gene isolation by TAR cloning (1). Although a high enrichment for the Tg.AC transgene sequence was observed with an ARSless vector, no positive clones were identified among a similar number of transformants when an ARS element was incorporated into the vector. Instead, the frequency of transformation was greatly increased as a result of the recircularization of the vector in yeast cells. Similar results, demonstrating a high efficiency of recircularization during yeast transformation, were obtained previously when adenovirus DNA was cloned using ARS-containing vectors

(15). Moreover, when *Alu*-containing ARS⁺ vectors were used for cloning of total human DNA (i.e. non-selective TAR cloning at all), the fraction of clones with the circularized vector was >70% (1). These results contradict recently published experiments of Bhargava *et al.* (16) who described the isolation of specific regions the human genomic DNA using an ARS-containing pClasper vector. The authors reported that the frequency of positive clones among the primary transformants varied from 0.05 to 3%. Surprisingly, this frequency is similar to that observed during TAR cloning with vectors lacking a yeast origin of replication [summarized by Kouprina and Larionov (2)]. Because transformants resulting from plasmid circularization should be the predominant transformants obtained in these circumstances, it is difficult to explain such a high efficiency of gene isolation using an ARS-containing vector. Discrepancy between our results and those published by Bhargava *et al.* (16) cannot be explained by differences in vector constructs (both constructs are pRS-based derivatives and contain the same ARS sequence) or by differences in chosen targeting sequences (circularization of vectors should not depend on targeting sequences). We also exclude that the discrepancy is due to the vectors being differently linearized before TAR cloning. (We observed approximately the same yield of circularized molecules with a vector cut by enzymes producing blunt and cohesive ends.) Though we cannot exclude that differences in results were caused by use of different yeast hosts, such an explanation is unlikely. Ineffective plasmid re-ligation during yeast transformation has been observed only in mutants on joining of non-homologous ends (17). Typically this class of mutants exhibits a low frequency of transformation and cannot be used for TAR cloning experiments (N.Kouprina and S.-H.Leem, unpublished data). Whatever the explanation of contradiction between our results and those published by Bhargava *et al.* (16) might be, it is clear that gene isolation is much more efficient if the vector lacks a yeast origin of replication. A selective cloning of genes with ARS-containing vectors would require development of new vectors with additional genetic markers allowing counter-selection of recombinant events from vector recircularization. A new generation of vectors may be essential for cloning and analysis of heterochromatin blocks (centromeric and telomeric regions) that may lack yeast ARS-like sequences.

NOTE ADDED IN PROOF

In recent experiments, we have demonstrated that the efficiency of TAR cloning of a single copy gene (the human HPRT gene) was the same with hooks varying from 380 to 60 bp. Yield of HPRT-positive clones was ~0.5%.

ACKNOWLEDGEMENTS

We would like to thank Dr Ron Cannon for his interest in this work and for providing us with the transgenic mice.

REFERENCES

1. Larionov, V., Kouprina, N., Graves, J., Chen, X.N., Korenberg, J.R. and Resnick, M.A. (1996) Specific cloning of human DNA as yeast artificial chromosomes by transformation-associated recombination. *Proc. Natl Acad. Sci. USA*, **93**, 491-496.

2. Kouprina, N. and Larionov, V. (1999) Selective isolation of mammalian genes by TAR cloning. In Boyle, A.L. (ed.), *Current Protocols in Human Genetics*. John Wiley and Sons, New York, NY, Vol. 1, pp. 5.17.1–5.17.21.
3. Stinchomb, D.T., Thomas, M., Kelly, J., Selker, E. and Davis, R.W. (1980) Eukaryotic DNA segments capable of autonomous replication in yeast. *Proc. Natl Acad. Sci. USA*, **77**, 4559–4563.
4. Kouprina, N., Annab, L., Graves, J., Afshari, C., Barrett, J.C., Resnick, M.A. and Larionov, V. (1998) Functional copies of a human gene can be directly isolated by transformation-associated recombination cloning with a small 3' end target sequence. *Proc. Natl Acad. Sci. USA*, **95**, 4469–4474.
5. Leder, A., Kuo, A., Cardiff, R.D., Sinn, E. and Leder, P. (1990) v-Ha-ras transgene abrogates the initiation step in mouse skin tumorigenesis: effects of phorbol esters and retinoic acid. *Proc. Natl Acad. Sci. USA*, **87**, 9178–9182.
6. Cannon, R.E., Spalding, J.W., Virgil, K.M., Faircloth, R.S., Humble, M.C., Lacks, G.D. and Tennant, R.W. (1998) Induction of transgene expression in Tg.AC (v-Ha-ras) transgenic mice concomitant with DNA hypermethylation. *Mol. Carcinog.*, **21**, 244–250.
7. Sikorski, R.S. and Hieter, P. (1989) A system of shuttle vectors and yeast host strains designed for efficient manipulation of DNA in *Saccharomyces cerevisiae*. *Genetics*, **122**, 19–27.
8. Cancilla, M., Graves, J., Matesic, L., Reeves, R., Tainton, K., Choo, K., Larionov, V. and Kouprina, N. (1998) Rapid cloning of mouse DNA as yeast artificial chromosomes by transformation-associated recombination (TAR). *Mamm. Genome*, **9**, 157–159.
9. Humble, M.C., Kouprina, N., Noskov, V., Graves, J., Garner, E., Resnick, M., Tennant, R.W., Larionov, V. and Cannon, R.E. (2000) Radial TAR cloning from the Tg.AC mouse. *Genomics*, **70**, 292–299.
10. Jinks-Robertson, S., Michelitch, M. and Ramcharan, S. (1993) Substrate length requirements for efficient mitotic recombination in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.*, **13**, 3937–3950.
11. Lorenz, M.C., Muir, R.S., Lim, E., McElver, J., Weber, S.C. and Heitman, J. (1995) Gene disruption with PCR products in *Saccharomyces cerevisiae*. *Gene*, **158**, 113–117.
12. Manivasakam, P., Weber, S.C., McElver, J. and Schiestl, R.H. (1995) Micro-homology mediated PCR targeting in *Saccharomyces cerevisiae*. *Nucleic Acids Res.*, **23**, 2799–2800.
13. Hua, S.-B., Qui, M., Chan, E., Zhu, L. and Luo, Y. (1997) Minimal length of sequence homology required for *in vivo* cloning by homologous recombination in yeast. *Plasmid*, **38**, 91–96.
14. Oldenburg, K.R., Vo, K.T., Michaelis, S. and Paddon, C. (1997) Recombination-mediated PCR-directed plasmid construction *in vivo* in yeast. *Nucleic Acids Res.*, **25**, 451–452.
15. Ketner, G., Spencer, F., Tugendreich, S., Connelly, C. and Hieter, P. (1994) Efficient manipulation of the human adenovirus genome as an infectious yeast artificial clone. *Proc. Natl Acad. Sci. USA*, **91**, 6186–6190.
16. Bhargava, J., Shashikant, C.S., Carr, J.L., Juan, H., Bentley, K.L. and Ruddle, F.H. (1999) Direct cloning of genomic DNA by recombinogenic targeting method using a yeast-bacterial shuttle vector, pClasper. *Genomics*, **62**, 285–288.
17. Lewis, L.K. and Resnick, M.A. (2000) Typing up loose ends: nonhomologous end-joining in *Saccharomyces cerevisiae*. *Mutat. Res.*, **451**, 71–89.