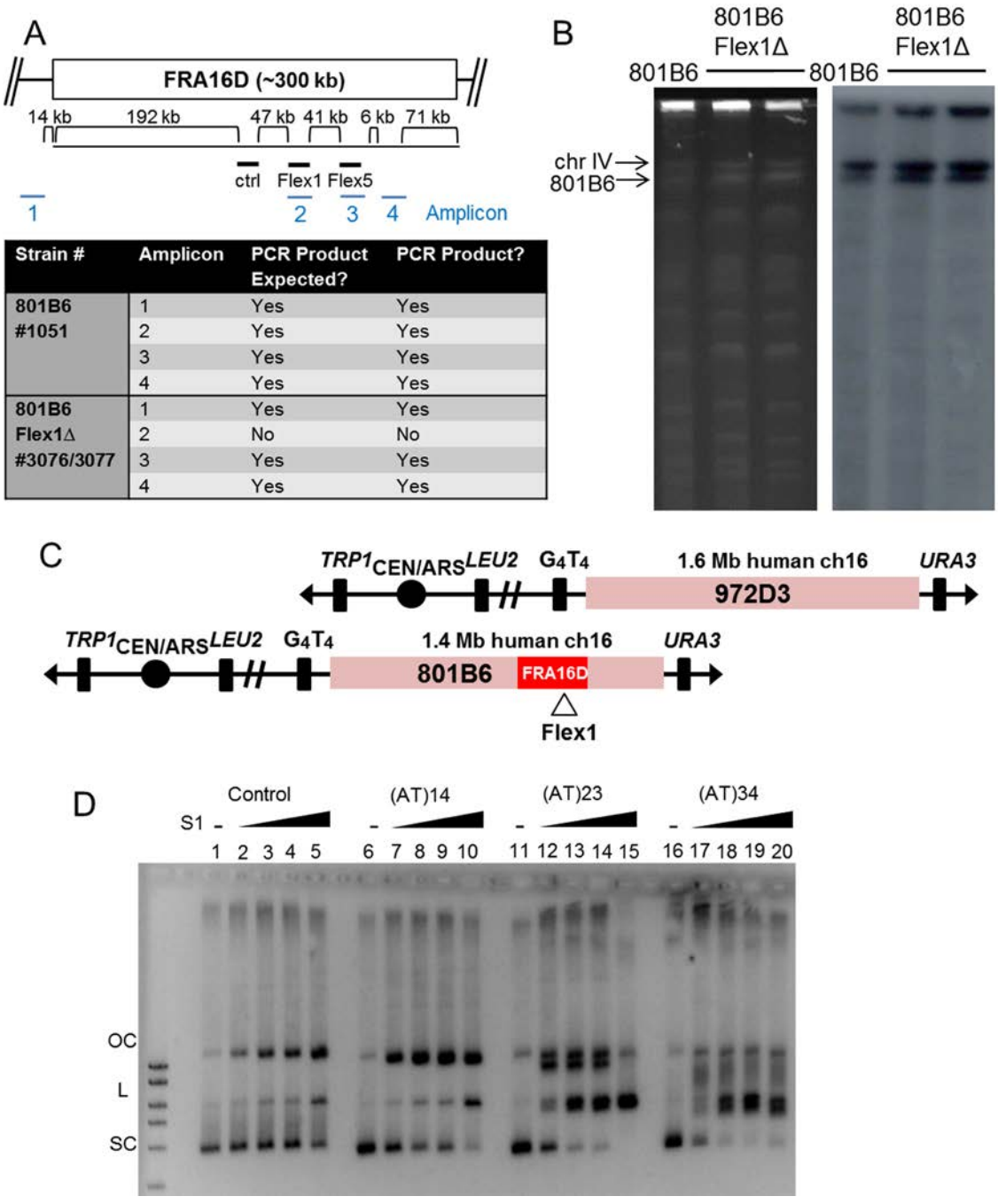
Cell Reports, Volume *27*

Supplemental Information

Sequence and Nuclease Requirements for Breakage and Healing of a Structure-Forming (AT)n Sequence within Fragile Site FRA16D

Simran Kaushal, Charles E. Wollmuth, Kohal Das, Suzanne E. Hile, Samantha B. Regan, Ryan P. Barnes, Alice Haouzi, Soo Mi Lee, Nealia C.M. House, Michael Guyumdzhyan, Kristin A. Eckert, and Catherine H. Freudenreich



Strain #	Amplicon	PCR Product Expected?	PCR Product?
801B6 #1051	1	Yes	Yes
	2	Yes	Yes
	3	Yes	Yes
	4	Yes	Yes
801B6 Flex1Δ #3076/3077	1	Yes	Yes
	2	No	No
	3	Yes	Yes
	4	Yes	Yes

Figure S1. Confirmation of FRA16D YAC integrity (related to Figure 1). (A) Large FRA16D YAC structure was verified by PCR amplifying the indicated amplicons. (B) Overall size of the 801B6 YAC (~1400 kb) was verified by pulsed field gel electrophoresis of intact chromosomes (left panel) followed by a Southern blot using a probe to *TRP1* (right panel). The probe binds to the *TRP1* marker on the YAC (~1500 kb) as well as the *trp1-289* allele on chromosome IV. The 801B6 YAC contains Flex1 (AT)34 by PCR and sequencing. (C) Diagram of YACs containing human chromosome 16 sequences. Chromosome 16 boxes are lined up according to their genomic coordinates. (D) S1 nuclease cleavage assay on plasmids (related to middle panel of Figure 1C). Full gel image of a representative 1% agarose gel showing S1 nuclease cleavage titration (0U, 1U, 1.75U, 2.5U and 5U) of plasmids containing indicated Flex1 or control sequences.

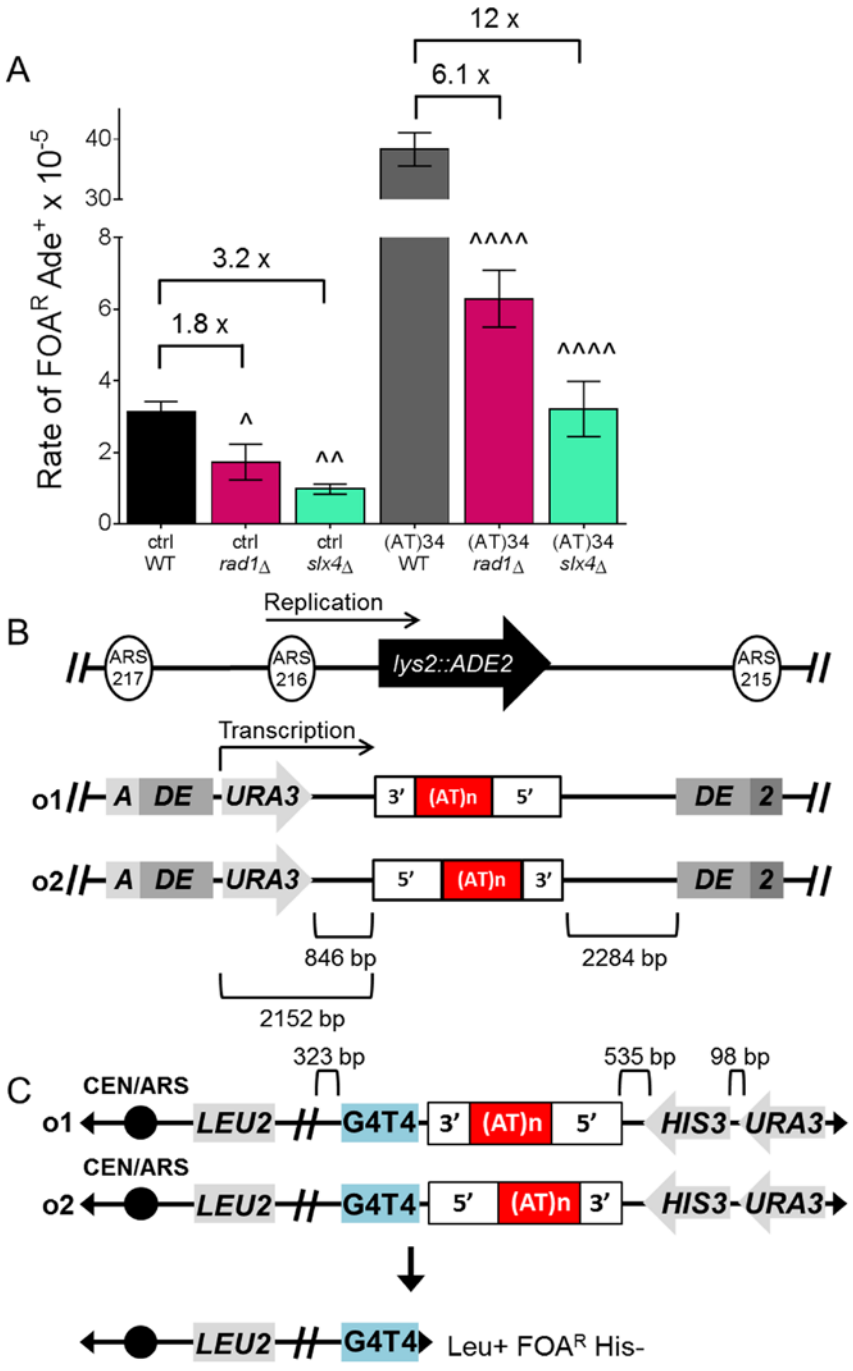


Figure S2. Fragility Assay constructs and DDRA results for *slx4*Δ and *rad1*Δ strains (related to Figures 1-5). (A) DDRA rates for *slx4*Δ and *rad1*Δ strains. Fold decreases compared to WT and statistical decrease compared to WT values using an unpaired t-test are indicated ^ $p<0.05$, ^^ $p<0.01$, ^^^ $p<0.001$, and ^^^^ $p<0.0001$. Rates in Table S2. (B) A detailed depiction of the DDRA fragility assay cassette at the *LYS2* locus on chromosome II is shown. (C) A detailed depiction of the YAC end loss assay and the yeast artificial chromosome showing Flex1 in orientations 1 (o1) and 2 (o2).

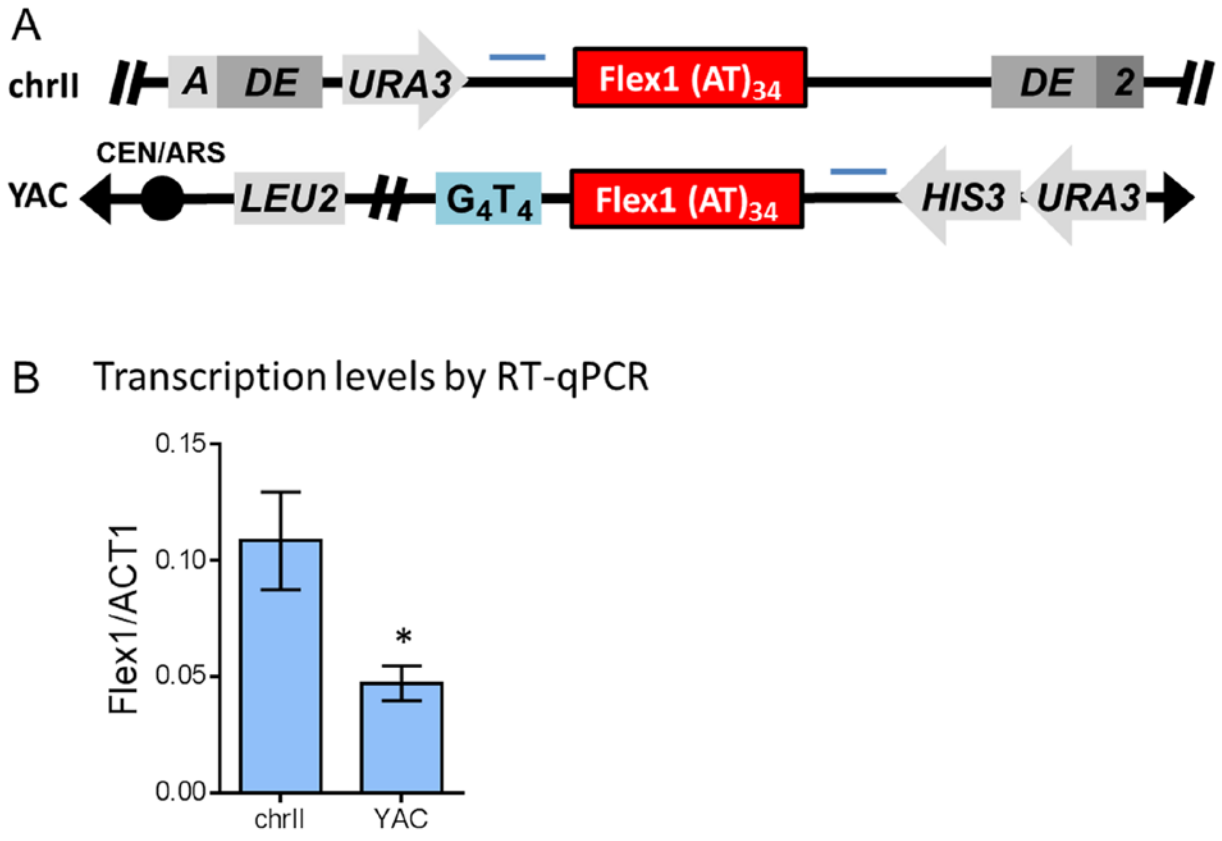


Figure S3. Flex1 is transcribed at the both the chromosome II DDRA locus and the YAC locus (related to Figure 1D). (A) Blue bars indicate area amplified from cDNA and quantified by qPCR using primers 1254 and 1255 (Table S4). Primer locations were chosen based on previous data that transcription arises from read-through of the *URA3* gene (Su and Freudenreich, 2017). (B) Flex1 transcripts as detected by RT-qPCR, normalized to ACT1. Data are from 3 separate RNA preparations with 1-2 separate cDNA and qPCR preparations per RNA sample (see Table S7). Chromosome II and YAC strains used are Flex1 S5'(AT)34S3' orientation 1. *p <0.05 compared to chrII.

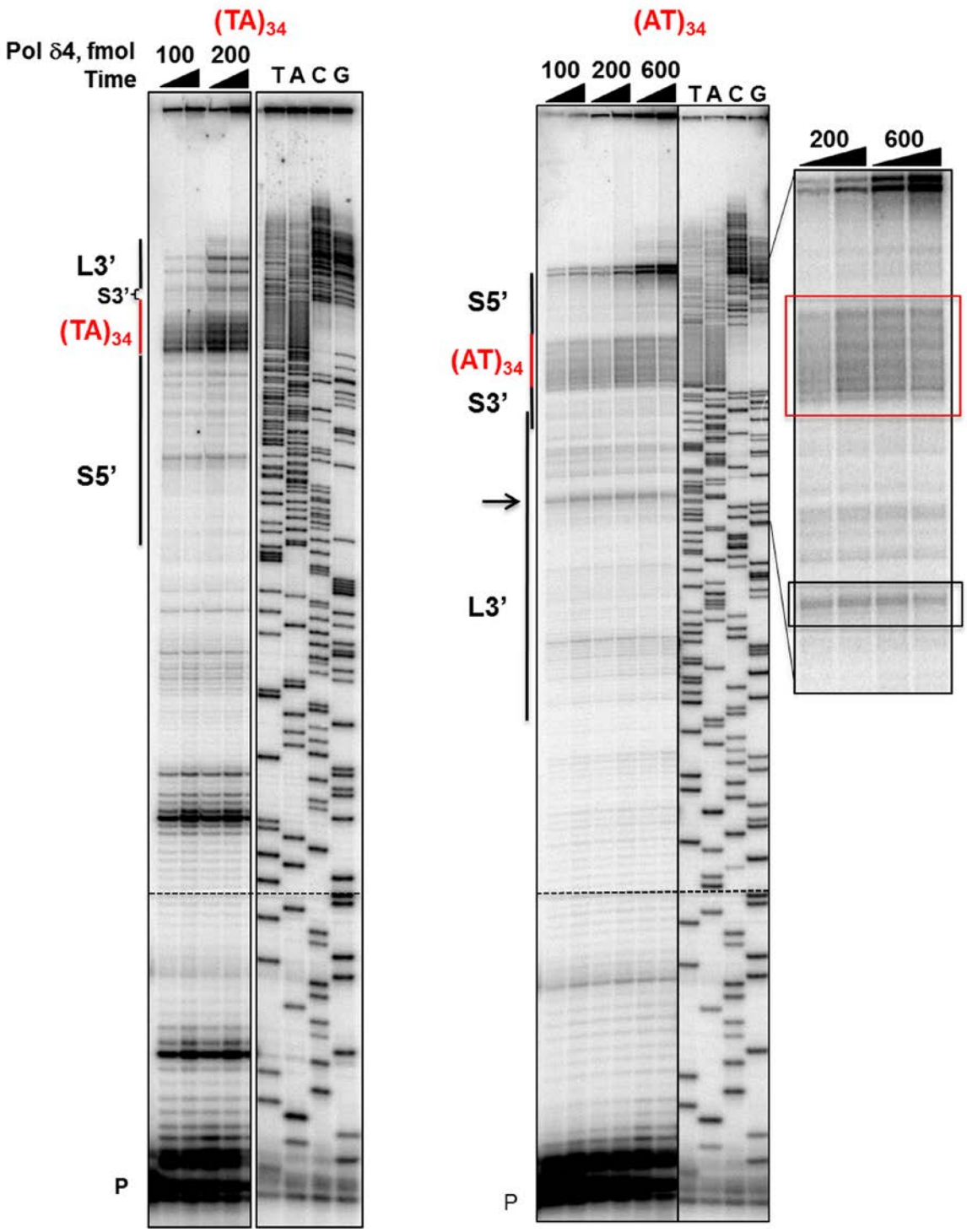


Figure S4. Flex1 (AT)34 stalls human polymerase delta on both DNA strands (related to Figures 2 and 5). *In vitro* DNA synthesis of Flex1 with (AT)34 and a L3' flanking sequence by the 4-subunit human polymerase δ holoenzyme (Pol δ4), showing pause sites at the predicted hairpin in the L3' sequence (arrow and black boxed area, right-hand gel). Pausing at the AT run is evident whether the (TA)34 or (AT)34 repeat is the template strand. Sequence outside of the marked area is composed of the plasmid backbone. TACG, dideoxy sequencing ladder of the DNA template.

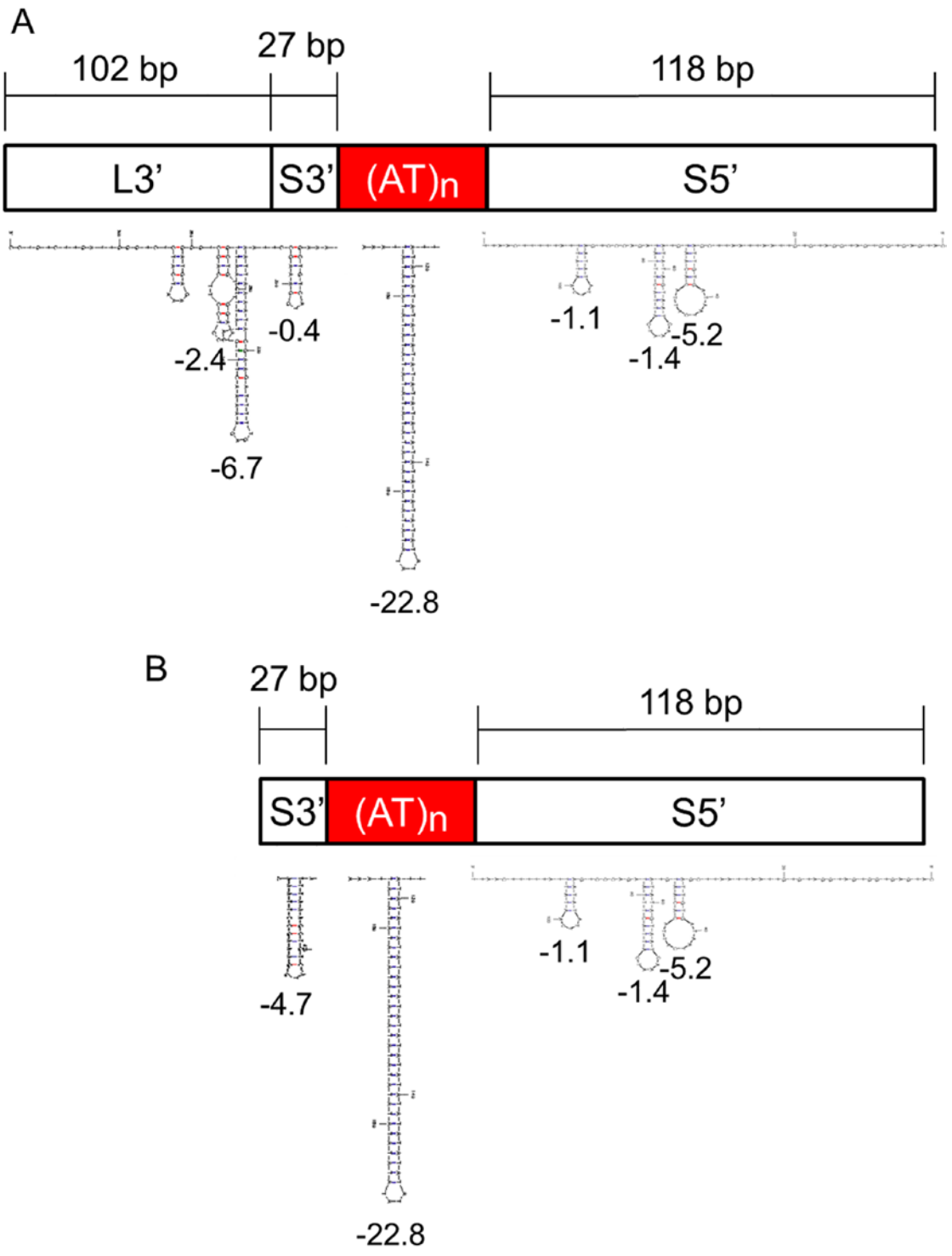


Figure S5. Secondary structure predictions for Flex1 with various flanking sequences (related to Figure 5). Secondary structure predictions for sequences contained within Flex1 with a L3' (A) and S3' (B) flanking sequence by MFold. ΔG values of each predicted hairpin are reported below the structure using folding conditions: 37°C folding temperature, 1 mM Na^+. Note that the sequence between hairpins is non-contiguous for illustration purposes.

Table S1. % FOA^R colonies for large FRA16D YACs. Related to Figure 1A.

YAC strain	# of Experiments	Average % FOA^R	SEM	p value	p compared to
972D3	3	4.1	0.2646		
801B6	6	18.1	1.4241	0.0003	972D3
801B6 Flex1Δ	5	12.9	0.5687	0.01613	801B6

Table S2. DDRA fragility assay data. Related to Figures 1B, 2C, 3, 4, 5A, 5B, 5D and S2A.

FRA16D sequence	Deleted gene(s) or treatment	# of Experiments	Average FOA^R x 10^{-5}	SEM	p value	p compared to
ctrl		6	3.1	0.2883		
ctrl	+HU	4	22.7	5.4501	0.0020	ctrl
Flex1 (AT)14		3	4.5	0.6028	0.0499	ctrl
Flex1 (AT)14	+HU	3	26.3	1.6586	0.0002	Flex1 (AT)14
Flex1 (AT)23		5	16.2	1.2178	<0.0001	ctrl
Flex1 (AT)23	+HU	3	31.9	1.2785	0.0002	Flex1 (AT)23
Flex1 (AT)34		7	38.3	2.7815	<0.0001	ctrl
Flex1 (AT)34	+HU	3	122.4	37.0016	0.0058	Flex1 (AT)34
ctrl	*mus81*Δ	3	3.8	0.3180	0.2203	ctrl
Flex1 (AT)14	*mus81*Δ	3	4.2	0.3606	0.6913	Flex1 (AT)14
Flex1 (AT)23	*mus81*Δ	3	3.9	0.1667	0.0003	Flex1 (AT)23
Flex1 (AT)34	*mus81*Δ	3	13.8	0.7219	0.0005	Flex1 (AT)34
Flex1 (AT)34	*mms4*Δ	4	6.7	0.9127	<0.0001	Flex1 (AT)34
Flex1 (AT)34	*mms4-9A*	5	36.5	3.6398	0.6983	Flex1 (AT)34
Flex1 (AT)34	*rtt107*Δ	5	27.5	3.6814	0.0372	Flex1 (AT)34
Flex1 (AT)34	*rtt107*Δ *mms4-9A*	4	35.7	5.0382	0.6258	Flex1 (AT)34
Flex1 (AT)34	*rad51*Δ	3	37.2	0.57735	0.8070	Flex1 (AT)34
Flex1 (AT)34	*rad51*Δ *mus81*Δ	3	17.4	4.0720	0.0032	Flex1 (AT)34
Flex1 (AT)34	*yen1*Δ	4	80.9	10.3907	0.0007	Flex1 (AT)34
Flex1 (AT)34	*slx1*Δ	3	29.3	0.7881	0.0750	Flex1 (AT)34
Flex1 (AT)34	*rad1*Δ	3	6.3	0.7937	<0.0001	Flex1 (AT)34
Flex1 (AT)34	*slx4*Δ	5	3.2	0.7736	<0.0001	Flex1 (AT)34

Flex1 (AT)34	*mus81*Δ *yen1*Δ	3	12.3	1.2583	0.0004	Flex1 (AT)34
Flex1 (AT)34	*mus81*Δ *slx1*Δ	3	10.4	2.5989	0.0003	Flex1 (AT)34
ctrl	*yen1*Δ	3	3.3	0.2517	0.7238	ctrl
ctrl	*slx1*Δ	3	2.9	0.4583	0.6659	ctrl
ctrl	*rad1*Δ	3	1.7	0.4978	0.0341	ctrl
ctrl	*slx4*Δ	3	1.0	0.1362	0.0016	ctrl
ctrl	*sae2*Δ	3	3.3	1.0817	0.8455	ctrl
Flex1 (AT)34	*sae2*Δ	4	7.6	1.2743	<0.0001	Flex1 (AT)34
Flex1 (AT)34 S3’ o2		4	11.6	2.4052	<0.0001	Flex1 (AT)34
Flex1 (AT)34 L3’ o2		9	2.1	0.5431	0.0002	Flex1 (AT)34 S3' o2
I-SceI only		6	344.5	81.6916		
I-SceI S3’		7	385.3	57.1993	0.6832	I-SceI only
I-SceI L3’		7	98.3	41.6164	0.0016	I-SceI S3'

All Flex1 constructs contain the S3’ flanking sequence and are in orientation 1 unless otherwise noted.

Table S3. YAC fragility assay data. Related to Figures 4B and 5C.

FRA16D sequence	**Deleted gene(s)**	**# of Experiments**	**Average FOA^R His^+ x 10^{-6}**	**SEM**	**p value**	**p compared to**
Flex1 (AT)34 S3' o1		3	11.1	0.9207	0.0167	(AT)23-S3' o1
Flex1 (AT)34	*yen1*Δ	3	12.2	1.3043	0.5287	(AT)34-S3' o1
Flex1 (AT)34	*mus81*Δ	3	5.8	0.4177	0.0061	(AT)34-S3' o1
Flex1 (AT)34	*slx1*Δ	3	6.5	0.9244	0.0232	(AT)34-S3' o1
Flex1 (AT)34	*rad1*Δ	3	4.6	1.0366	0.0094	(AT)34-S3' o1
Flex1 (AT)34	*slx4*Δ	3	5.9	0.5859	0.0087	(AT)34-S3' o1
Flex1 (AT)34	*mus81*Δ *rad1*Δ	5	5.2	0.9528	0.0108	(AT)34-S3' o1
Flex1 (AT)34	*slx1*Δ *rad1*Δ	3	7.8	1.9150	0.2511	(AT)34-S3' o1
Flex1 (AT)34	*slx1*Δ *mus81*Δ *rad1*Δ	5	7.7	2.4192	0.3402	(AT)34-S3' o1
Flex1 (AT)23 S3'		3	6.4	0.7513		

Flex1 (AT)34 L3'		3	0.3	0.0876	0.0003	(AT)34-S3' o1
Flex1 (AT)34 S3' o2		3	15.4	1.3528	0.0044	(AT)23-S3' o1
Flex1 (AT)34 L3' o2		3	14.4	5.0560	0.8578	(AT)34-S3' o2

All Flex1 constructs contain the S3' flanking sequence and are in orientation 1 unless otherwise noted.

Table S4. **Oligonucleotides**. Related to STAR Methods.

Oligo Name	**CF Oligo Stock #**	**Purpose**	**Sequence**
ura3rev	3	Check for *URA3* absence	TCCTGTTGCTGCCAAGCTAT
ura3rev	4	Check for *URA3* absence	TCCCAGCCTGCTTTTCTGTA
ura3for2	5	Check for *URA3* absence	TGCTGCTACTCATCCTAG
URA3 internal reverse	1223	Check for *URA3* absence	GCTTAACTGTGCCCTCCATGG
RT-PCR_F1_upstream_for	1254	To measure levels of Flex1 transcripts	AACTGTTGGGAAGGGCGATC
RT-PCR_F1_upstream_rev	1255	To measure levels of Flex1 transcripts	TGAGTCGTATTACAATTCACTGGC
TRP1_222b_int_for	1711	Southern TRP1 probe for	GGCGTGTTTCGTAATCAACC
TRP1_127bp_int_rev	1712	Southern TRP1 probe rev	GGCGTCAGTCCACCAGCTAA
P1_for_252bp_chk	1807	FRA16D amplicon 1 for	GCATATGAGAATACTCATACT CAG TGCTGC
P1_110bp_chk	1704	FRA16D amplicon 1 rev	CCATGCACTCTGGTGTACCA
P3_for_642bp_chk	1840	FRA16D amplicon 2 for	GTGTGAATACCAGGTGGTAGGGATTATGTG
P3_rev_120bp_chk	1841	FRA16D amplicon 2 rev	ACAGAACTAACCCAGAGATGGTTTCTCATC
F5His_For	1545	FRA16D amplicon 3 for	GGGAGTCCTAGATCAAGGTG
P4_rev_752bp_chk	1809	FRA16D amplicon 3 rev	GAACTCAGATAAAGATAAGGCCTATGGTTC
P5P5B_for_672bp_chk	1810	FRA16D amplicon 4 for	AAAACTTTGCTGGAGAACATCACCAATCAC

P5P5B_rev_428bp_chk	1811	FRA16D amplicon 4 rev	TTCTGAGAAACTGTCACAGCCAAGAAGATG
F1_420down	1267	Checking Flex1::KANMX6 in FRA16D YAC	GCTGAAGTCACAAGATCTTAGGATGGGGTG
pBL007for	679	Screening for pBL007 transformants with insert	AAGCATATTTGAGAAGATGCGGCCAGC
pBL007rev	680	Screening for pBL007 transformants with insert	GGAATAAGGGCGACACGGAAATGTTGA
Flex1_pBL007_seq_For	1032	PCR and sequencing of insert in pBL007 and chrII locus	ACTCACTATAGGGCGAATTG
Flex1_pBL007_seq_Rev	1033	PCR and sequencing of insert in pBL007 and chrII locus	CCAACTGATCTTCAGCATCT
5'LYS2_pBL007_integr_For	1028	PCR of 5’ cassette in chrII locus	AAGTAACAAGCAGCCAATAG
5'LYS2_pBL007_integr_Rev	1029	PCR of 5’ cassette in chrII locus	CATGTGTCAGAGGTTTTCAC
3'LYS2_pBL007_integr_For	1030	PCR of 3’ cassette in chrII locus	CTCGGAATTAACCCTCACTA
3'Lys2junctionrev	1047	PCR of 3’ cassette in chrII locus	GCAAAGTGGTGATAGAGTTC
T7	2	PCR and sequencing of insert in pHZ-HIS3MX6 and YAC	TAATACGACTCACTATAGGG
M13R	1343	PCR and sequencing of insert in pHZ-HIS3MX6 and YAC	CAGGAAACAGCTATGACC
His3Revsk	375	PCR from *HIS3MX6* to *URA3* to confirm modified YAC	TTAGATAAATCGACTACGGCAC
URA3 for	832	PCR from *HIS3MX6* to *URA3* to confirm modified YAC	CAGTACTCTGCGGGTGTATACAG
ILV1_for	1465	PCR of 5’ junction of pGAL-I-SceI nuclease	CTCTGCGCTATATCTTTGGG

		cassette	
GAL1,10_chk	1466	PCR of 5’ junction of pGAL-I-SceI nuclease cassette	CGCTTCGCTGATTAATTACCCCAG
I-SceI_for2	1511	Creation of I-SceI insert for cloning (3’ end anneals to 1512)	gatctaGAATTCggtactgcgggatatcgtccattccgacagTAGGGATAACAGGGTAAT
I-SceI_rev2	1512	Creation of I-SceI insert for cloning (3’ end anneals to 1511)	tatcgaGAATTCagcgcgacgtcgcttgcggtattcggATTACCCTGTTATCCCTActgt
I-SceI_for2_short	1513	Creation of I-SceI insert for cloning	gatctaGAATTCggtactgc
I-SceI_rev2_short	1514	Creation of I-SceI insert for cloning	tatcgaGAATTCagcgcgac
M13 Forward (-20)	n.a.	Pol δ4 polymerase pausing assay	GTAAAACGACGGCCAG
G40	n.a.	Pol δ4 polymerase pausing assay	GCATGCCTGCAGGTCG
G40-16mer, PAGE-purified	n.a.	Pol δ4 polymerase pausing assay	GCATGCCTGCAGGTCG
gBlock	n.a.	EcoRI-S5-I-SceI-S3-EcoRI	AGCGTAGAATTCTGTTACCATGAGTGGTGATGGATGTGTTAATTAATTCGATTGTGATAATCATTACACAATGTATATAGTAATCAAATCATTACTTTATAGACCCTGAATATATTCAATATTTATTTTTCAATTTAGGGATAACAGGGTAATTTAAAGCTGTCATGGAAAGCCTTAAAGCAGTATGAATTCTCTGAC
gBlock	n.a.	EcoRI-S5-ISceI-L3-EcoRI	AGCGTAGAATTCTGTTACCATGAGTGGTGATGGATGTGTTAATTAATTCGATTGTGATAATCATTACACAATGTATATAGTAATCAAATCATTACTTTATAGACCCTGAATATATTCAATATTTATTTTTCAATT

			TAGGGATAACAGGGTAATT TAAAGCTGTCATGGAAAGC CTTAAAGTTAAAATACGAA GATTTTTGAGAAAAACTTT GCATATTTTAATTGCTGTCT GGAATCCTCCTTCAGCTGG GATGAGAAATCATCTCTGG GTTAGTTCTGTCCCAGTATG AATTCTCTGAC

Table S5. Plasmids. Related to STAR Methods.

Plasmid	**CF Plasmid stock#**	**Description**	**Source**
pFA6a-KANMX6	136	Template for one-step gene replacement by PCR	(Wach et al., 1994)
pBL007	223	*ADE2* nt 512-1480 *URA3*	this study
pBL007+ctrl	387/388	*ADE2* nt 512-1480 *URA3*-EcoRI-ctrl-BamHI	this study
pBL007+S5'-(AT)14-S3' o1	565/566	*ADE2* nt 512-1480 *URA3*-EcoRI-Flex1(AT)14-EcoRI	this study
pBL007+S5'-(AT)23-S3' o1	516/517	*ADE2* nt 512-1480 *URA3*-EcoRI-Flex1(AT)23-EcoRI	this study
pBL007+S5'-(AT)34-S3' o1	351	*ADE2* nt 512-1480 *URA3*-EcoRI-Flex1(AT)34-EcoRI	this study
pHZ-*HIS3MX6*	466	G_4T_4 *HIS3MX6 URA3*	this study
pHZ-*HIS3MX6*+S5'-(AT)34-S3' o1	513	G_4T_4 *HIS3MX6 URA3* EcoRI-Flex1(AT)34-S3' o1-EcoRI	this study
pHZ-*HIS3MX6*+S5'-(AT)34-S3' o2	559, 560	G_4T_4 *HIS3MX6 URA3* EcorI-Flex1(AT)34-S3' o2-EcoRI	this study
pHZ-*HIS3MX6*+S5'-(AT)34-L3' o2	512	G_4T_4 *HIS3MX6 URA3* EcoRI-Flex1(AT)34-L3' o1-EcoRI	this study

pBL007+I-SceI	519	*ADE2* nt 512-1480 *URA3*-EcoRI-I-SceI-EcoRI	this study
pBL007+S5'-I-SceI-S3'	571	*ADE2* nt 512-1480 *URA3*-EcoRI-Flex1 S5'-I-SceI-S3'-EcoRI	this study
pBL007+S5'-I-SceI-L3'	581	*ADE2* nt 512-1480 *URA3*-EcoRI-Flex1 S5'-I-SceI-L3'-EcoRI	this study
pGSHU	524	pFA6a-pGAL1-I-SceI-HYG-klURA3	(Storici et al., 2003)

Table S6. AT series termination probability. Related to Figures 2A and 2B.

	AT14			**AT23**			**AT34**		
	% synthesis	**Term. Prob.**		**% synthesis**	**Term. Prob.**		**% synthesis**	**Term. Prob.**	
		AT only	**S5'(AT) S3'**		**AT only**	**S5'(AT) S3'**		**AT only**	**S5'(AT) S3'**
Rep. 1	96	0.25	0.74	119	0.40	0.77	115	0.73	0.87
Rep. 2	86	0.25	0.71	97	0.28	0.61	113	0.64	0.82
Rep. 3	87	0.21	0.68	83	0.27	0.56	113	0.62	0.83
AVG	90	0.24	0.71	100	0.32	0.65	114	0.66	0.84
s.d.	5.5	0.023	0.028	18	0.069	0.11	1.15	0.059	0.026

One way ANOVA values are as follows: AT14 vs. AT23 p= 0.2573, AT14 vs. AT34 p= 0.0002, AT23 vs AT34 p= 0.0006.

Table S7. RT-qPCR data. Related to Figure S3 and STAR Methods RT-PCR experiment.

Flex1 locus	**# of Experiments**	**Mean transcript levels**	**SEM**	**p value**	**p compared to**
chrII	5	0.11	0.0210		
YAC	5	0.05	0.0074	0.025	chrII