Supplementary Data
Supplementary figures 1-6 and figure legends, and Supplementary tables 1-2

MutSβ abundance and Msh3 ATP hydrolysis activity are important drivers of CTG•CAG repeat expansions

Norma Keogh[1], Kara Chan[2], Guo-Min Li[2,3] and Robert S. Lahue[1,4,*]

[1]Centre for Chromosome Biology and [4]NCBES Galway Neuroscience Centre

National University of Ireland Galway

Newcastle Road

Galway, Ireland

H91TK33

[2]Department of Toxicology and Cancer Biology, University of Kentucky College of Medicine, Lexington, KY 40536

[3]Department of Biochemistry and Molecular Biology, Norris Comprehensive Cancer Center, University of Southern California Keck School of Medicine, Los Angeles, CA 90033

*corresponding author

Legends to supplementary figures

Supplementary Figure 1. Schematic showing overview of shuttle vector assay and details of shuttle vectors used for measuring expansions and contractions. **(A)** Workflow of shuttle vector assay. SVG-A cells are seeded at $4.0x10^5$ cells and transfected with the appropriate shuttle vector. The plasmid is harvested and purified after 48 hours, then transformed into *S. cerevisiae* as a biosensor. Yeast cells are plated onto SC-histidine plates for total transformants for both assays and onto SC-his-uracil plates for genetic selection for contraction assays. Replica plating onto canavanine plates (100 μg/ml) is required for genetic selection for expansion assay only. Colony counts and PCR confirmation are required to

calculate expansion/contraction frequency. **(B)** Schematic showing pBL302 reporter plasmid system used to measure CTG expansions. Expansion of ≥ 4 CTG repeats causes out of frame translation of the *CAN1* gene and confers canavanine resistance. In these studies, expansion of + 12 repeats was the largest expansion observed. **(C)** Reporter plasmid pBL247 is used to measure CTG contractions for the contraction assay. Contraction of CTG repeats from 33 repeats to ≤ 28 repeats causes an Ura^+ phenotype, which allows genetic selection of contractions. A maximum contraction size of -30 CTG repeats was observed in these studies.

Supplementary Figure 2. Immunoblot analysis of *Msh3* variant cell lines with Msh3 C-terminal antibody, Abcam Ab154521, raised against amino acid positions 916-1137 of Msh3. **(A)** Immunoblot showing Msh3 protein expression in $Msh3^{+/+}$ and Msh3 variant cell lines. No detectable Msh3 expression in $Msh3^{-/-}$ cell line, confirming knockout of *Msh3*. Second band, marked with *, is non-specific as noted by the manufacturer. **(B)** Schematic showing CRISPR/Cas9 target region and immunogen regions of both Msh3 antibodies used in screening cell lines.

Supplementary Figure 3. Growth curve analysis of $Msh3^{-/-}$, $Msh3^{+/+}$and $Msh3^{1.7X}$ cell lines, showing no detectable difference in cell growth. Error bars denote ±SEM, n=3.

Supplementary Figure 4. Estimation of MutSβ abundance as fraction of overall MutS complexes in SVG-A cells. **(A)** Overview of experimental approach and calculation of Msh3/Msh2 capture efficiencies and MutSβ abundance. FT stands for Flow-through. **(B)** Representative immunoblot of co-immunoprecipitated Msh3 and Msh2 showing IP input and IP flow-through from wild type SVG-A cells. In this example there was 11% of total Msh3 in flow-through, correspondingly the capture efficiency was calculated at 84%. For Msh2, 78%

was measured in flow-through, indicating that 22% of Msh2 protein was captured by the anti-Msh3 antibody. MutSβ abundance is calculated using the formula (22/84)*100% to give a value of 27%. (**C)** Quantitative data of Msh3 and Msh2 capture efficiencies from co-IP data, average values shown ±SEM, *n*=4.

Supplementary Figure 5. Characterisation of $Msh3^{E976A}$ Walker B mutant cell line, in comparison to $Msh3^{+/+}$ cell line. **(A)** Growth curve analysis of $Msh3^{E976A}$ cells in comparison to $Msh3^{+/+}$ showing no detectable defect in cell growth. *n*=3. **(B)** Quantitative analysis of immunoprecipitated Msh3 and Msh2 protein abundance normalised to protein input and relative to $Msh3^{+/+}$cells. A partial defect in MutSβ complex formation is observed in $Msh3^{E976A}$ cell lines. *n*=3, **P*=0.023 for Msh3 and $P=1.5\text{x}10^{-4}$ for Msh2. For both panels, error bars denote ±SEM.

Supplementary Figure 6. Characterisation of Msh3 polymorphic cell lines in comparison to $Msh3^{+/+}$ wild type cells. (**A**) Quantitative analysis of Msh protein expression normalised to actin and relative to $Msh3^{+/+}$ cell line. $Msh3^{T363I}$ add-back cell line has an average Msh3 expression level 75% that of wild type expression, while the value for $Msh3^{T363I}$ add-back cell was 71% of control. *n*=7 for Msh3 wild-type or *n*=4 for polymorphic Msh3 proteins; *n*=3-4 for Msh2 and Msh6 protein levels in all cell lines. **P* = 0.035 compared to Msh2 level in $Msh3^{+/+}$ cells; and ***P*=0.0085, both compared to Msh6 level in $Msh3^{+/+}$ cells. **(B)** Quantitative analysis of immunoprecipitated Msh3 and Msh2 protein abundance normalised to protein input and relative to $Msh3^{+/+}$cells. A decrease in immunoprecipitated Msh2 protein levels in observed in $Msh3^{T363I}$ cells; *n*=3, **P*=0.004; and in immunoprecipitated Msh2 protein levels in observed in $Msh3^{T1045A}$ cells; *n*=3, **P*=0.005. **(C)** Growth curve analysis of

Msh3^{T363I} and *Msh3*^{T1045A} cell lines in comparison to *Msh3*$^{+/+}$ cells, showing no detectable difference in cell growth, *n*=3. For all panels, error bars denote ±SEM.

Supplementary Table 1. Mutagenic primers were designed using NEBasechanger program (New England Biolabs). Primers were used in pairs to perform site-directed mutagenesis using the Q5 site-directed mutagenesis kit (New England Biolabs) to create Msh3 ATPase and polymorphic variant cell lines. Mutagenic nucleotides are shown in bold and underline.

Supplementary Table 2. Values of *n* and *P* for all figures.

Figure S1

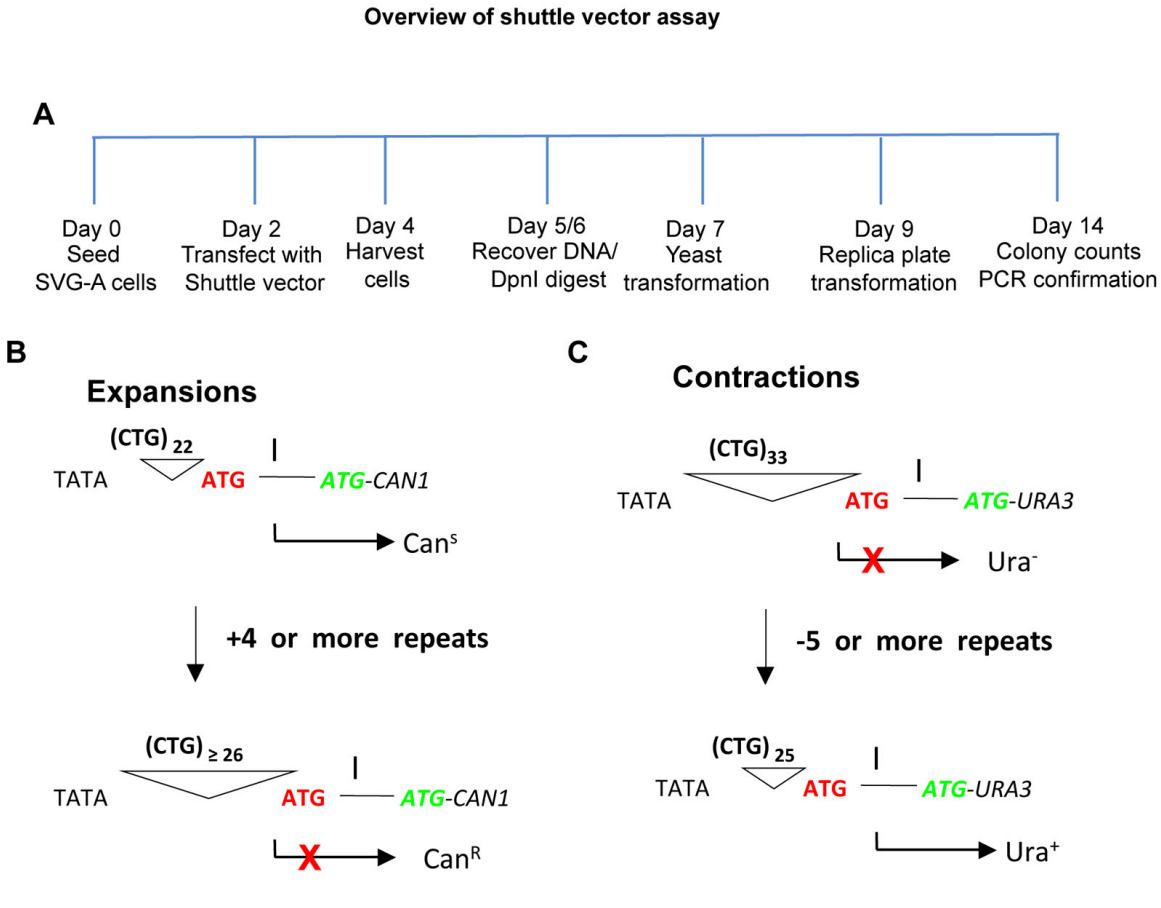
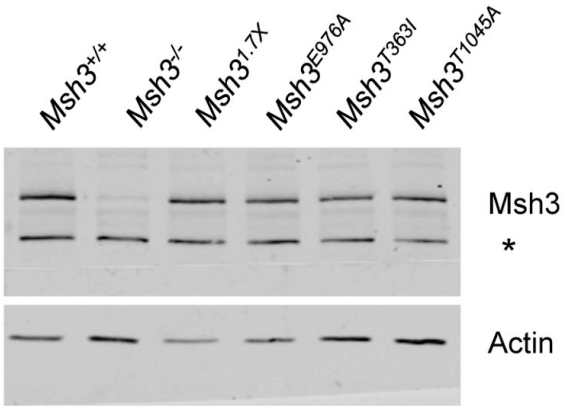

Overview of shuttle vector assay
A
Day 0
Seed
SVG-A cells
Day 2
Transfect with
Shuttle vector
Day 4
Harvest
cells
Day 5/6
Recover DNA/
DpnI digest
Day 7
Yeast
transformation
Day 9
Replica plate
transformation
Day 14
Colony counts
PCR confirmation
B
Expansions
(CTG) 22
TATA
ATG
ATG-CAN1
Can^S
+4 or more repeats
(CTG) ≥ 26
TATA
ATG
ATG-CAN1
Can^R
C
Contractions
(CTG)33
TATA
ATG
ATG-URA3
Ura-
-5 or more repeats
(CTG) 25
TATA
ATG
ATG-URA3
Ura+

Figure S2

A



B

Deletion in *Msh3*$^{-/-}$ Δaa 116-118

Antibody	Immunogen
BD Msh3 611390	Human Msh3 aa. 136-349
Abcam Msh3 Ab154521	Human Msh3 aa. 916-1137

Figure S3

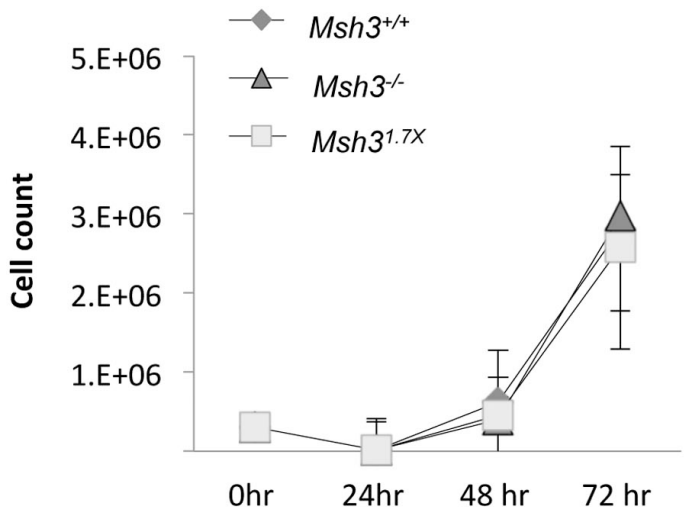


Figure S4

Estimation of MutSβ abundance

A

- Measure Msh3 capture efficiency

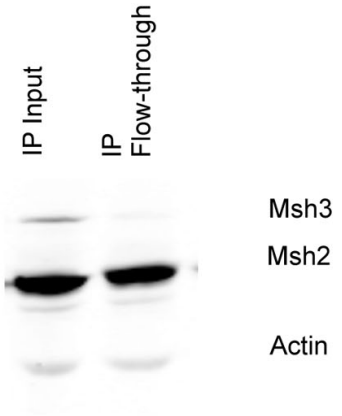
$$1 - \left(\frac{Msh3_{FT}/Actin_{FT}}{Msh3_{Input}/Actin_{Input}} \right) * 100\%$$

- Measure Msh2 capture efficiency

$$1 - \left(\frac{Msh2_{FT}/Actin_{FT}}{Msh2_{Input}/Actin_{Input}} \right) * 100\%$$

- MutSβ abundance = Msh2 capture efficiency/ Msh3 capture efficiency

B



C

Protein	Capture efficiency	SEM	n
Msh3	84%	11%	4
Msh2	22%	6%	4
MutSβ abundance	26%	6%	4

Figure S5

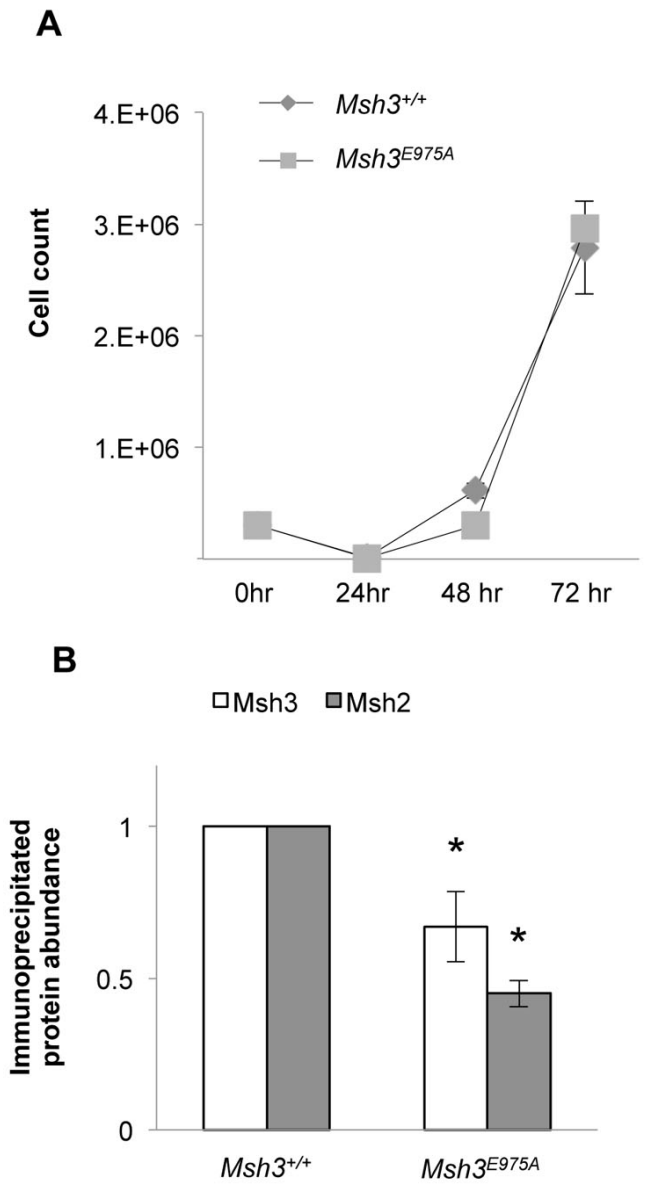
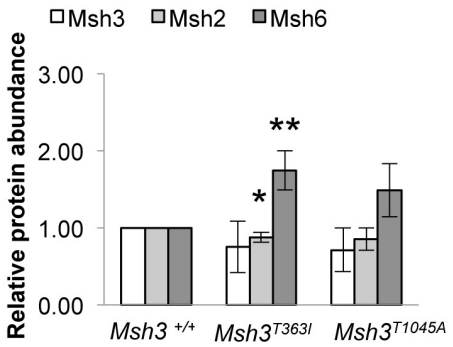
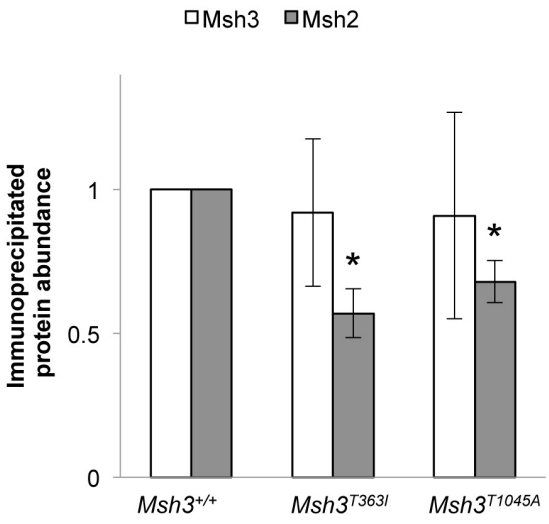

A
Msh3+/+
Msh3E975A
Cell count
4.E+06
3.E+06
2.E+06
1.E+06
0hr
24hr
48 hr
72 hr
B
Msh3
Msh2
Immunoprecipitated protein abundance
1
0.5
0
*
*
Msh3+/+
Msh3E975A

Figure S6

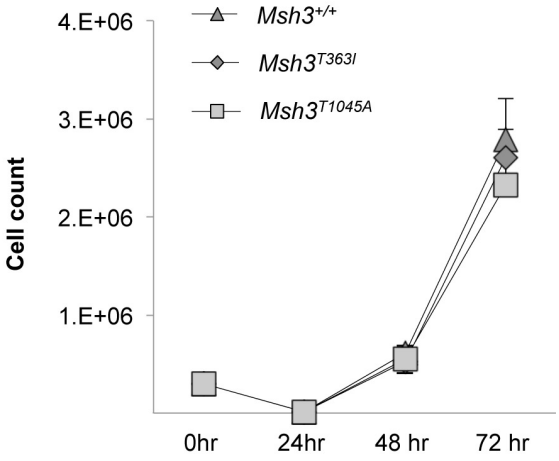
A



B



C



Supplementary Table 1

Msh3 variant	Forward primer	Reverse primer
E975A	CTAGGAAGAGGGACGAGCACTCATG	**GGC**ATCCAAGATAACCAAGGACTGTGATG
T363I	ATGACTGAT**ATT**TCTACCAGCTATC	TATCTCATCAACATTTACAGC
T1045A	GAACAAGTCCCTGATTTTGTCACCTTC	TGC**GGC**GCCTGGATCCAGTTTGCTTTC

Supplementary Table 2
n and P values

Figure	Entry	Comparator	n	P value
1C	Msh3 abundance in Msh3-/- cells	Msh3 abundance in Msh3+/+ cells	9	2.23E-20
1C	Msh2 abundance in Msh3-/- cells	Msh2 abundance in Msh3+/+ cells	6	0.135
1C	Msh6 abundance in Msh3-/- cells	Msh6 abundance in Msh3+/+ cells	6	0.058
1E	Expansion frequency in Msh3-/- cells	Expansion frequency in Msh3+/+ cells	3 - 4	0.0019
1F	Contraction frequency in Msh3-/- cells	Contraction frequency in Msh3+/+ cells	3	0.3
2B	Msh3 abundance in Msh3 1.7x cells	Msh3 abundance in Msh3+/+ cells	9	1.87E-06
2B	Msh2 abundance in Msh3 1.7x cells	Msh2 abundance in Msh3+/+ cells	5 - 7	0.082
2B	Msh6 abundance in Msh3 1.7x cells	Msh6 abundance in Msh3+/+ cells	5 - 7	0.032
2C	Expansion frequency in Msh3 1.7x cells	Expansion frequency in Msh3+/+ cells	3	0.163
2D	Contraction frequency in Msh3 1.7x cells	Contraction frequency in Msh3+/+ cells	3	0.209
2D	Contraction frequency in Msh3 2.9x cells	Contraction frequency in Msh3+/+ cells	3	0.290
2F	IP Msh3 abundance from Msh3-/- cells	IP Msh3 abundance from Msh3+/+ cells	4	5.44E-13
2F	IP Msh2 abundance from Msh3-/- cells	IP Msh2 abundance from Msh3+/+ cells	4	4.30E-11
2F	IP Msh3 abundance from Msh3 2.9x cells	IP Msh3 abundance from Msh3+/+ cells	4	1.66E-04
2F	IP Msh2 abundance from Msh3 2.9x cells	IP Msh2 abundance from Msh3+/+ cells	4	4.96E-03
3D	Expansion frequency in Msh3 E975A cells	Expansion frequency in Msh3+/+ cells	3	0.037
3E	Contraction frequency in Msh3 E975A cells	Contraction frequency in Msh3+/+ cells	4 - 5	0.065
3G	ATPase activity MutSβ-Msh3E976A minus DNA	ATPase activity wild type MutSβ minus DNA	3	0.028

Supplementary Table 2 continued

3G	ATPase activity MutSβ-Msh3E976A plus DNA	ATPase activity wild type MutSβ plus DNA	3	0.001

Supplementary Table 2 continued

4D	Expansion frequency in Msh3 T363I cells	Expansion frequency in Msh3+/+ cells	3	0.300
4D	Expansion frequency in Msh3 T1045A cells	Expansion frequency in Msh3+/+ cells	3	0.289
Supp 5B	IP Msh3 abundance from Msh3 E975A cells	IP Msh3 abundance from Msh3 +/+ cells	3	0.023
Supp 5B	IP Msh2 abundance from Msh3 E975A cells	IP Msh2 abundance from Msh3+/+ cells	3	1.50E-04
Supp 6A	Msh3 T363I abundance	Msh3 wild type abundance	4 - 7	0.17
Supp 6A	Msh3 T1045A abundance	Msh3 wild type abundance	4 - 7	0.10
Supp 6A	Msh2 abundance from Msh3 T363I cells	Msh2 abundance from Msh3 +/+ cells	3 - 4	0.035
Supp 6A	Msh2 abundance from Msh3 T1045A cells	Msh2 abundance from Msh3 +/+ cells	3 - 4	0.13
Supp 6A	Msh6 abundance from Msh3 T363I cells	Msh6 abundance from Msh3 +/+ cells	3 - 4	0.0085
Supp 6A	Msh6 abundance from Msh3 T1045A cells	Msh6 abundance from Msh3 +/+ cells	3 - 4	0.075
Supp 6B	IP Msh3 abundance from Msh3 T363I cells	IP Msh3 abundance from Msh3 +/+ cells	3	0.38
Supp 6B	IP Msh3 abundance from Msh3 T1045A cells	IP Msh3 abundance from Msh3 +/+ cells	3	0.41
Supp 6B	IP Msh2 abundance from Msh3 T363I cells	IP Msh2 abundance from Msh3 +/+ cells	3	0.004
Supp 6B	IP Msh2 abundance from Msh3 T1045A cells	IP Msh2 abundance from Msh3 +/+ cells	3	0.005