Translesion polymerases drive microhomology-mediated break-induced replication leading to complex chromosomal rearrangements

Cynthia J. Sakofsky [1] *, Sandeep Ayyar [2, 4] *, Angela K. Deem[2, 5], Woo-Hyun Chung[3, 6], Grzegorz Ira[3], and Anna Malkova [1 #]

[1] Department of Biology, University of Iowa, Iowa City, IA 52242

[2] Indiana University Purdue University Indianapolis (IUPUI), Indianapolis, IN 46202

[3] Department of Molecular & Human Genetics, Baylor College of Medicine, Houston, Texas 77030

[4] Current address: Department of Biomedical Informatics, Stanford University, CA 94305

[5] Current address: Department of Genomic Medicine, University of Texas MD Anderson Cancer Center, Houston, Texas 77230

[6] Current address: College of Pharmacy, Duksung Women's University, Seoul 132-714, Republic of Korea

[#] **Contact:** anna-malkova@uiowa.edu

* These authors contributed equally to this work.

FIGURE S1

A.


(i)
3'
5'
ADE1
5'
ADE3
3'
STOP
(ii)
ADE1
ADE3
(iii)
ADE1
LYS+
ADE3

B.

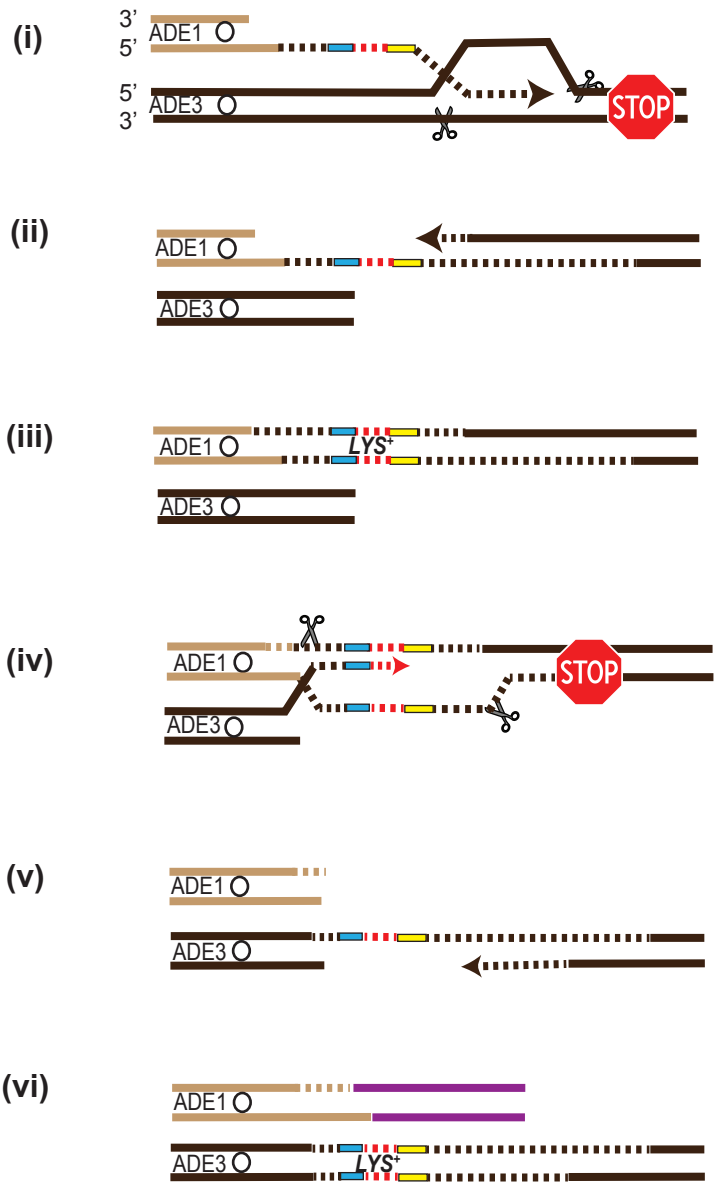

(i)
3'
5'
ADE1
5'
ADE3
3'
STOP
(ii)
ADE1
ADE3
(iii)
ADE1
LYS+
ADE3
(iv)
ADE1
STOP
ADE3
(v)
ADE1
ADE3
(vi)
ADE1
ADE3
LYS+

FIGURE S2

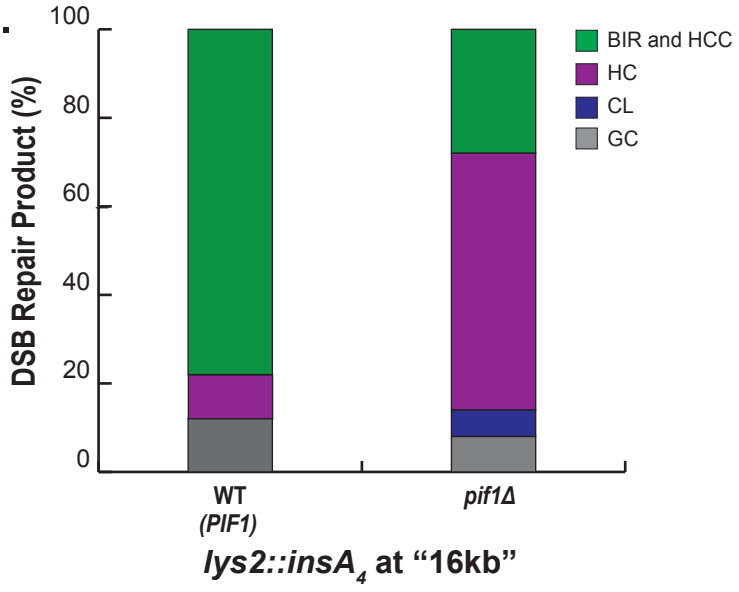
A.

Relevant Phenotype	*lys2::insA*$_4$ at MAT									*lys2::insA*$_4$ at 16kb	
	WT (PIF1)	*pif1Δ*	*pif1Δ rev3Δ*	*pif1Δ rev1Δ*	*pif1Δ rev1-cd*	*pif1Δ rev3Δ tlc1Δ*	*pif1Δ tlc1Δ*	*pif1Δ rad30Δ*	*pol32Δ*	*WT (PIF1)*	*pif1Δ*
Ade^{+} Leu^{+}	36	5	0	0	2	0	0	3	2	220	42
Ade^{+}Leu^{-}	1724	528	118	147	181	96	148	144	47	1511	310
Ade^{-red} Leu^{-}	50	128	44	59	67	62	71	39	102	5	30
Ade^{-white} Leu^{-}	165	229	47	67	72	46	71	56	29	92	158

B.

Chr III Structure	*lys2::insA*$_4$ at MAT									*lys2::insA*$_4$ at 16kb	
	WT (PIF1)	*pif1Δ*	*pif1Δ rev3Δ*	*pif1Δ rev1Δ*	*pif1Δ rev1-cd*	*pif1Δ rev3Δ tlc1Δ*	*pif1Δ tlc1Δ*	*pif1Δ rad30Δ*	*pol32Δ*	*WT (PIF1)*	*pif1Δ*
Rearranged	2 [8%]	16 [29%]	4 [27%]	4 [27%]	3 [20%]	4 [27%]	3 [14%]	3 [20%]	4 [27%]	2 [8%]	9 [31%]
Not Rearranged	24 [92%]	39 [71%]	11 [73%]	11 [73%]	12 [80%]	11 [73%]	19 [86%]	12 [80%]	11 [73%]	23 [92%]	20 [69%]

C.



D.

Relevant Phenotype	*lys2::insA*$_4$ at MAT	
	WT (*PIF1*)	***pif1Δ***
Ade^{+} Leu^{+}	80 [1%]	4 [0.5%]
Ade^{+}Leu^{-}	5586 [96%]	804 [58.5%]
Ade^{-red} Leu^{-}	83 [2%]	266 [19%]
Ade^{-white} Leu^{-}	77[1%]	298 [22%]

E.

```
ORIGINAL  AGTGTTTGCCCACGTCAGATCCTGGAAAACGGGAAAGGTTCCGTTCAGGACGCTACTTGTGTATAAGAGTCAGCGTCAGGGCCAAGGATGAA
2193B7    AGTGTTTGCCCACGTCAGATCCTGGAAAACGGGAAAGGTTCCGTTCAGGACGCTACTTGTGTATCGTCCTGAACGGA-----CCAAGGATGAA
```

F.

```
ORIGINAL  AGTGTTTGCCCACGTCAGATCCTGGAAAACGGGAAAGGTTCCGTTCAGGACGCTACTTGTG--TATAAGAGTCAGCGTCAGGGCCAAGGATGAA
DGC_23    AGTGTTTGCCCACGTCAGATCCTGGAAAACGGGAAAGGTTCCGGGATCTGACGTGGGCAAACATATAAGAGTCAGCGTCAGGACCAAGGATGAA
```

SUPPLEMENTAL DATA

Figure S1. Various mechanisms for the formation of half-crossover events. Related to Figure 4. **(A)** HCC with a rearranged donor chromosome and *LYS2* in the recipient. (i) Resolution of D-loop in BIR intermediate containing MMBIR mutation. (ii) Formation of HC event. (iii) Stabilization of broken donor by ectopic recombination or *de novo* telomere formation as represented by purple lines. **(B)** HCC with a rearranged recipient chromosome and *LYS2* in the donor. Steps (i and ii) are the same as in (A). (iii-v) invasion of the broken donor into a HC product and resolution of the intermediate leading to the formation of secondary HC and a broken *ADE1*-containing fragment. (vi) Stabilization of *ADE1*-containing fragment by ectopic recombination or *de novo* telomere formation as represented by purple lines.

Figure S2. Details of the distribution of DSB repair outcomes. Related to Figures 1 and 3. The results of phenotypic analysis of DSB repair outcomes **(A)** and PFGE of representative Ade^+ Leu^- outcomes showing the percent rearrangements in chromosome III (represent HCC events) **(B).** Data from (A) were used to calculate final distribution of DSB outcomes shown in **(C)** and in Figure 1B (See Experimental Procedures for details of calculations). **(D)** Phenotypic analysis of Lys^+ mutants used to calculate the distribution of repair outcomes shown in Figure 3D. **(E and F)** Two examples of MMBIR events obtained in *pif1Δ* strains with reporters at *MAT* (E) and at 16kb (F) that contain imperfect microhomologies near the position of the first template switch. Imperfect microhomologies (include microhomologies located $\geq$1bp from the junctions

and mismatched bases) are underlined. Colors of letters are like in Figure 3A. See Table S4, S5, and S6 for a full list and description of mutations.

SUPPLEMENTAL EXPERIMENTAL PROCEDURES

Yeast Strains

All yeast strains (Table S1) are isogenic to AM1003 strain that is disomic for chromosome III (Deem et al., 2008). In these strains, HO-induced DSBs were introduced into a truncated copy of chromosome III (recipient) by the addition of galactose. Control strains lacking an HO-endonuclease recognition cut site were obtained by plating on YEP-Gal media followed by selection of Ade^+ Leu^+ colonies, resulting from DSB repair by gene conversion (Deem et al., 2011). Strains containing *rev3Δ::BSD*, *rev1Δ::BSD*, or *tlc1Δ::BSD* disruptions were constructed by transformation with a PCR-derived blasticidin (BSD) marker (TEF/BSD from Invitrogen) flanked by terminal sequences matching the first and last 80bp of the open reading frame of each gene (Wach et al., 1994) and were confirmed by PCR and phenotypic analysis. In the case of strains containing *tlc1Δ::BSD* disruption, freshly obtained independent transformants were used in experiments immediately upon confirmation of *TLC1* disruption. Strains containing *rad30::ble*^{r} were constructed by transformation with PCR-derived phleomycin-resistant (*ble*^{r}) cassette (Gueldener et al., 2002) flanked by terminal sequences matching the first and last 80bp of the open reading frame of *RAD30* gene (Wach et al., 1994) and were confirmed by PCR. Strains containing *pif1Δ::KANMX* were constructed by transformation with PCR-derived *KANMX* module similar to (Wilson et al., 2013). To construct strain AM2268 containing *rev3Δ::URA3*, the strain AM1523 containing

rev3Δ::KANMX (Deem et al., 2011) was transformed with M3927 plasmid similar to (Voth et al., 2003). Strains with *pif1-m2* mutations were constructed using plasmid pVS31 (Schulz and Zakian, 1994) and confirmed by sequencing similar to (Wilson et al., 2013). Strains containing a catalytic mutation of *REV1* (*rev1-cd*) were constructed using the "pop-in-pop out" method with plasmid pRS306-rev1-cd (Northam et al., 2014) that was digested with *EcoRI* for genomic integration (pop-in step) followed by selection of Ura^- (pop-out step). Strains with *rev1-cd* mutation were confirmed by Sanger sequencing.

Analysis of Chromosome III Structure

Chromosome rearrangements were determined by physical analysis of chromosome III in $Ade^+Leu^-Lys^-$ and $Ade^+Leu^-Lys^+$ strains using PFGE and hybridization with *ADE1*-, and *ADE3*-specific probes (Deem et al., 2008) and *LYS2*- specific probes (Saini et al., 2013). Rearrangements were defined as chromosomes that deviated from their expected size of 355kb (donor chromosome) and 345kb (BIR repair product) (See Figure 3E for example of gel image and hybridization results).

Distribution of DSB repair events

To determine the distribution of DSB repair events among non-selected outcomes, yeast cultures were grown in leucine drop-out medium for ~20 hours and then diluted and plated on YEP-Gal plates similar to (Deem et al 2008). Colonies that formed on YEP-Gal plates were then replica plated onto appropriate omission media to determine the fraction of DSB repair events with the following phenotypes: Ade^+ Leu^+(GC), Ade^- ^{white} Leu^- (HC), Ade^{-red} Leu^- (CL), and Ade^+ Leu^- (Figure S2A). HC events included Ade^-

^{white} Leu^{-} outcomes as well as a portion of Ade^{+} Leu^{-}. The number of Ade^{+} Leu^{-} HCs was assumed to equal the number of Ade^{-white} Leu^{-} outcomes based on random segregation during mitosis (similar to (Deem et al., 2008; Wilson et al., 2013; Smith et al., 2009)). Therefore, to determine total HC events, the number of Ade^{-white} Leu^{-} outcomes was multiplied by two and Ade^{+} Leu^{-} events were then adjusted by subtracting the number of Ade^{-white} Leu^{-} colonies. The remaining Ade^{+} Leu^{-} outcomes included both HCC and BIR events, which were phenotypically indistinguishable and thus remained grouped together. The presence of HCCs in this group was verified for a representative set of Ade^{+} Leu^{-} events by PFGE that detected rearrangements in chromosome III (Figure S2B).

To determine the frequency of various DSB repair outcomes among Lys^{+} mutations, yeast were first grown on lysine drop-out media to obtain mutation rates, followed by replica plating of Lys^{+} colonies onto appropriate omission media to determine the phenotypes of the DSB repair outcomes. The fraction of each repair event was then multiplied by the Lys^{+} mutation rate (Figure 3D, S2D).

The efficiency of *Gal::HO* induction was confirmed by determining the fraction of Ade^{+} Leu^{-} events among Ade^{+} in all strains with a DSB cut site in at least three independent experiments per strain. Ade^{+} colonies obtained from cells plated on adenine drop-out medium at 7hrs after the addition of galactose were replica-plated onto leucine drop-out plates and the proportion of Ade^{+} Leu^{-} was determined using $\geq$ 30 colonies per experiment (Table S7).

Analysis of Lys$^+$ mutation spectra

Lys$^+$ mutation spectra were determined by amplification of a portion of the *LYS2* gene followed by Sanger sequencing using primers described in (Deem et al 2011). Lys$^+$ mutations were identified using the Codon Code Aligner DNA Sequence Analysis Program: http://www.codoncode.com/aligner/. Ade$^+$Leu$^-$Lys$^+$ mutants were further analyzed by PFGE as described above.

SUPPLEMENTAL REFERENCES

Carvalho, C.M., Bartnik, M., Pehlivan, D., Fang, P., Shen, J., and Lupski, J.R. (2012). Evidence for disease penetrance relating to CNV size: Pelizaeus-Merzbacher disease and manifesting carriers with a familial 11 Mb duplication at Xq22. Clin Genet *81*, 532-541.

Carvalho, C.M., Vasanth, S., Shinawi, M., Russell, C., Ramocki, M.B., Brown, C.W., Graakjaer, J., Skytte, A.B., Vianna-Morgante, A.M., Krepischi, A.C., *et al.* (2014). Dosage changes of a segment at 17p13.1 lead to intellectual disability and microcephaly as a result of complex genetic interaction of multiple genes. AM J Hum Genet *95*, 565-578.

Carvalho, C.M., Zhang, F., Liu, P., Patel, A., Sahoo, T., Bacino, C.A., Shaw, C., Peacock, S., Pursley, A., Tavyev, Y.J., *et al.* (2009). Complex rearrangements in patients with duplications of *MECP2* can occur by fork stalling and template switching. Human molecular genetics *18*, 2188-2203.

Gueldener, U., Heinisch, J., Koehler, G.J., Voss, D., and Hegemann, J.H. (2002). A second set of loxP marker cassettes for Cre-mediated multiple gene knockouts in budding yeast. Nucleic Acids Res *30*, e23.

Liu, P., Erez, A., Nagamani, S.C., Bi, W., Carvalho, C.M., Simmons, A.D., Wiszniewska, J., Fang, P., Eng, P.A., Cooper, M.L., *et al.* (2011b). Copy number gain at Xp22.31 includes complex duplication rearrangements and recurrent triplications. Human Mol Genet *20*, 1975-1988.

Voth, W.P., Jiang, Y.W., and Stillman, D.J. (2003). New 'marker swap' plasmids for converting selectable markers on budding yeast gene disruptions and plasmids. Yeast *20*, 985-993.

Wach, A., Brachat, A., Pohlmann, R., and Philippsen, P. (1994). New heterologous modules for classical or PCR-based gene disruptions in *Saccharomyces cerevisiae*. Yeast *10*, 1793-1808.