Yellman, 2020 Supplemental manuscript

An IT toolset

The basic toolset for integrative targeting are the plasmid PCR templates for the IT cassettes (Table 1) and the plasmids and chromosomally-integrated constructs for expression of I-SceI (Table 1 and Table 2). In addition, there is a list of genomic target (GT) sites (Figure S1 and Table S2) that may be good sites for a variety of genome engineering applications. Finally, a set of yeast strains with the *IT4* cassette integrated is provided ready to use for chromosomal integrations. The following discussion summarizes the resources not fully described in the main paper.

Genomic target (GT) sites

Chromosomal loci amenable to the integration of DNA have a wide variety of genome engineering applications, such as expression of individual genes or multi-gene pathways, reporters of promoter activity or protein stability, and control of gene dosage. Intergenic regions are easily identified in yeast genomes, and integrations sites chosen to minimize the potential for interference with neighboring genes. Good candidate sites can be found in large intergenic regions, as far as possible from the start codon of a neighboring gene, at least 200 bp from stop codons, and away from origins of replication or other functional elements. It is best to avoid AT-rich DNA sequences, which are particularly common in yeast terminators of transcription and origins of replication, and also required for the activity of subset of promoters (Yarger, Armilei and Gorman, 1986; Chen *et al.*, 1996; Sugihara, Kasahara and Kokubo, 2011; Miura *et al.*, 2019). Transposons and their remnants are also potential sites, although repetitive DNA elements within a transposon should be identified and their effects considered (Sakai, Shimizu and Hishinuma, 1990).

A set of potentially useful chromosomal sites (referred to as GTs) is provided (Figure S1 and Table S2). The sites were chosen by manual inspection of *S. cerevisiae* chromosomes, and include loci linked to centromeres, telomeres, and the rDNA array, but all are from regions with actively transcribed nearby genes.

Yeast strains with *GT:IT4* integrations, ready to use for integration

Six yeast strains are provided with *IT4* integrated into individual GT sites (Table 2). As in other applications of the IT method, integrations do not require a selectable marker, DNA constructs have been successfully integrated into *GT1* and *GT2*, but the use of any particular integration site requires a minimum of phenotypic analysis to reveal potential negative consequences. The suitability of *S. cerevisiae* chromosomal sites for heterologous expression can depend on local effects which are difficult to predict and best evaluated by the integration of a transcriptional reporter (Flagfeldt *et al.*, 2009).

Yeast strains for *GAL* expression control of chromosomally-integrated genes

The *IT4* cassette has been integrated into the native *GAL1* and *GAL10* genes, and the strains are provided (Table 2), The chromosomal *gal1::IT4* and *gal10::IT4* constructs can be used as integration targets that place genes under *GAL* control, and their presence can be followed by a Gal^- phenotype. The stable expression of foreign genes under *GAL* control may be a useful for large-scale assays of conditional gene expression or overexpression phenotypes. When introducing foreign DNA into the *gal1::IT* and *gal10::IT* cassettes, it is important to express I-

SceI from a construct that does not carry a repair template for the DSB site. For example, when replacing a *gal1::IT* cassette, I-SceI should not be expressed from a *pGAL1* construct, but from *pGAL10*. Conversely, when replacing *gal10::IT* cassettes, I-SceI should be under *pGAL1* control. The I-SceI expression plasmids (Table 1) provide flexibility to avoid this problem.

REFERENCES

Chen, S. *et al.* (1996) "Transcriptional terminators of RNA polymerase II are associated with yeast replication origins," *Nucleic Acids Research*, 24(15), pp. 2885–2893. doi: 10.1093/nar/24.15.2885.

Flagfeldt, D. B. *et al.* (2009) "Characterization of chromosomal integration sites for heterologous gene expression in Saccharomyces cerevisiae.," *Yeast (Chichester, England)*, 26(10), pp. 545–551. doi: 10.1002/yea.

Miura, O. *et al.* (2019) "A strong structural correlation between short inverted repeat sequences and the polyadenylation signal in yeast and nucleosome exclusion by these inverted repeats," *Current Genetics*. Springer Berlin Heidelberg, 65(2), pp. 575–590. doi: 10.1007/s00294-018-0907-8.

Sakai, A., Shimizu, Y. and Hishinuma, F. (1990) "Integration of heterologous genes into the chromosome of Saccharomyces cerevisiae using a delta sequence of yeast retrotransposon Ty," *Applied Microbiology and Biotechnology*, 33(3), pp. 302–306. doi: 10.1007/BF00164526.

Sugihara, F., Kasahara, K. and Kokubo, T. (2011) "Highly redundant function of multiple AT-rich sequences as core promoter elements in the TATA-less RPS5 promoter of Saccharomyces cerevisiae," *Nucleic Acids Research*, 39(1), pp. 59–75. doi: 10.1093/nar/gkq741.

Yarger, J. G., Armilei, G. and Gorman, M. C. (1986) "Transcription terminator-like element within a Saccharomyces cerevisiae promoter region.," *Molecular and Cellular Biology*, 6(4), pp. 1095–1101. doi: 10.1128/mcb.6.4.1095.

Yellman 2020, Figure S1

Figure S1. Genomic targets (GTs): chromosomal loci amenable to the integration of foreign DNA. Approximate locations of GT sites in the genome of *S. cerevisiae*. Chromosomes lengths are proportional to their actual sizes, and centromeres are marked with black ovals. The *IT4* cassette was integrated at six GTs (fuchsia) on different yeast chromosomes.

Table S1. Oligonucleotides used in this study.

oligo	sequence (5'-3')	notes
To integrate IT cassettes into *SPO12(352-375)* (genomic target sequences in light font, *IT* cassette annealing in bold)		
IT24F	GGGATCTTTTAAAACAGGATCCGAAGAGGAAACTCCAACTACAACAAAGATTTGCACGGA**CGTCACGACCTGCGAT**	*spo12::IT*
IT24R	GGATTGACTTTCTTTTTCTTTCCGAACATCTTAACTTTGTGCTCATTCAATTTCAATGAGCATGG**GGCTGTCAGGCGTGCACGAT**	
To repair *spo12(352-375)::IT* with phosphorylation site mutations (genomic identity in light font, mutations in bold)		
12AAF	CCGAAGAGGAAACTCCAACTACAACAAAGATTTGCAG**CA**CCGACTGACAGGCTGGTAG**CA**CCATGCTCATTGAAATTGAATGAGCACAAAGTTAAGATG	*spo12-S118A, S125A*
12AAR	CATCTTAACTTTGTGCTCATTCAATTTCAATGAGCATGG**TG**CTACCAGCCTGTCAGTCGG**TG**CTGCAAATCTTTGTTGTAGTTGGAGTTTCCTCTTCGG	
12DDF	CCGAAGAGGAAACTCCAACTACAACAAAGATTTGCAG**AT**CCGACTGACAGGCTGGTAG**AT**CCATGCTCATTGAAATTGAATGAGCACAAAGTTAAGATG	*spo12-S118D, S125D*
12DDR	CATCTTAACTTTGTGCTCATTCAATTTCAATGAGCATGG**AT**CTACCAGCCTGTCAGTCGG**AT**CTGCAAATCTTTGTTGTAGTTGGAGTTTCCTCTTCGG	
12EEF	CCGAAGAGGAAACTCCAACTACAACAAAGATTTGCAG**AA**CCGACTGACAGGCTGGTAG**AA**CCATGCTCATTGAAATTGAATGAGCACAAAGTTAAGATG	*spo12-S118E, S125E*
12EER	CATCTTAACTTTGTGCTCATTCAATTTCAATGAGCATGG**TT**CTACCAGCCTGTCAGTCGG**TT**CTGCAAATCTTTGTTGTAGTTGGAGTTTCCTCTTCGG	
To synthesize IT cassettes that precise replace yeast genes (genomic target sequences in light font, IT cassette annealing in bold)		
IT36F	TGAAGTTCAAGAGAAGTGGGTAAGAATACATTGAACTAAGCCTTCATATATTTACAAAAC**CGGACGTCACGACCTGCG**	*pre5::IT*
IT36R	AAGGCTTCTTTGGTCTTGCTTTGACAATTGCTTAGTGGAAACGATACTTTGTGTCCAGAC**GGCTGTCAGGCGTGCACG**	
IT37F	ACCTTCAAGTTAGTGCTCAGCAGTTCATCAGGATCAGTTGCATAAACACAATTACACAAA**CGGACGTCACGACCTGCG**	*pre6::IT*
IT37R	TGCATTTTTATTATTGCTGTTATTTTATATAGGTTTTATGCCCAATATATATCGCCGTTT**GGCTGTCAGGCGTGCACG**	
IT39F	TTTTTAGTTTCAATTAGTAGTAAGCAACCATAAGACACCAATCAACACAGTTCTATAATT**CGGACGTCACGACCTGCG**	*pre8::IT*
IT39R	AATTGGCTTTCTTTTGGATAAGTTGAGTGAGATGGGTGATTGGCGGGGATAATTATAACA**GGCTGTCAGGCGTGCACG**	
IT40F	AGTTTTCGATCAGTCTCTATTTTAATAATTGATTATTGGATATAGTTAGTAGTGTTAAAC**CGGACGTCACGACCTGCG**	*pre9::IT*
IT40R	AGATATGTTTCTATGCGTACATATTTATATAAGCATGAAGTCAAACAATACTTTCCAACC**GGCTGTCAGGCGTGCACG**	
IT41F	AGCTTTCAAACACATTCGAGCGTCGCATCATCAAATTAACAAAAGCATAACTCTTCAGCA**CGGACGTCACGACCTGCG**	*pre10::IT*
IT41R	CACGTGAGTTCATATTATTTCAACTCTTTGGTTCTTCTTAACGTATTATCAGAATGTCAC**GGCTGTCAGGCGTGCACG**	
IT43F	AACTATAGAATACAAGTCAATAAGAACATAAATTCCAATTGTCTAATATAACTAGAAAGG**CGGACGTCACGACCTGCG**	*pup2::IT*
IT43R	TTTGTAAGGTTTTCTTTCTTTGTTAGTGTGACGTTGGTATTTACCTTTATGTAACTATAT**GGCTGTCAGGCGTGCACG**	

IT60F	CAATTGAGATATTTTTTTGAACCATCGTAATAAATCAAGGAAGTTAGGAACTATACGGA**CGGACGTCACGACCTGCG**	*rpt1::IT*
IT60R	GTTGTACTTGTTGTTTTTATTAACCCTTCTCATAATATATAATAAAAAAGTACATTGACG**GGCTGTCAGGCGTGCACG**	
IT61F	GTGTCAAGCTTACAGTCAATTCAAGTTAAAGAAGAGACAGCGTTTTTAATAGAAAATAAA**CGGACGTCACGACCTGCG**	*rpt2::IT*
IT61R	GCTTTGTTGAAGCATTACTTATATACGTAAATCTATTATAGAAAAATGTAAATGGGCACA**GGCTGTCAGGCGTGCACG**	
IT62F	GTTATACAAGAGGCAATCGGATAAAATTCAAGTATTTATAAAGTAAAAACCACAACAAGA**CGGACGTCACGACCTGCG**	*rpt3::IT*
IT62R	TGTCTAGATAGAAATGTACGTAATACTTTATGTAGACCTACGTTGAAGCATATATGGTAA**GGCTGTCAGGCGTGCACG**	
IT64F	GTGACTTATTTGTGCATTATAGAGGTGAGAACAAATTGGAAAGTTTTGATTTTAGTTTAAG**CGGACGTCACGACCTGCG**	*rpt5::IT*
IT64R	GGATTGTTCTAATATGTAGATATGTGAATGGCGGCTTGATAAATCAAAATATTATTATTT**GGCTGTCAGGCGTGCACG**	
IT45F	CTTCATGCTAAATCATATAAGGGCAGAGACGAAGCAAAGCGAAAAAAACATATTACAATC**CGGACGTCACGACCTGCG**	*scl1::IT*
IT45R	CCATACATATATTGAGTGTTGACGCGTGTGATTTCACATTATGTTGTGGCAGGAAGTATG**GGCTGTCAGGCGTGCACG**	
To amplify IT cassettes for integration into selected genomic (GT) target sites (genomic target sequences in light font, IT cassette annealing in bold)		
IT07F	GATCGTCATAACCAAATATTGCTAGAAACAAAAGCGTTCGGGATAGGTCAAACCATGATGAAC**CGGACGTCACGACCTGCGAT**	*GT1::IT*
IT07R	GTCAGACATTCCTAAATTTGCCCTTTTTAAAAAATAGCACGCTTTAAGATATGCGTGTGGAACTA**GGCTGTCAGGCGTGCACGAT**	
IT23F	GCGTATATATACAGAAAATTTTATTATTTTATATTCAAAAACAAGAAAACAAAAGAGAAAACATTAACA**CGGACGTCACGACCTGCGAT**	*GT2::IT*
IT23R	CATACAGGATTTAAAGAAACAGTTGTAAAAAGTATTACTTTCTAGGACTATTAGTTATGAAAATTACA**GGCTGTCAGGCGTGCACGAT**	
IT08F	GCCATTTTCTTGGACTGTAAATCATACTTGGATGTTGTGCATTAGTCAATAATCGGTTCTTGTTCC**CGGACGTCACGACCTGCGAT**	*GT5::IT*
IT08R	CACCAAAAGGACCGCCGCATGATTTACCATAATTATTTCTCCCTTCATTTACATGTAATCGTT**GGCTGTCAGGCGTGCACGAT**	
IT09F	GAAGGAGGTCACAACCGCGAGAGAAAGTAAGATGTCATGTTTTTGAATATGAACCTGGAG**CGGACGTCACGACCTGCGAT**	*GT8::IT*
IT09R	CCTACATTTCAGAGTTAAAATAAATCTAAAGGTGCCACTGGAAAGAGTTATATGCGAG**GGCTGTCAGGCGTGCACGAT**	
IT10F	CAAGACCTATATATTAAAGTAAAACAACATTGAATCAGATCGCATACTGGTGTAG**CGGACGTCACGACCTGCGAT**	*GT11::IT*
IT10R	CCTTCCCAAACGTGTCCCACACCAAATTGAGCACCCTGCTCGATTATCATATGTCTACTTCACGT**GGCTGTCAGGCGTGCACGAT**	
IT11F	GGGGTGCTTTGGGTTTATTTAATCGTCAAATATAGTCCGATTTGTCTCATTATTCTTGTGTC**CGGACGTCACGACCTGCGAT**	*GT12::IT*
IT11R	GGAAAAATTGCAGAATATCTAGGTAGTAGTCGCTTCAATAAATTTTTTTCC**GGCTGTCAGGCGTGCACGAT**	

Table S2. Genomic target (*GT*) sites.

locus	chromosome	SGD coordinates	comments
GT1	XIV, left arm	335522/335523	*CBK1-YGP1* intergenic
GT2	IV, left arm	443742/443743	5967 bp left of *CEN4; ATP16-MCD1* intergenic
GT3	V, right arm	152859/152860	755 bp right of *CEN5*
GT4	II, right arm	45436/45437	455 bp right of *CEN2*
GT5	XII, right arm	449338/449339	~2.5kb left of the rDNA array
GT6	VII, left arm	324086/324087	*SRM1-TOS8* intergenic, tightly linked to *GT16*
GT7	XI, left arm	439817/439818	311 bp left of *CEN11*
GT8	XIII, right arm	512029/512030	*NCW1-PKR1* intergenic
GT9	XII, left arm	150359/150360	468 bp left of *CEN12*
GT10	V, left arm	151409/151410	577 bp left of *CEN5*
GT11	III, right arm	116855/116856	856 bp right of *CEN3, YCR001W-CDC10* intergenic
GT12	V, right arm	555845/555846	subtelomeric, *FAU1-TOG1* intergenic
GT13	VII, left arm	16719/16720	subtelomeric, *ADH4-ZRT1* intergenic
GT14	XIV, right arm	629371/629372	496 bp right of *CEN14*
GT15	XIII, right arm	268744/268745	595 bp right of *CEN13*
GT16	VII, left arm	328174/328175	*TOS8-VPS45* intergenic, in transposon *YGLWdelta6*

Table S3. Reaction conditions for PCR synthesis of *IT* cassettes.

Notes:

The first few amplification cycles of the PCR synthesis use an annealing temperature appropriate to the common regions of all synthesis primers. Subsequent amplification cycles use a higher annealing temperature based on the T_m of the full-length primers. We show 70ºC as the higher T_m, but that value should be adjusted depending on the specific primer sequences used. The PCR products should be purified over a DNA minicolumn before being used in a high-efficiency PEG/LiOAc transformation of yeast.

predicted product sizes (base pairs):
IT1: 1260 + ~ 80 = ~ 1340
IT2, *IT3*: 1278 + ~ 80 = ~ 1358
IT4, *IT5*: 1296 + ~ 80 = ~ 1376

REACTION MIXTURE

	reagent	**volume used per rxn (µl)**	**final concentration**
H_2O	pure water	7.5	-
RXN mix	2x KAPA HiFi reaction mixture	10	1x
$MgCl_2$	in reaction mixture	-	2.5mM
dNTPs	in reaction mixture	-	0.3mM
polymerase	in reaction mixture	-	0.4U total
primers	mixed primers, 3µM each	2	0.3µM each
template	20x-diluted plasmid DNA	0.5	5ng total
each rxn	complete reaction mixture	20	-

THERMOCYCLER PROGRAM

	# of cycles	**denaturation**	**annealing**	**synthesis**
initial denaturation	1	96ºC / 60s		
1st amplification set	3	96ºC / 20s	56ºC / 30s	72ºC / 75s
2nd amplification set	30	96ºC / 20s	70ºC / 30s	72ºC / 75s
final amplification	1			72ºC / 120s
store	1		4ºC / hold	