Sarkies et al.

FANCJ coordinates two pathways that maintain epigenetic stability at G quadruplex DNA.

SUPPLEMENTARY INFORMATION

Sarkies et al.

Supplementary information: Table of contents

Supplementary Figures

Supplementary Tables

Sarkies et al. Figure S1

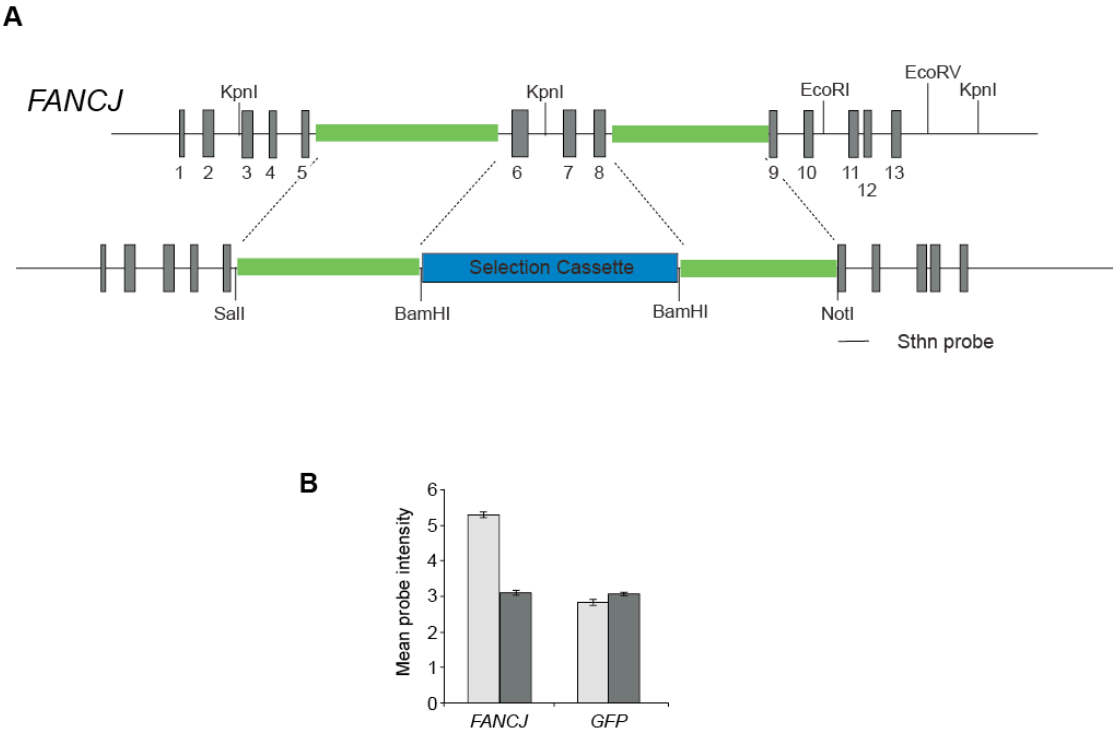


Figure S1. Generation of FANCJ-deficient DT40. A. The targeting construct to remove exons 6 – 8 of the chicken *FANCJ* locus, which disrupts the helicase domain of the protein, was constructed with primers FJ5F and FJ5R to generate the 5' arm and FJ3F and FJ3Y (Table S6) to generate the 3' arm. These were cloned into pBluescript as SalI – BamHI and BamHI – NotI fragments respectively. The drug selection cassette, conferring resistance to either blasticidin or puromycin (55) was inserted into BamHI site. Drug-resistant clones were screened for targeted integration by digestion of genomic DNA with EcoRV and BamHI followed Southern Blotting with a probe 3' of the targeting construct generated with the primers FJPF and FJPR. B. Loss of FANCJ mRNA was confirmed in the microarray analysis. Mean probe intensities for the *FANCJ* locus in wild type (light grey) and *fancj* (dark grey) cells, compared with the mean probe intensity for GFP, a non-expressed control, in both lines.

Sarkies et al. Figure S2

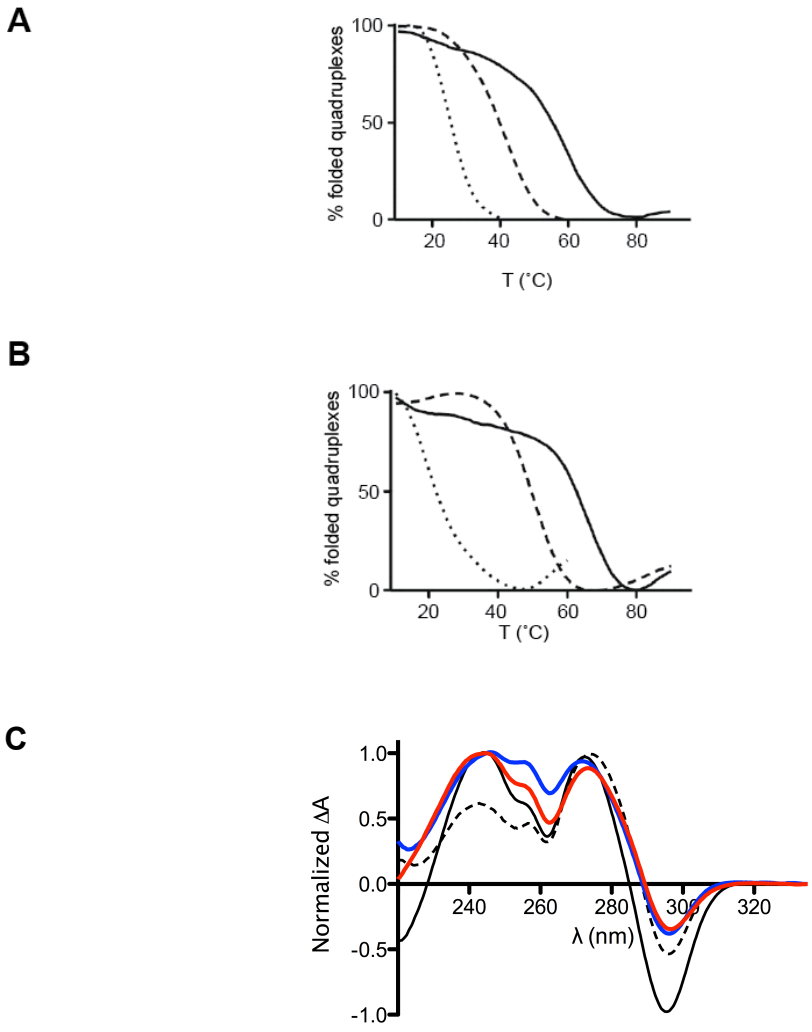


Figure S2. Further biophysical analysis of the *CD72 & BU1A* G4 DNA sequences. A. UV melting curve of the oligonucleotide CD72-F recorded at 295 nm in the presence of LiCl (dotted line), NaCl (dashed line) and KCl (solid line). B. UV melting curve of the oligonucleotide BU1A-F recorded at 295 nm in the presence of LiCl (dotted line), NaCl (dashed line) and KCl (solid line). C. Thermal difference spectra. Normalized differential absorbance signatures of CD72-F (blue), BU1A-F (red), HTelo (black solid line) and C-KIT1 (black dashed line) in presence of KCl. The differential absorbance signatures of CD72-F and BU1A-F are similar (negative minimum at λ = 295 nm and positive maxima at λ = 243, 273 nm) to the two quadruplex motifs formed by HTelo and C-KIT1. Conformational studies have been performed and structural assignments have been used to establish the quadruplex conformation adopted by HTelo (53) and C-KIT1 (54).

Sarkies et al. Figure S3

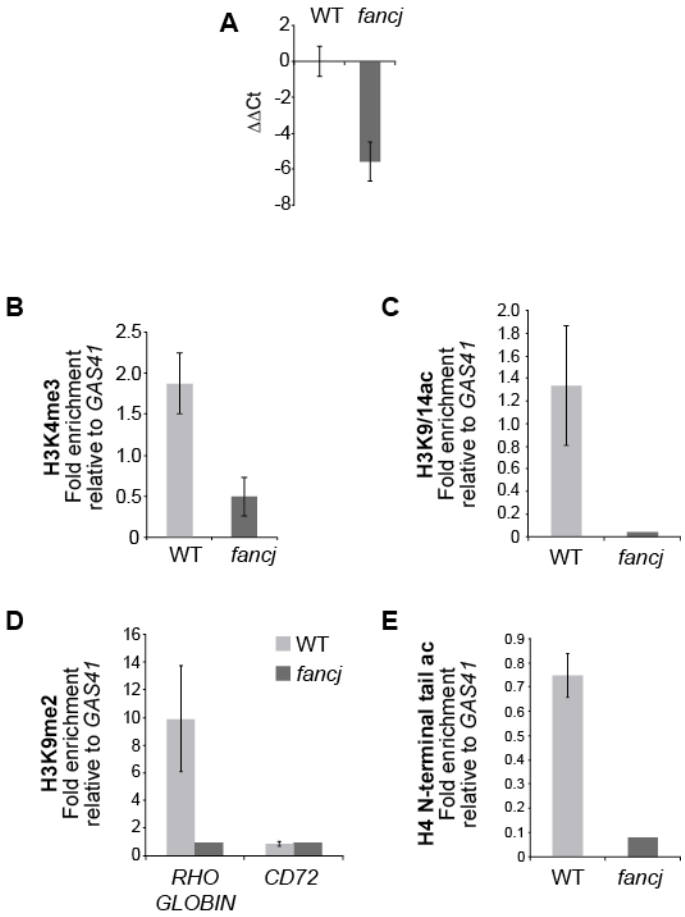


Figure S3. Chromatin changes at the *CD72* locus of *fancj* cells. A. qRT-PCR confirming reduction of CD72 expression in the *fancj* line. Error bars show standard deviation. B. Trimethylation of H3K4 at the *CD72* promoter in wild type (WT) and *fancj* cells. C. Acetylation of H3K9 and K14 at the *CD72* promoter in wild type and *fancj* cells. D. Dimethylation of H3K9 at the *CD72* promoter in wild type and *fancj* cells. The signal at the *RHOGLOBIN* promoter is shown for comparison. E. Acetylation of the H4 N-terminal tail at the *CD72* promoter in wild type and *fancj* cells. Error bars in B - E represent S.E.M..

Sarkies et al. Figure S4

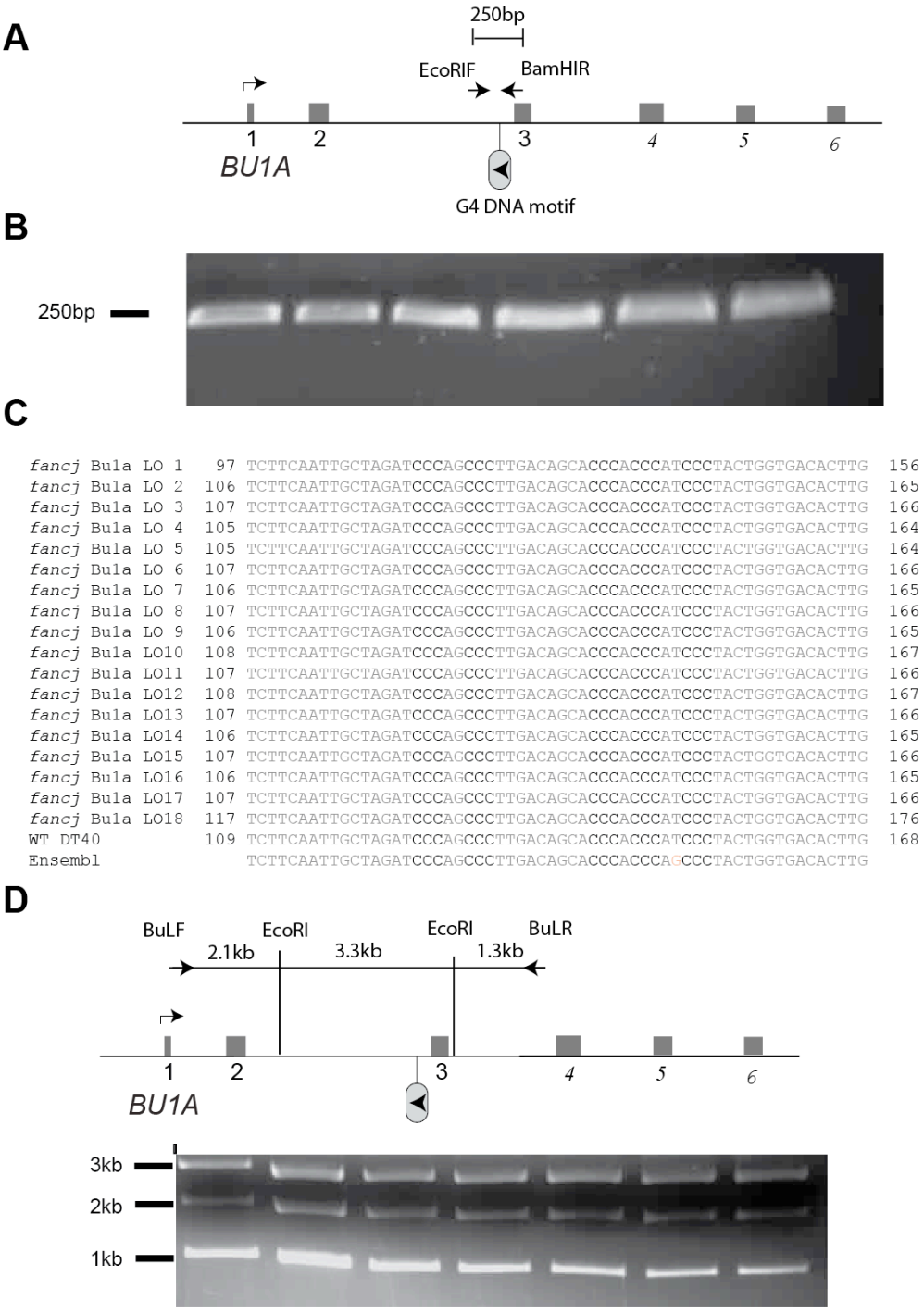
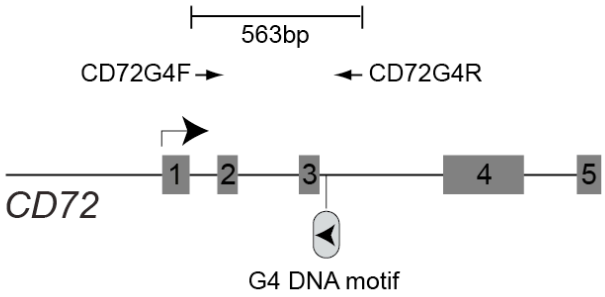


Figure S4. Genetic stability of the *BU1A* locus in *fancj* cells. A. Map of the Bu1a locus indicating the position of the primers to amplify the G-rich region, including the two overlapping G4 DNA motifs. B. Amplification of the G-rich region from 6 *fancj* Bu1a-loss variants. C. Sequences of the G4 DNA motifs from 18 *fancj* Bu1a-loss variant clones. D. Long range PCR of the extracellular domain of Bu1a from 6 *fancj* Bu1a-loss variant clones. The initial 6.7 kb fragment was digested with EcoRI.

Sarkies et al. Figure S5

A



B

```
fj1 360 ATGTGGCACTGGGACATGGGGACAGAGAGGGGTGAGGTTGGGCTGGGTGCCACAGACTAT 419
fj2 358 ATGTGGCACTGGGACATGGGGACAGAGAGGGGTGAGGTTGGGCTGGGTGCCACAGACTAT 417
fj3 360 ATGTGGCACTGGGACATGGGGACAGAGAGGGGTGAGGTTGGGCTGGGTGCCACAGACTAT 419
fj4 358 ATGTGGCACTGGGACATGGGGACAGAGAGGGGTGAGGTTGGGCTGGGTGCCACAGACTAT 417
fj5 360 ATGTGGCACTGGGACATGGGGACAGAGAGGGGTGAGGTTGGGCTGGGTGCCACAGACTAT 419
fj6 359 ATGTGGCACTGGGACATGGGGACAGAGAGGGGTGAGGTTGGGCTGGGTGCCACAGACTAT 418
```

Figure S5. Genetic stability of the *CD72* G4 DNA motifs in *fancj* cells. A. Map of the *CD72* locus indicating the position of the primers used to amplify the G4 DNA motif. B. Sequences of the G4 DNA motif from 6 CD72-loss variant *fancj* clones.

Sarkies et al. Figure S6

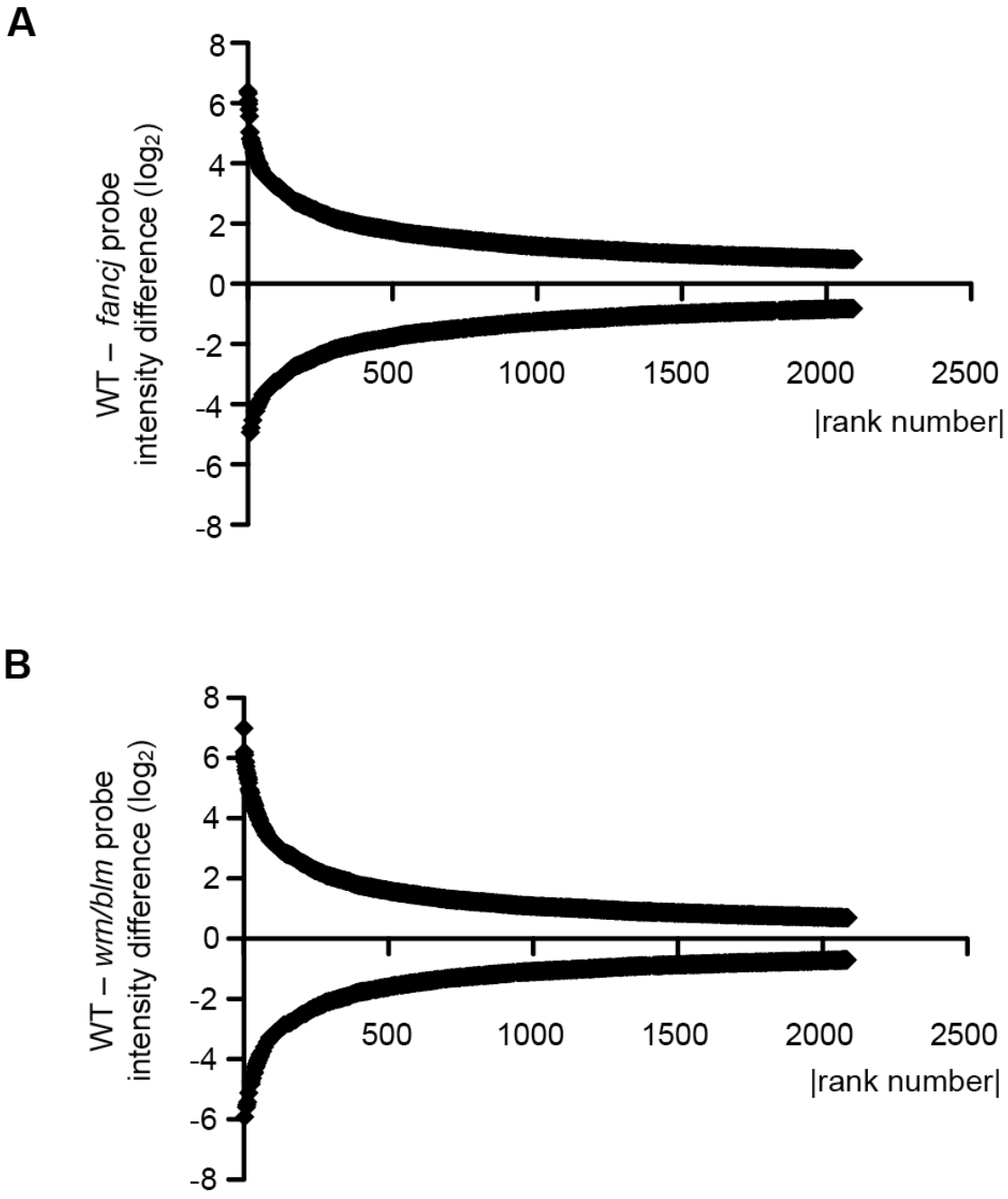


Figure S6. Gene deactivation and derepression in *fancj* and *wrn*/*blm* cells. Mod ranked plot of probe intensity differences for probes whose expression is altered by >0.25 $\log_2$ units in *rev1* compared with wild type with $p < 0.075$.

Sarkies et al. Figure S7

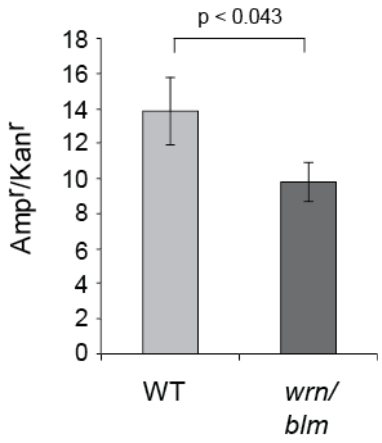


Figure S7. Efficiency G4 DNA replication in *wrn*/*blm* cells. Replication efficiency, in *wrn*/*blm* cells, of the *RHOGLOBIN* G4 DNA (14), incorporated into the leading strand template of the replicating plasmid pQ, shown as the ratio of Amp^r to Kan^r *E. coli* colonies. Amp^r plasmid contains the CD72-F oligo while Kan^r plasmid does not (14). The single mutants *wrn* and *blm* did not exhibit any decrease in the efficiency of replication of this plasmid. Attempts to perform this assay in *fancj* cells were unsuccessful due to reproducibly poor recovery of pQ.

Sarkies et al. Figure S8

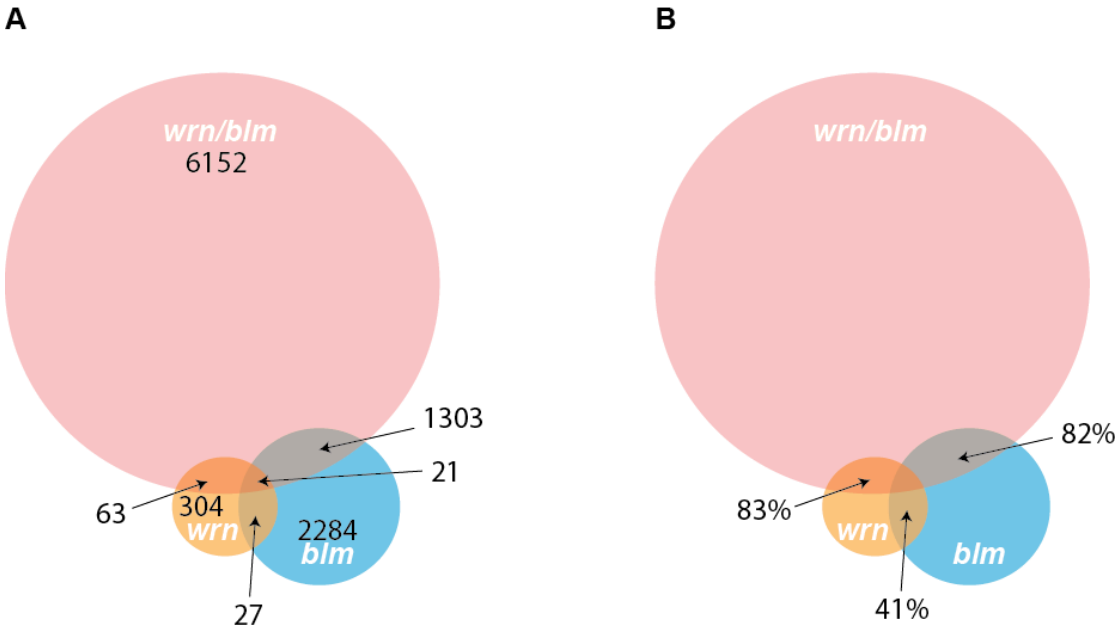


Figure S8. Redundancy between WRN and BLM in maintenance of epigenetic stability at G4 DNA. A. Venn diagram showing the number of probes statistically significantly perturbed ($p<0.05$) with a greater than 0.25 log units difference to WT DT40 in *wrn*, *blm* and *wrn/blm* mutants, and the overlap between these sets. Significance of overlap between sets (Fisher's hypergeometric distribution): *wrn* & *blm* p = NS; *wrn* & *wrn/blm* p = NS; *blm* & *wrn/blm* $p < 1 \times 10^{-15}$. B. Venn diagram as in Figure S3A showing the percentage of probes, within each set, which show the same direction of change in each pair of mutants. The expected codirectionality assuming independence between the conditions is approximately 50% for each set. Significance of co-directionality between sets (chi-squared test): *wrn* & *blm* p = NS; *wrn* & *wrn/blm* $p < 3 \times 10^{-7}$; *blm* & *wrn/blm* $p < 1 \times 10^{-17}$.

Sarkies et al. Figure S9

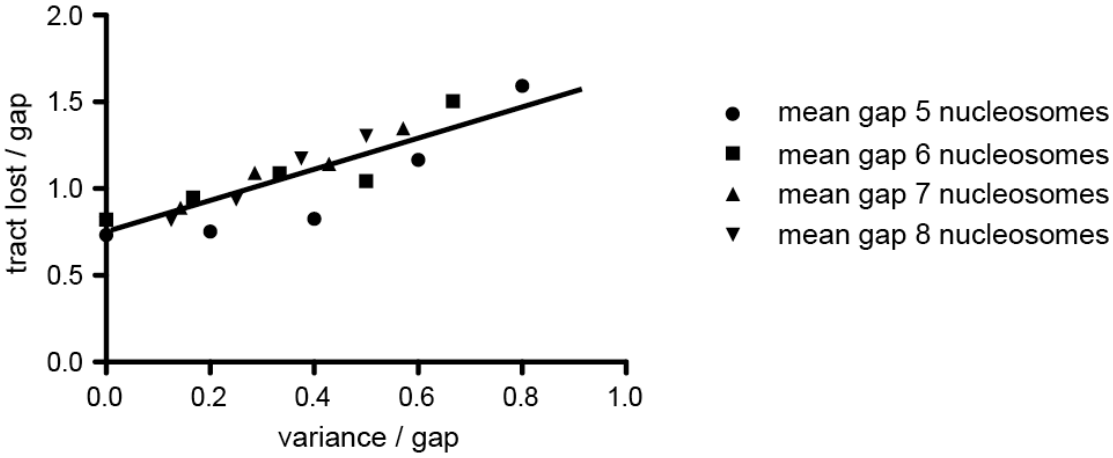


Figure S9. Effect of gap length variance on histone modification loss. Using the Zippee algorithm (14), replication fork stalling at a fixed position, with a probability of 0.5 per cell division, was simulated for 30 cell divisions, with 30 repetitions of the simulation per point. The probability of mark copying reaching two nucleosomes was 0.5, 3 nucleosomes 0.25 etc. Mean gap length was varied between 5 and 8 nucleosomes, which corresponds to a gap approximately 1 - 1.6 kbp, and variance was altered from 0 to a maximum of 0.8 of the gap length, assuming a normal distribution of gap lengths clamped at 0 nucleosomes (i.e. a gap length of less than 1 nucleosome in length). Random histone misincorporation was set at 1%. The final gap length after 30 cell divisions, as a fraction of the mean gap length, is plotted against the variance as a fraction of the mean gap length. This plot reveals a roughly linear relationship, with the gradient approximately the same for all the different gap lengths tested, which indicates that as the gap length variance increases, the final length of gap also increases. Thus, the final gap length after 30 cell divisions is not just determined by the mean gap length left whenever the replication fork stalls, but also by the maximum gap length that occurs. This suggests that the occurrence of rare events where long gaps are left leads to a longer tract of histone modification loss.

Sarkies et al. Table S1

Mutant line	Reference	Notes
rev1 (Cambridge)	(31)	
rev1 (Munich)	(47)	
rev1:hREV1	(13)	
rev1:hREV1$^{\Delta cat}$	(13)	
rev3	(56)	
polη	(57)	
rad18	(58)	
*pcna*K164R	(47)	
blm	(59)	
wrn	(32)	
blm/*wrn*	this study	*wrn* cells targeted with *blm* constructs (59)
fancj	this study	Figure S1
fancc	(60)	
rad54	(61)	
xrcc3	(62)	
xpa	(63)	

Table S1. DT40 mutants used in this study.

Sarkies et al. Table S2 - Legend

Identity of probes exhibiting dysregulated expression in *rev1*, *fancj*, *wrn*, *blm* and *wrn*/*blm* DT40. Table S2 is an Excel workbook listing all the probes whose signals are more than 0.25 $\log_2$ units increased or decreased from wild type with $p < 0.075$. There is a single sheet for each mutant and for the overlaps between mutants (gene sets A – C in Figure 6C). The probes are ordered in descending absolute intensity difference from wild type.

Sarkies et al. Table S3

Sequence	Melting temperature T_m (°C)		
	Li^+	Na^+	K^+
BU1A-1	*no transition*	30.0	61.5
BU1A-2	*no transition*	45.5	59.5
BU1A-F	20.5	50.5	65.5
CD72-1	25.5	36.0	52.0
CD72-2	25.0	40.0	59.5
CD72-F	25.0	42.0	59.5

Table S3A. Melting temperature of the studied G4 DNA oligonucleotides. Melting temperatures were calculated by following the UV absorbance at 295nm. Oligonucleotides were prepared at a final concentration of 4 μM in 10mM lithium cacodylate (pH 7.2) containing 1mM of EDTA and 100mM of MCl (where M is either Li, Na or K).

Sequence	CD characterisation			
		Li^+	Na^+	K^+
BU1A-1	maxima minima	261, 284 nm 235 nm	261, 291 nm 235 nm	264 nm 242 nm
BU1A-2	maxima minima	278 nm 235 nm	245, 292 nm 259 nm	264 nm 241 nm
BU1A-F	maxima minima	263, 281 nm 236 nm	253, 295 nm 230, 268 nm	264 nm 242 nm
CD72-1	maxima minima	260 nm 240 nm	260, 295 nm 238 nm	264, 292 nm 239 nm
CD72-2	maxima minima	259 nm 241 nm	258, 295 nm 237 nm	264, 291 nm 237 nm
CD72-F	maxima minima	260 nm 240 nm	259, 294 nm 238 nm	266, 289 nm 239 nm

Table S3B. CD characterisation of the studied G4 DNA oligonucleotides. Oligonucleotides were prepared at a final concentration of 10 μM in 10mM lithium cacodylate (pH 7.2) containing 1mM of EDTA and 100mM of MCl (where M is either Li, Na or K).

Sarkies et al. Table S4

correlation / probe sets	log_2*wrn/blm* + log_2*rev1* *vs.* log_2*fancj*	log_2*wrn/blm* *vs.* log_2*fancj*	log_2*rev1* *vs.* log_2*fancj*	n	p ($R_{wrn/blm+rev1\ vs.\ fancj} > R_{wrn/blm\ vs.\ fancj}$)	p ($R_{wrn/blm+rev1\ vs.\ fancj} > R_{rev1\ vs.\ fancj}$)
(fancj + wrn/blm) – rev1	0.74	0.85	0.35	3890	1	1×10^{-36}
(fancj + rev1) – wrn/blm	0.81	0.59	0.80	88	1.7×10^{-3}	0.43
wrn/blm + rev1	0.82	0.69	0.62	172	2.3×10^{-3}	3.6×10^{-5}

Table S4. Correlation coefficients (R) for comparison of gene expression changes in overlap sets of *rev1* and *wrn/blm* against *fancj*. The correlations tested are indicated in the headers to columns 2 - 4. The probe sets tested are indicated in the headers to rows 2 - 4. For example '*(fancj + wrn/blm) – rev1*' means the probes overlapping between the *fancj* and *wrn/blm* sets but excluding those that also overlap with *rev1*. The correlations coefficient R is indicated as are the number (n) of probes analysed. Only in the overlap between *wrn/blm* and *rev1* (Row 4) is the correlation coefficient for log_2*wrn/blm* + log_2*rev1* statistically significantly greater ($p<0.01$) than both log_2*wrn/blm* with log_2*fancj* and log_2*rev1* with log_2*fancj*.

Sarkies et al. Table S5

ChIP, qPCR and PCR primers

Name	Sequence	Reference
*RHOGLOBIN*promF	5′-TCAGTGTGCACAAGGTGTGG	(64)
*RHOGLOBIN*promR	5′-GAAGGGTGAGGGAAGTGCC	(64)
*LYSC*promF	5′-CCACATTGTATAAGAAATTTGGCAA	(65)
*LYSC*promR	5′-AAAACGCCTCTTGAGTATACAGAA	(65)
*GAS41*promF	5′-CGTGAACTGCGCGAAGAAG	(65)
*GAS41*promR	5′-CCCCCGCCACCTACCA	(65)
*CD72*promF	5′-CAAGTTCTGCTTCTGTAAGAGCC	This study
*CD72*promR	5′-AGGTCAGTGTAGAGCACTCCTTG	This study
*BU1A*promF	5′-CTCTGTAGCCAGATCGTCTTCTC	This study
*BU1A*promR	5′-GTGTCAGCTCATCTAGGCAAATC	This study
*BU1A*expnF	5′-CTGTTACTGATGGCTCTGCTACC	This study
*BU1A*expnR	5′-CTCCAGTTTCAGACATCTCTTGG	This study
*CD72*expnF	5′-CAAGGAGTGCTCTACACTGACCT	This study
*CD72*expnR	5′-GACTCTCAGCCTCATCCATACC	This study
*ACTB*expnF	5′-TGTCCACCTTCCAGCAGATGT	(65)
*ACTB*expnR	5′-AGTCCGGTTTAGAAGCATTTGC	(65)
*BU1A*SthnprobeF	5′-GCTTCTGACATCATTTCCTGAATGC	This study
*BU1A*SthnprobeR	5′-ACCTTATGTTCTCCATTGACACGC	This study
*BU1A*G4seqF	5′-GG**GAATTC**AAGGCTGACTCTCCTCTG AAGCTA	This study
*BU1A*G4seqR	5′-GC**GGATCC**GGAGCACATCACTAAGTA ACCAGAC	This study
*CD72*G4F	5′- CTTGGTATGGATGAGGCTGAGAGTC	This study
*CD72*G4R	5′- AGAAGAACCAAGTGAGGAGTATGG	This study
BU1ALF	5′- TGCAGCTAGACCAGAGTAGGTATT	This study
BU1AL4	5′- GATCTCCATAGACAGATGAGGAC	This study

G quadruplex experiments

BU1A-1	5′-GGGCTGGGTGGGTGCTGTCAAGGG
BU1A-2	5′-GGGTGGGTGCTGTCAAGGGCTGGG
BU1A-F	5′-GGGCTGGGTGGGTGCTGTCAAGGGCTGGG
CD72-1	5′-GGGACATGGGGACAGAGAGGGGTGAGGTTGGG
CD72-2	5′-GGGGACAGAGAGGGGTGAGGTTGGGCTGGG
CD72-F	5′-GGGACATGGGGACAGAGAGGGGTGAGGTTGGGCTGGG
HTelo	5′-GGGTTAGGGTTAGGGTTAGGG
C-KIT1	5′-GGGAGGGCGCTGGGAGGAGGG
Bu1a pQ G4 1	5′-AATTCGGGCTGGGTGGGTGCTGTCAAGGGCTGGG
Bu1a pQ G4 2	5′-AATTCCCCAGCCCCTTGACAGCACCCACCCAGCCC
CD72 pQ G4 1	5′-AATTCACCCAGCCCAACCTCACCCCTCTCTGTCCCCATGTCCC
CD72 pQ G4 2	5′-AATTCGGGACATGGGGACAGAGAGGGGTGAGGTTGGGCTGGGT

Construction of *fancj* DT40

FJ5F	5′-GC**GTCGAC**ACAGCTCTTATGCCATGATTATTAG
FJ5R	5′-GC**GGATCC**CATATTCAACTTCATCAGAACCC
FJ3F	5′- GC**GGATCC**AGTAGGAATCTCAGTCCCTTCATG
FJ3R	5′-TAA**GCGGCCGC**CCACAAACTTTGAAAGGCATTCC
FJPF	5′-GCAAAGTCAGTACAGATGCAAGACAACC
FJPR	5′-AGTCAGCTGGTAGAAACAGGGTATG

Supplementary Methods

UV spectroscopy

UV melting curves were collected using a Varian Cary 400 Scan UV-visible spectrophotometer by following the absorbance at 295 nm. Oligonucleotides solutions were prepared at a final concentration of 4 μM in 10 mM lithium cacodylate (pH 7.2) containing 1mM EDTA and 100 mM MCl (where M is either Li, Na or K). The samples were annealed by heating to 95 ˚C for 10 minutes and then slowly cooled to room temperature. Samples were transferred to a 1cm path-length quartz cuvette, covered with a layer of mineral oil, placed in the spectrophotometer and equilibrated at 10 ˚C for 5 minutes. Samples were then heated to 95 ˚C and cooled to 10˚C at a rate of 1 ˚C/min, with data collection every 1 ˚C during both melting and cooling. T_m values were obtained from the minimal of the first derivative of the melting curve. Thermal differential spectra were obtained by subtracting the UV spectra collected at 20 ˚C from the one collected at 80 ˚C of the oligonucleotide solutions prepared as described above.

Supplementary References

55. Arakawa, H., Lodygin, D. and Buerstedde, J.M. (2001) Mutant loxP vectors for selectable marker recycle and conditional knock-outs. *BMC Biotechnol*, **1**, 7.

56. Ambrus, A., Chen, D., Dai, J., Bialis, T., Jones, R.A. and Yang, D. (2006) Human telomeric sequence forms a hybrid-type intramolecular G-quadruplex structure with mixed parallel/antiparallel strands in potassium solution. *Nucleic Acids Res.*, **34**, 2723-2735.

57. Phan, A.T., Kuryavyi, V., Burge, S., Neidle, S. and Patel, D.J. (2007) Structure of an unprecedented G-quadruplex scaffold in the human c-kit promoter. *J. Am. Chem. Soc.*, **129**, 4386-4392.

58. Sonoda, E., Okada, T., Zhao, G.Y., Tateishi, S., Araki, K., Yamaizumi, M., Yagi, T., Verkaik, N.S., van Gent, D.C., Takata, M. *et al.* (2003) Multiple roles of Rev3, the catalytic subunit of polzeta in maintaining genome stability in vertebrates. *Embo J*, **22**, 3188-3197.

59. Kawamoto, T., Araki, K., Sonoda, E., Yamashita, Y.M., Harada, K., Kikuchi, K., Masutani, C., Hanaoka, F., Nozaki, K., Hashimoto, N. *et al*. (2005) Dual roles for DNA polymerase eta in homologous DNA recombination and translesion DNA synthesis. *Mol Cell*, **20**, 793-799.

60. Yamashita, Y.M., Okada, T., Matsusaka, T., Sonoda, E., Zhao, G.Y., Araki, K., Tateishi, S., Yamaizumi, M. and Takeda, S. (2002) RAD18 and RAD54 cooperatively contribute to maintenance of genomic stability in vertebrate cells. *Embo J*, **21**, 5558-5566.

61. Wang, W., Seki, M., Narita, Y., Sonoda, E., Takeda, S., Yamada, K., Masuko, T., Katada, T. and Enomoto, T. (2000) Possible association of BLM in decreasing DNA double strand breaks during DNA replication. *Embo J*, **19**, 3428-3435.

62. Niedzwiedz, W., Mosedale, G., Johnson, M., Ong, C.Y., Pace, P. and Patel, K.J. (2004) The Fanconi anaemia gene FANCC promotes homologous recombination and error-prone DNA repair. *Mol Cell*, **15**, 607-620.

63. Bezzubova, O., Silbergleit, A., Yamaguchi-Iwai, Y., Takeda, S. and Buerstedde, J.M. (1997) Reduced X-ray resistance and homologous recombination frequencies in a RAD54-/- mutant of the chicken DT40 cell line. *Cell*, **89**, 185-193.

64. Takata, M., Sasaki, M.S., Tachiiri, S., Fukushima, T., Sonoda, E., Schild, D., Thompson, L.H. and Takeda, S. (2001) Chromosome instability and defective recombinational repair in knockout mutants of the five Rad51 paralogs. *Mol Cell Biol*, **21**, 2858-2866.

65. Okada, T., Sonoda, E., Yamashita, Y.M., Koyoshi, S., Tateishi, S., Yamaizumi, M., Takata, M., Ogawa, O. and Takeda, S. (2002) Involvement of vertebrate polkappa in Rad18-independent postreplication repair of UV damage. *J Biol Chem*, **277**, 48690-48695.

66. Wang, J., Liu, H., Lin, C.M., Aladjem, M.I. and Epner, E.M. (2005) Targeted deletion of the chicken beta-globin regulatory elements reveals a cooperative gene silencing activity. *J Biol Chem*, **280**, 23340-23348.

67. Myers, F.A., Lefevre, P., Mantouvalou, E., Bruce, K., Lacroix, C., Bonifer, C., Thorne, A.W. and Crane-Robinson, C. (2006) Developmental activation of the

lysozyme gene in chicken macrophage cells is linked to core histone acetylation at its enhancer elements. *Nucleic Acids Res*, **34**, 4025-4035.