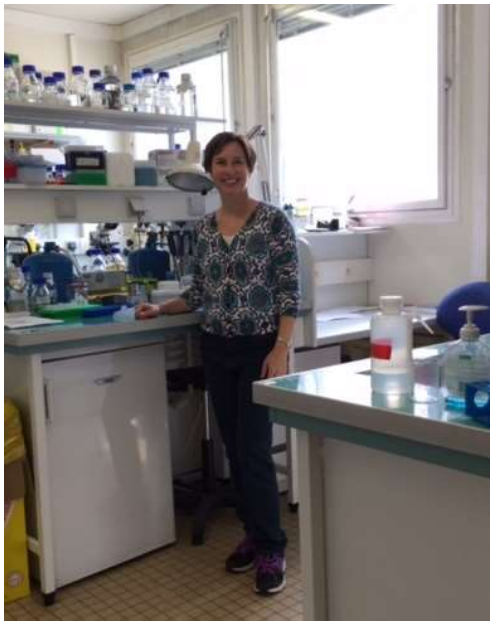
# Restarted replication forks drive CAG repeat instability

Michaela A. Gold[1], Jenna M. Whalen[1], Karine Freon[2], Zixin Hong[1], Ismail Iraqui[2], Sarah A. E. Lambert[2] and Catherine H. Freudenreich[1]

[1]Department of Biology, Tufts University, Medford, MA USA
[2]Institut Curie, CNRS, Orsay, France



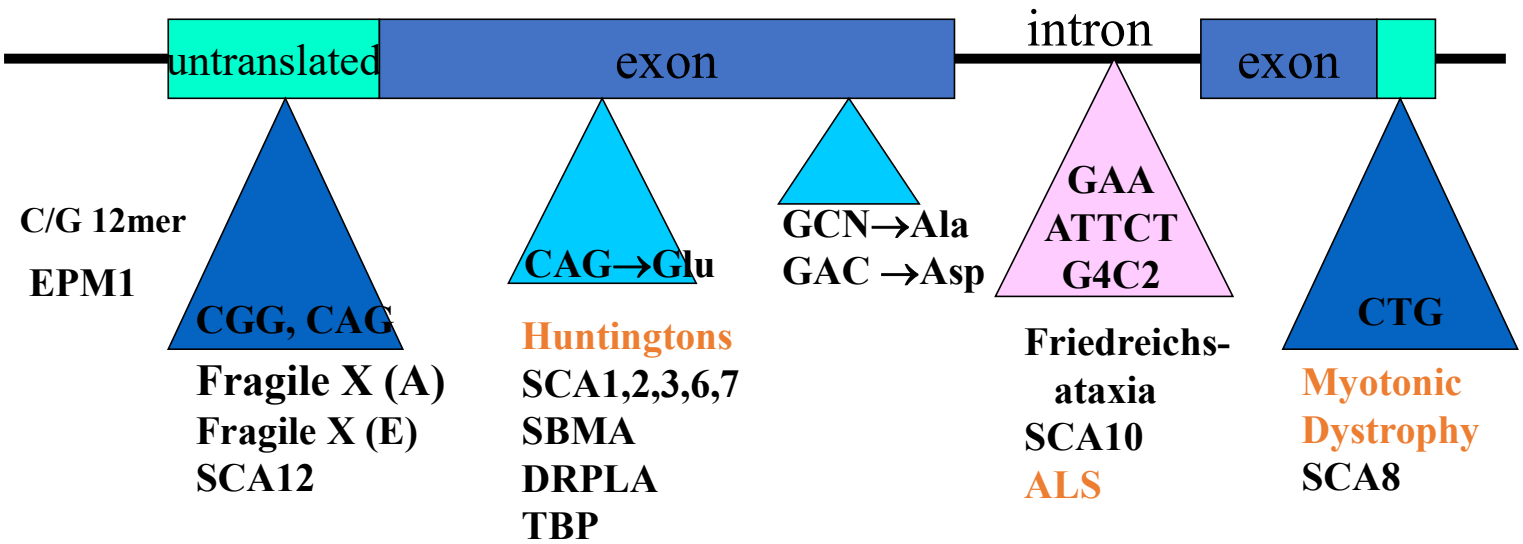
Catherine on sabbatical in Sarah Lambert's lab

Michaela Gold, Catherine Freudenreich & Jenna Whalen (Tufts group) visiting Notre Dame Cathedral

Karine Freon & Sarah Lambert

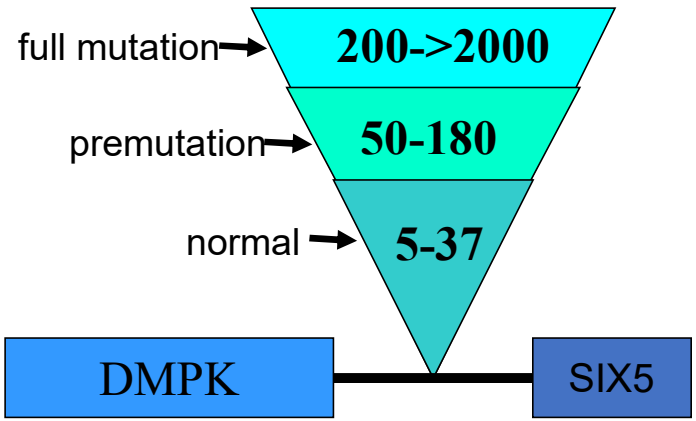
Kevin (Zixin) Hong

# Repeat Expansion Diseases

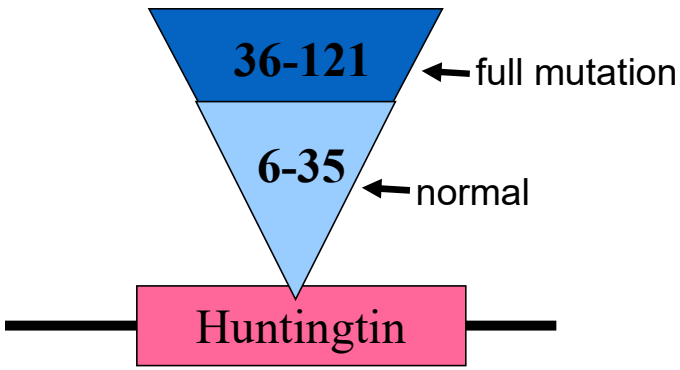


# Trinucleotide Repeat Expansions are Dynamic Mutations

## Myotonic Dystrophy (CTG)




## Huntington's Disease (CAG)




Expansions occur Intergenerationally and Somatically

6-34 REPEATS (STABLE LENGTH) —1→ 35-38 REPEATS (THRESHOLD) —2→ EXPANSIONS → 40-150 REPEATS (HD), 40-2000 REPEATS (DM)

# When do Expansions Occur?

- In the Germline

  - Paternal (HD, SCA) and Maternal expansion biases (FRAX, DM1) exist

  - Paternal age expansion bias- in dividing premeiotic and postmeiotic spermatogonia.

  - Paternal contractions happen in proliferating germ stem cells that lack methylation (FRAXA, FRDA, SCA8)

  - Expansions in maternal germline in Fragile X and myotonic dystrophy

  - Germline expansions are important in disease inheritance (intergenerational)

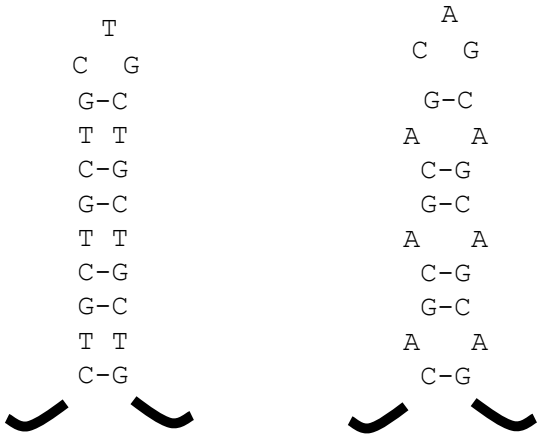
- In Somatic tissue

  - Early embryogenesis – cells dividing quickly

  - Mature somatic tissues – brain

  - Somatic expansions have an important contribution to disease progression

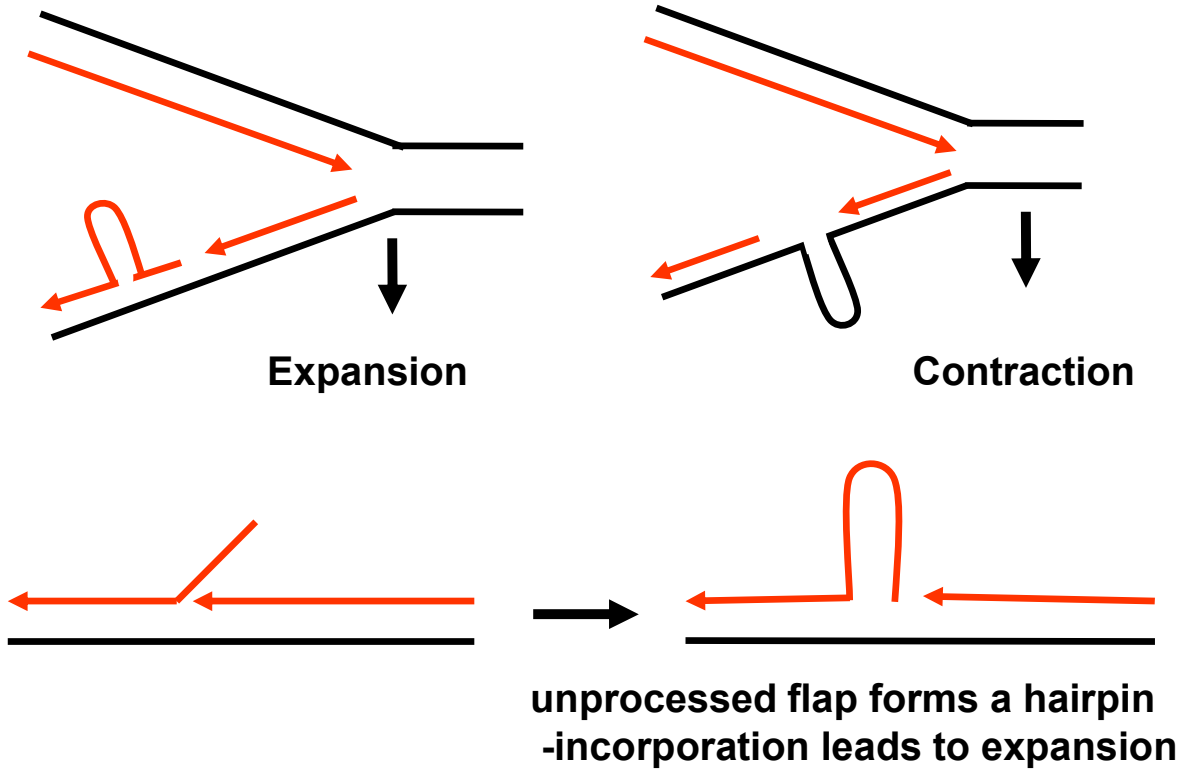
# How do Expansions Occur?
## -during either replication or repair (requires DNA synthesis)




Tm CTG > Tm CAG

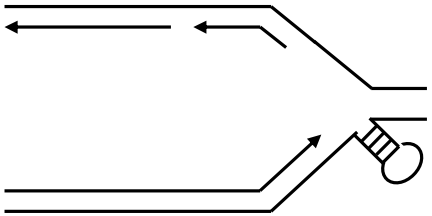
CTG and CAG repeats form hairpin structures that interfere with DNA replication and DNA repair

# Tract-length changes can occur during Replication

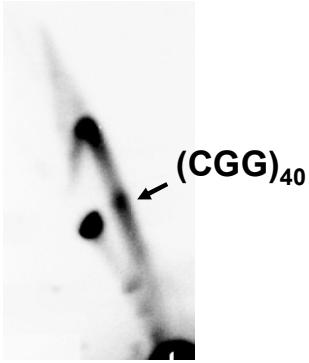



- Direction of replication determines the expansion-contraction bias (bacteria, yeast, human cells)
- Disease loci tend to be in the expansion-prone orientation
  - CAG on the lagging strand template at the DM1 locus (Cleary..Pearson, 2010)
- Mutation or inhibition of replication proteins increases expansion and contraction frequency

# Structure-forming Trinucleotide Repeats Interfere with DNA Replication

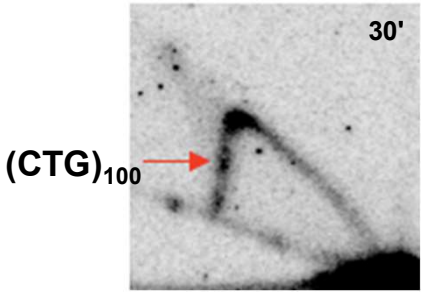


On a yeast replicating plasmid




Anand…Freudenreich, NAR, 2012

On a yeast chromosome




Nguyen, Viterbo..Richard, Freudenreich, NAR 2017

Repeat names indicate the sequence on the lagging strand template

# Question: What stages of replication are most prone to repeat instability?

- Slippage or template switch during normal replication progression through the repeat?
- Slippage or misalignment at a replication stall caused by the repeat?
- During subsequent replication of the repeat by a restarted fork that has altered properties?
- After breakage at the repeat during replication: expansion or contraction occurring during BIR?

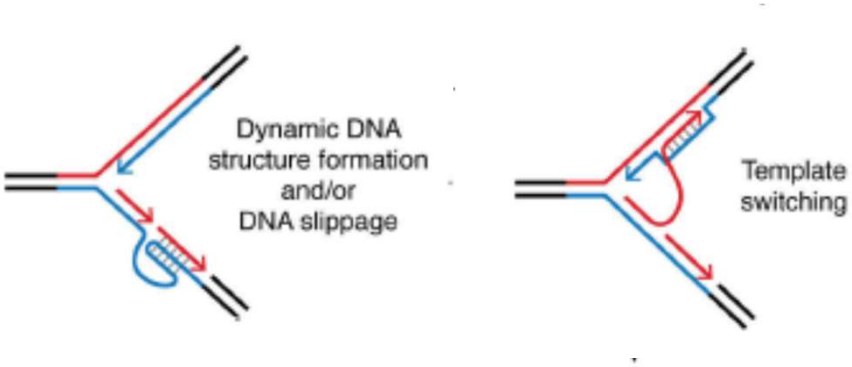



Figure from Khristich & Mirkin, 2020

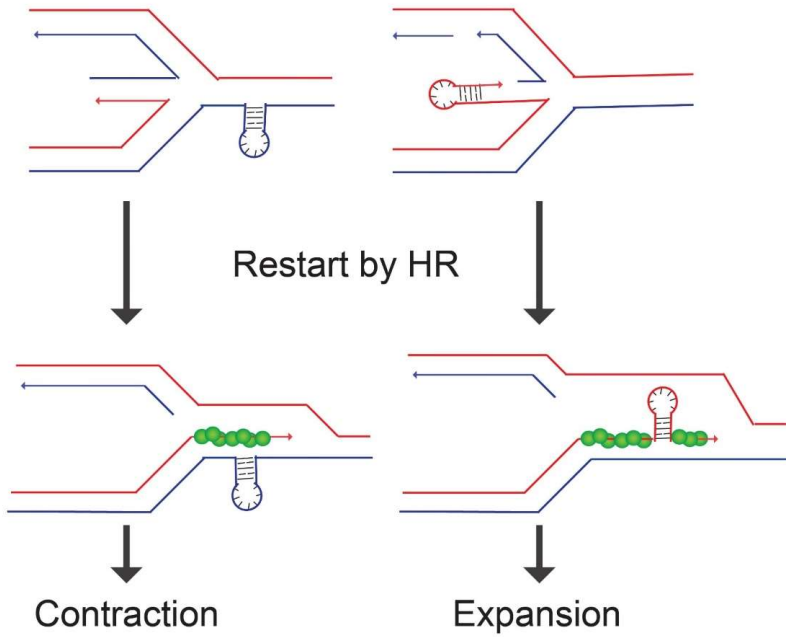



Figure from Polleys, House, Freudenreich 2017

# Question: What stages of replication are most prone to repeat instability?

- Slippage or template switch during normal replication progression through the repeat?
- Slippage or misalignment at a replication stall caused by the repeat?
- During subsequent replication of the repeat by a restarted fork that has altered properties?
- After breakage at the repeat during replication: expansion or contraction occurring during BIR?

**The Idea: separate the repeat-mediated stall from the restarted fork**

-Induce restart with a protein-mediated stall STOP

-Restarted fork travels through a CAG repeat CAG repeat

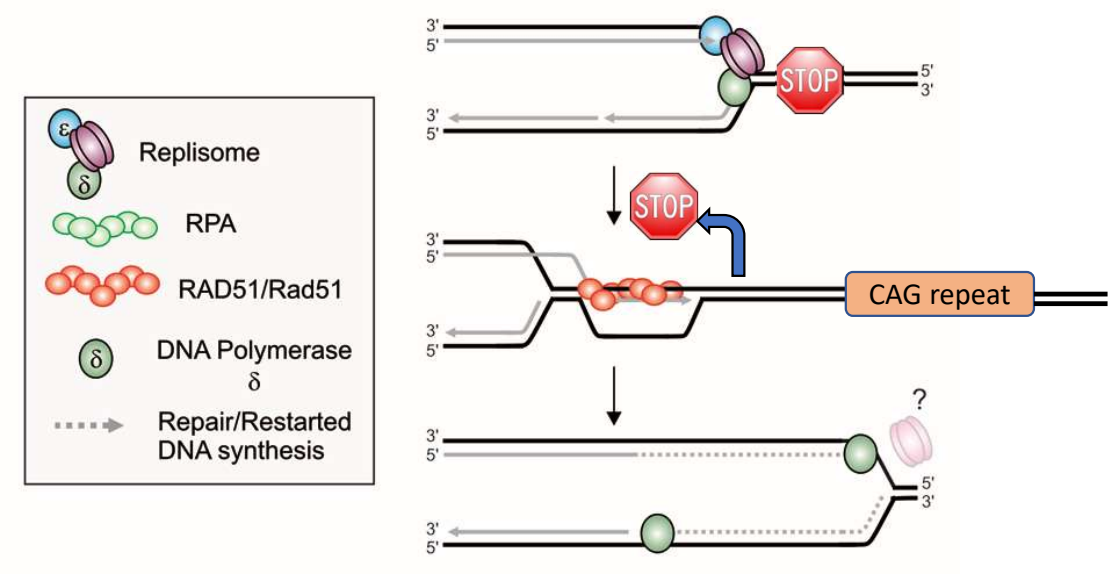
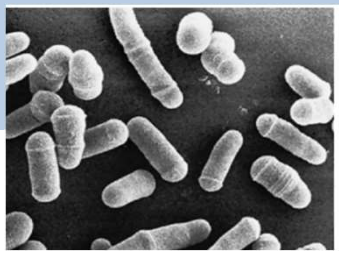
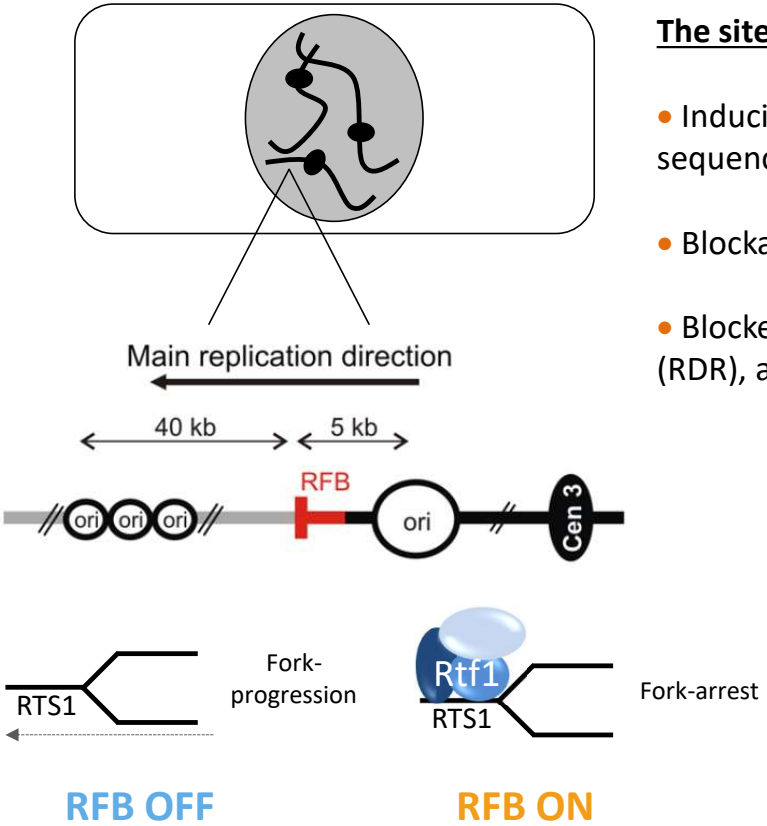



Figure from Sarah Lambert

# Creating a single and local replication stress site with the power of yeast genetics



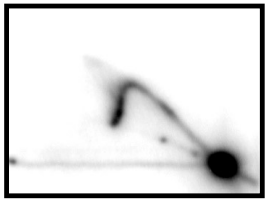
*Schizosaccharomyces pombe*
fission yeast



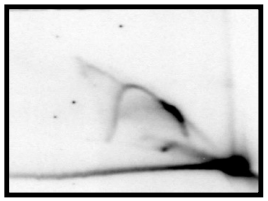

**The site specific *RTS1*-RFB**

- Inducible *polar* fork arrest mediated by the binding of Rtf1 to *RTS1* sequence
- Blockade at the same, unique *locus* at Chromosome III in > 90% of cells
- Blocked forks are restarted by Recombination-Dependent Replication (RDR), a form of non-canonical DNA synthesis

RFB OFF



RFB ON

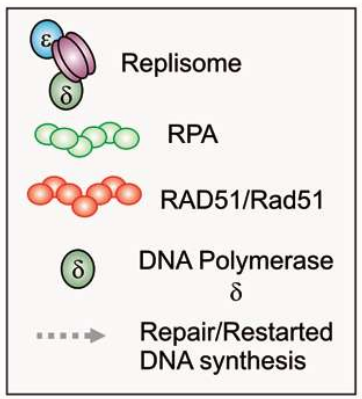


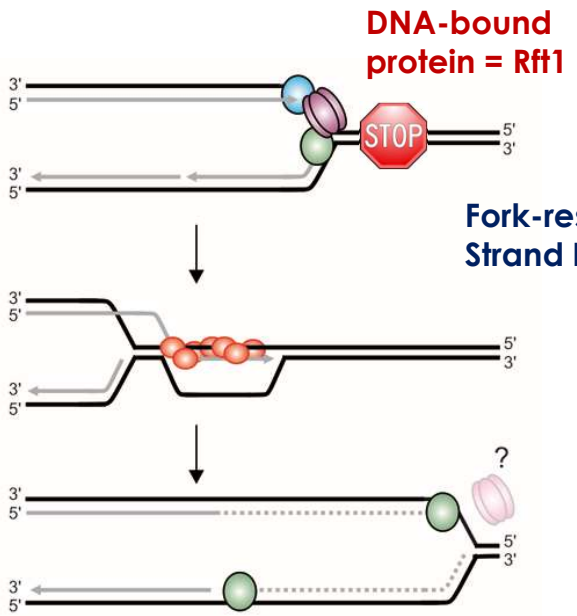
*Lambert Cell 2005; Mizuno Gene & Dev 2009; Lambert Mol Cell 2010; Iraqui PLoS Genetics 2012; Mizuno Nature 2013; Tsang J. Cell Science 2014; Nguyen Elife 2015, Miyabe NSMB 2015*

# Recombination-Dependent Replication (RDR) as a fork-restart pathway

Lambert, Murray, Whitby and Carr labs

*Lambert Cell 2005*
*Mizuno Gene & Dev 2009*
*Lambert Mol Cell 2010*
*Iraqui PLoS Genetics 2012*
*Mizuno Nature 2013*
*Tsang J. Cell Science 2014*
*Nguyen Elife 2015*
*Miyabe NSMB 2015*
*Jalan Elife 2019*
*Ait Saada Life Sci Alliance 2019*
*Hardy PLoS Genetics 2019*
*Naiman Nat Com 2021*




**Fork-resection** (MRN-Ctp1-Ku: initial; Exo1: Long-range; Fft3$^{Fun30/SMARCAD1}$)
**Strand Invasion** (Dependent on **Rad52** & Rad51 strand exchange activity)

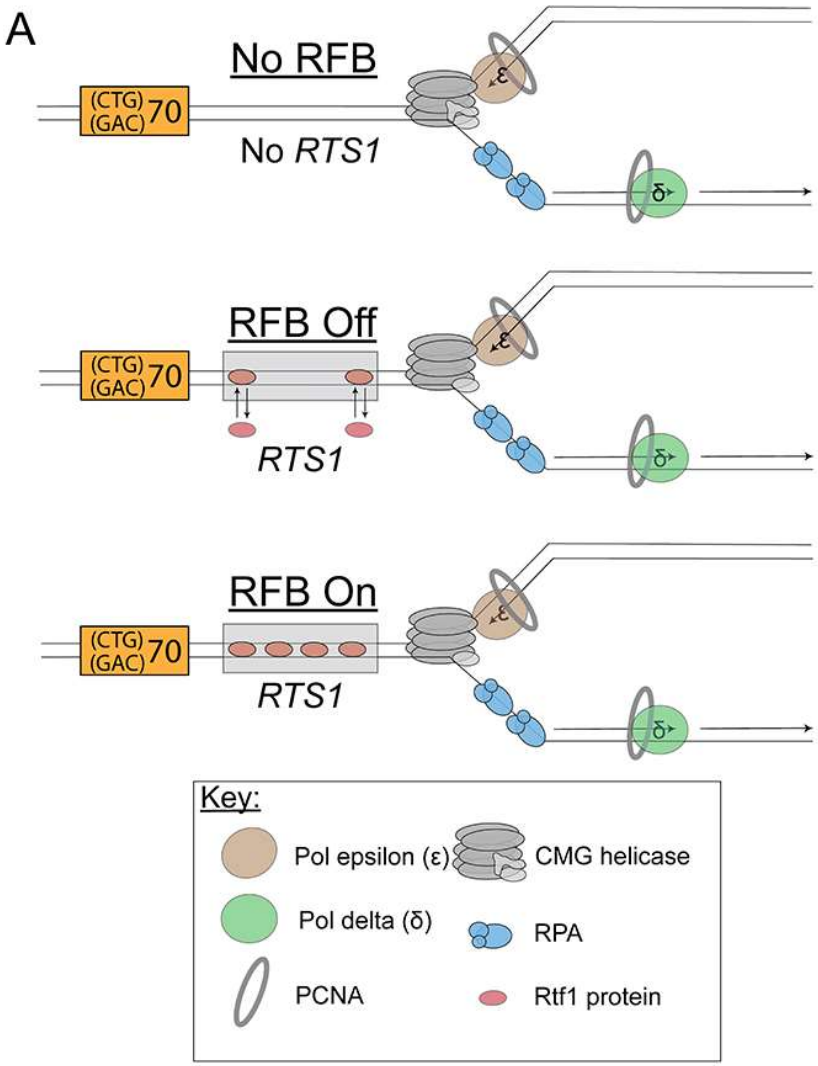
**Restarted fork (in ~ 20min):**

- Semi-conservative DNA synthesis (≠ BIR, = Canonical fork)
- Pol δ/ Pol δ DNA synthesis (≠ Canonical fork)
- Error-prone, liable to slippage mutations and template switch
- Prone to GCRs (Gross Chromosomal Rearrangements)

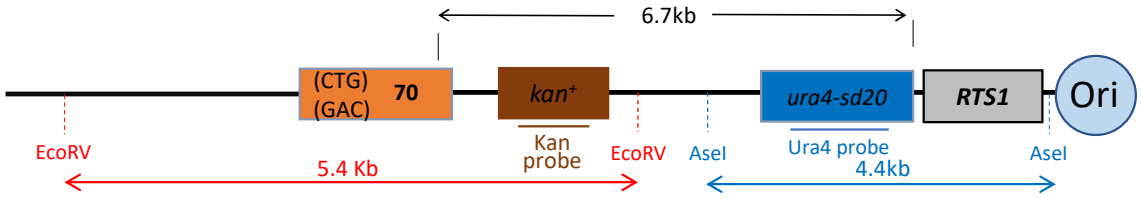
# The System:

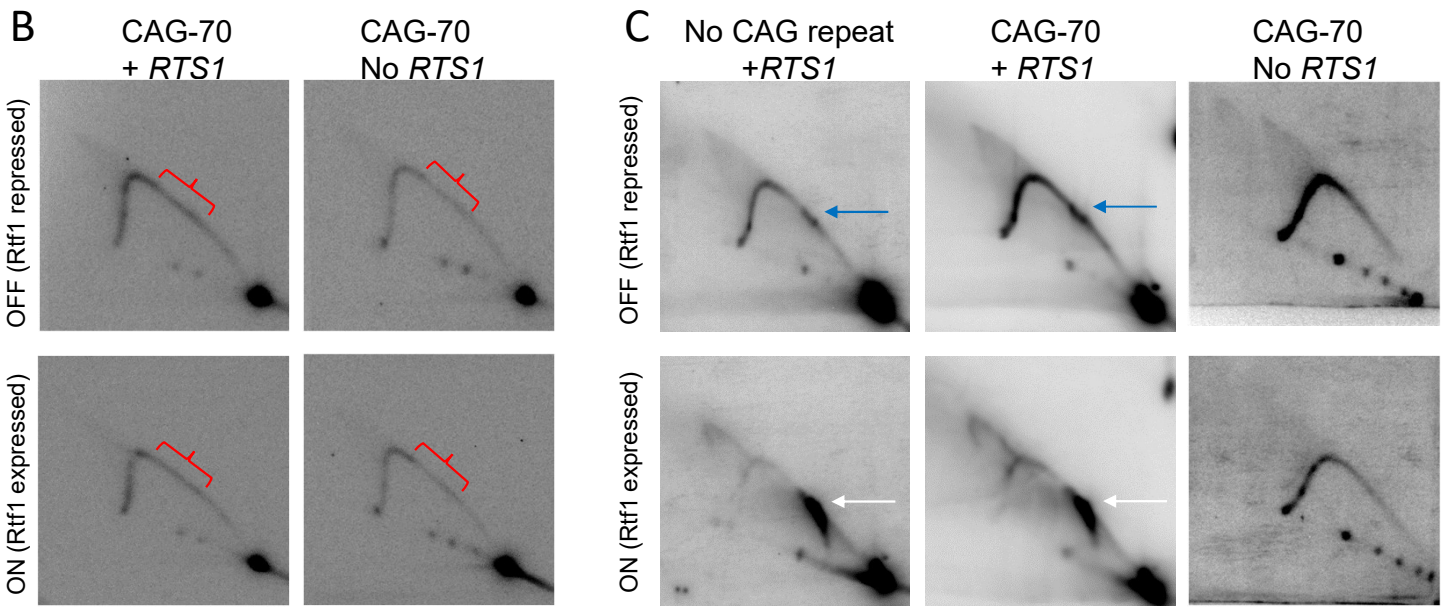
**Put a (CAG)70 repeat tract after an inducible replication fork barrier (RFB)**

**Does replication through the repeat by the restarted fork cause repeat expansions or contractions?**



# Analysis of Replication Intermediates in the Region

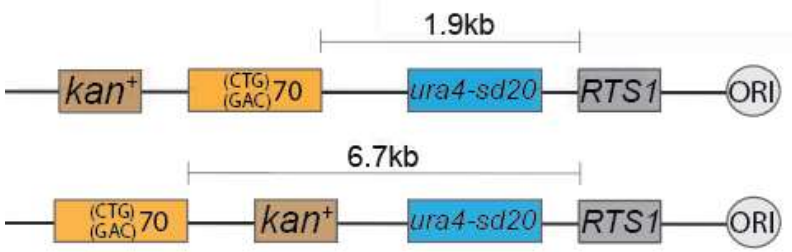


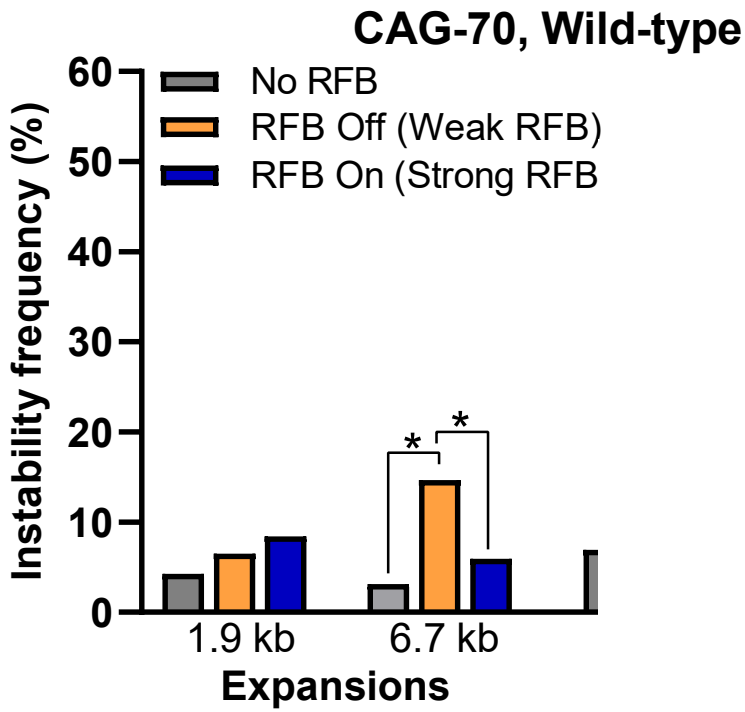




CAG-70-containing fragment
(EcoRV digest, Kan probe)

*RTS1*-containing fragment
(Ase1 digest, Ura4 probe)

# Both CAG Expansions and Contractions increase after induction of a nearby Replication Fork Barrier (RFB)



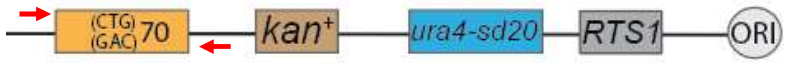



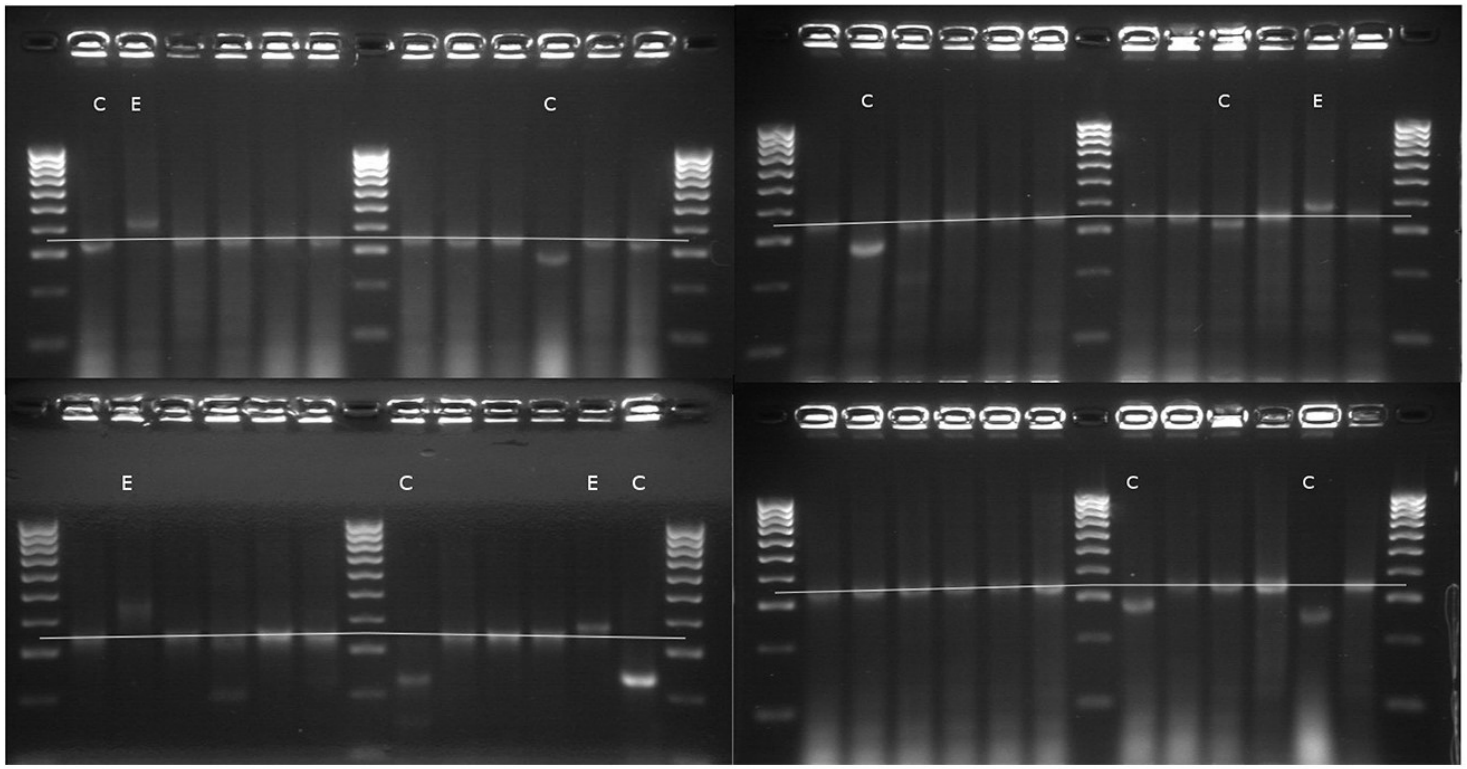

**Conclusions:**

1. The RFB increases CAG expansions

   Surprisingly, the effect is more dramatic for a WEAK RFB and increases further from the RFB

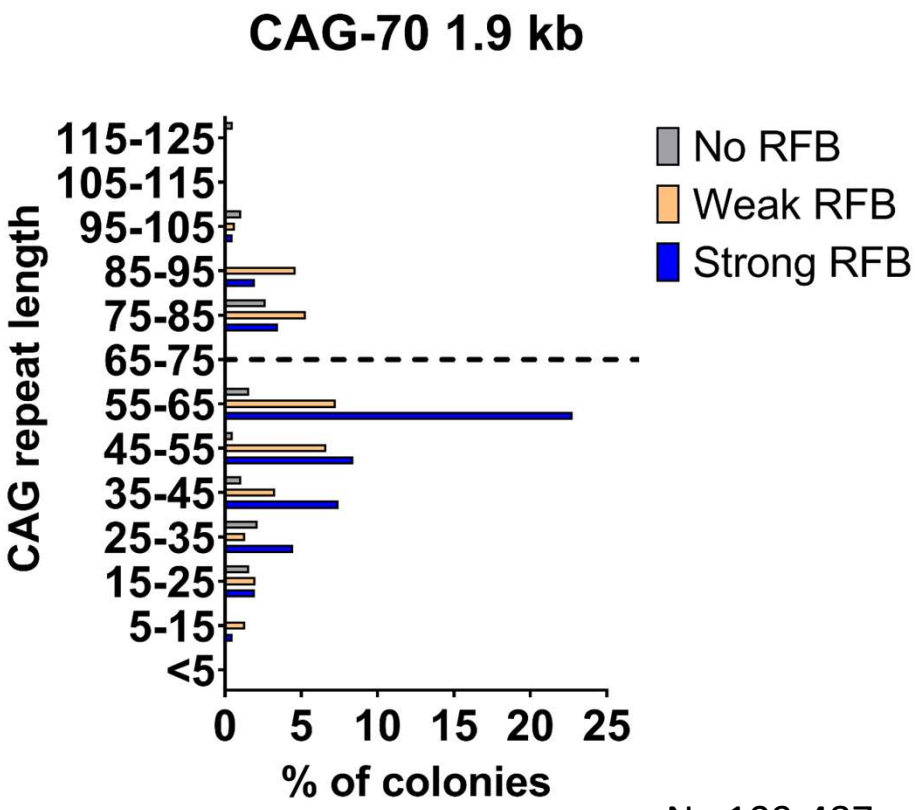
# PCR Assay to detect both CAG Expansions and Contractions

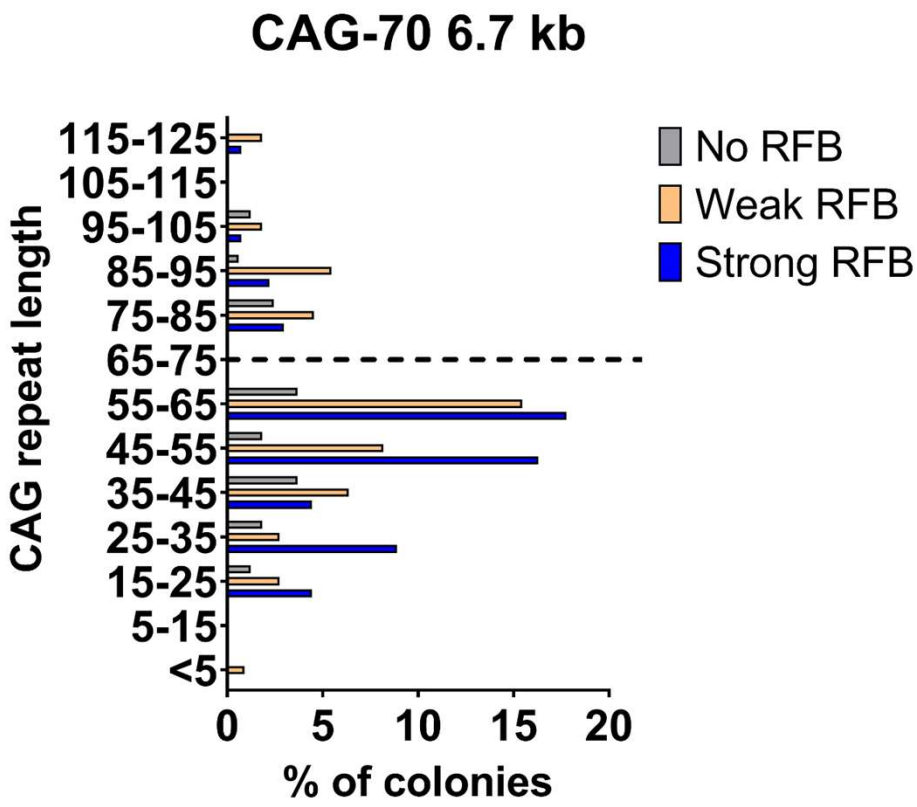


# Distribution of CAG Expansion and Contraction Sizes

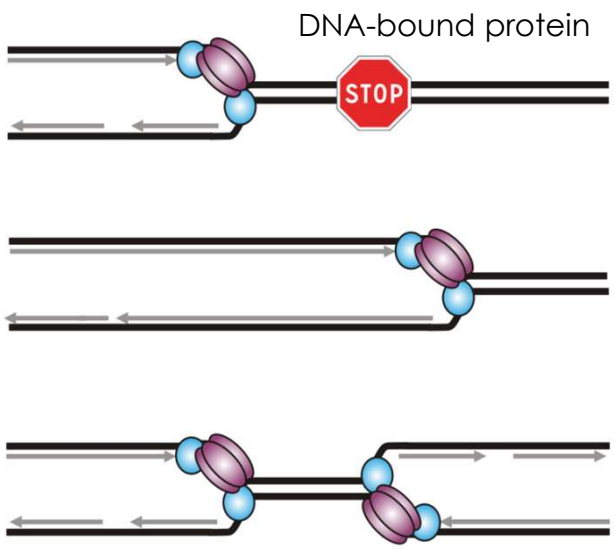
Expansions ranged from +5 to +50 repeats
Contractions ranged from -5 to -65 repeats




N= 109-487 per group

# Mechanisms of replication fork rescue



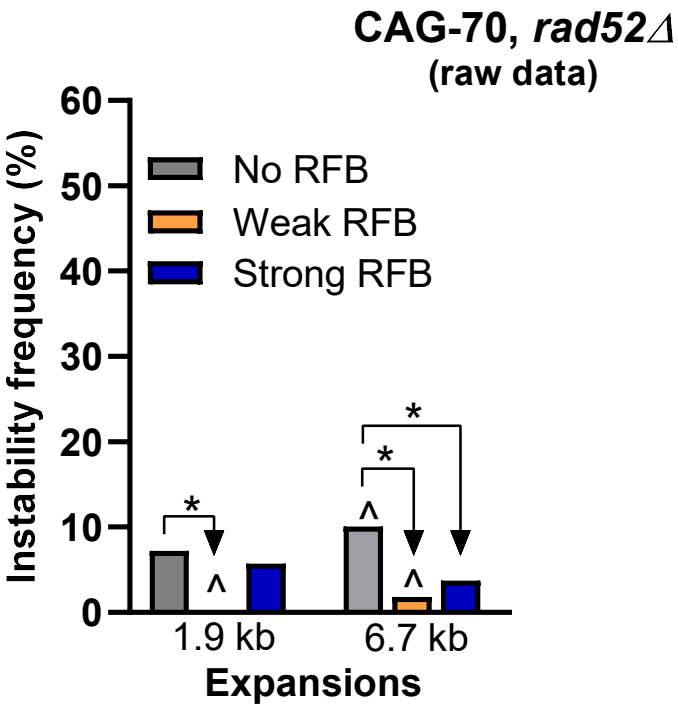

**Fork-arrest**

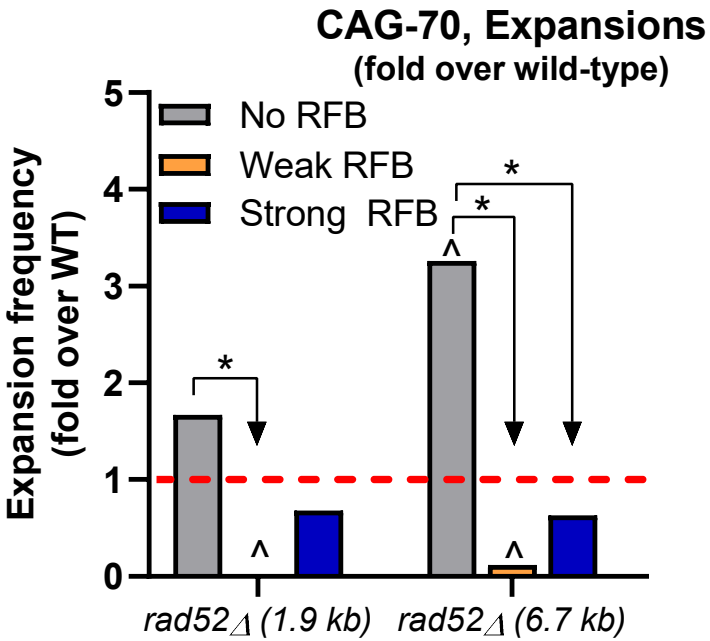
**Fork-restart**

RFB system is **Rad52-dependent**
(Re-priming is another option)

**Rescue by converging fork**
(Firing of dormant origins)
Not **Rad52-dependent**

# Rad52-dependent fork restart is required to cause RFB-dependent CAG expansions




CAG expansions are happening during Rad52-dependent fork restart

Contractions may be occurring by another mechanism

# How are expansions occurring during fork restart (by what mechanism)?

*S. pombe* Rad8
*S. cerevisiae* Rad5
human HTLF
needed for template switch

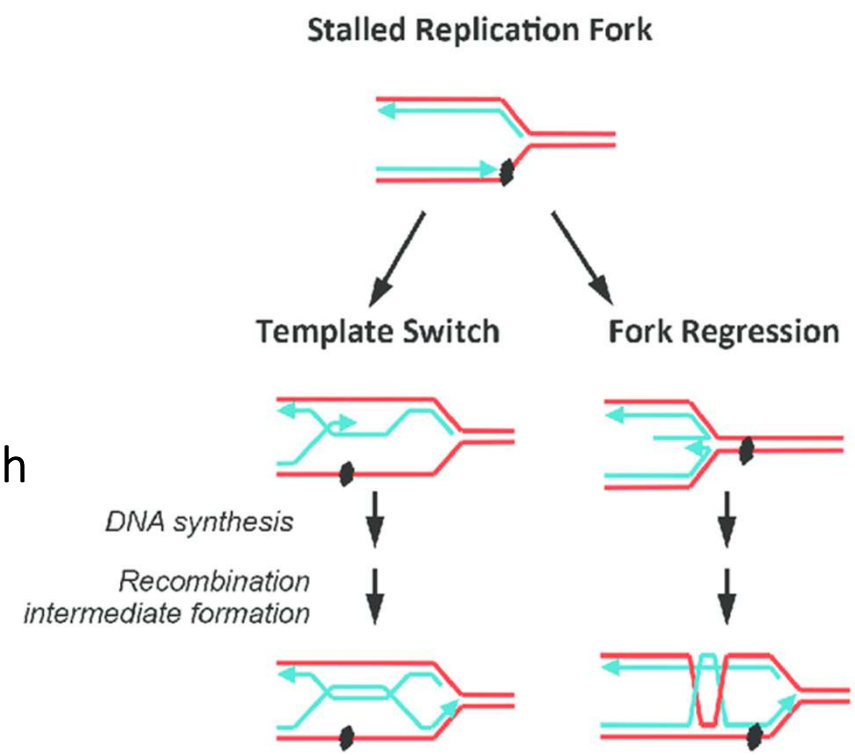
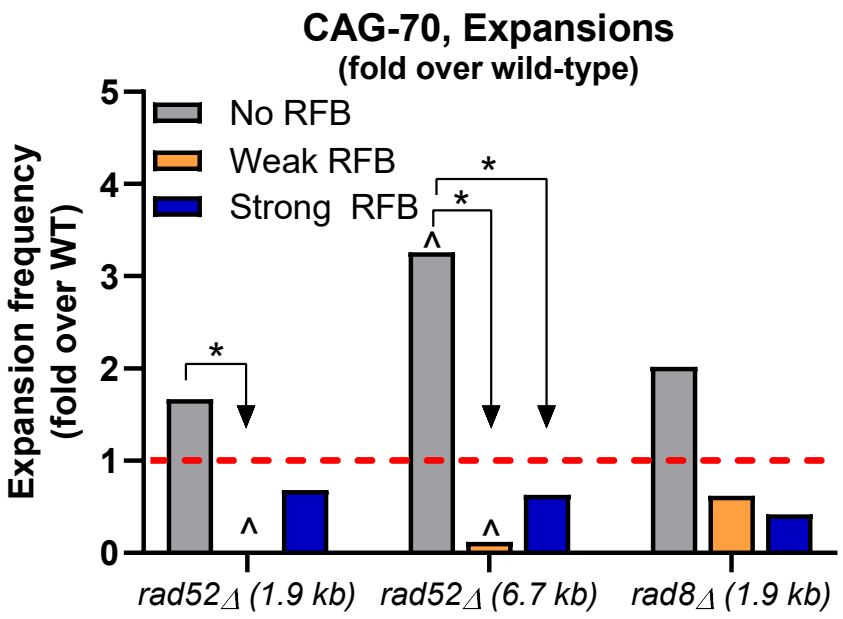
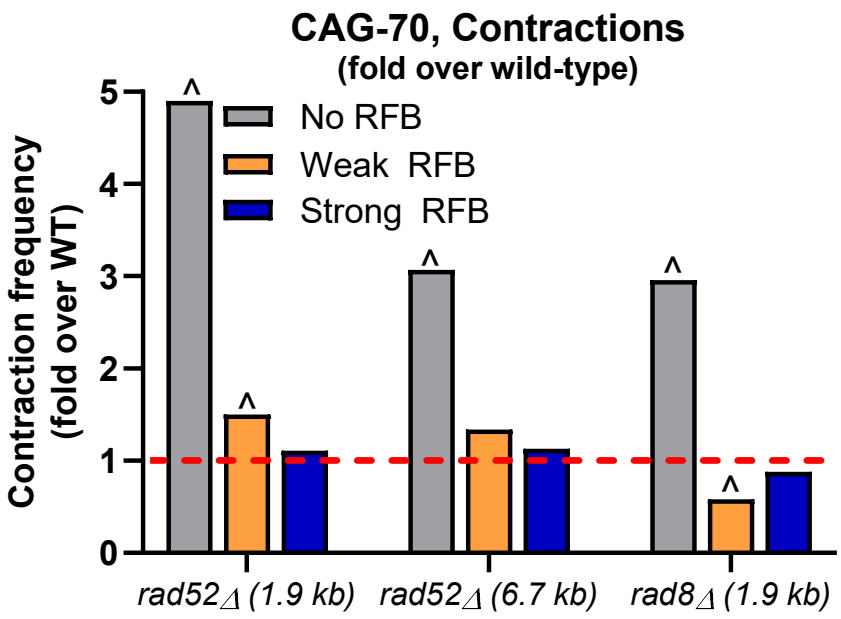



Figure from Bonner & Zhao, 2016

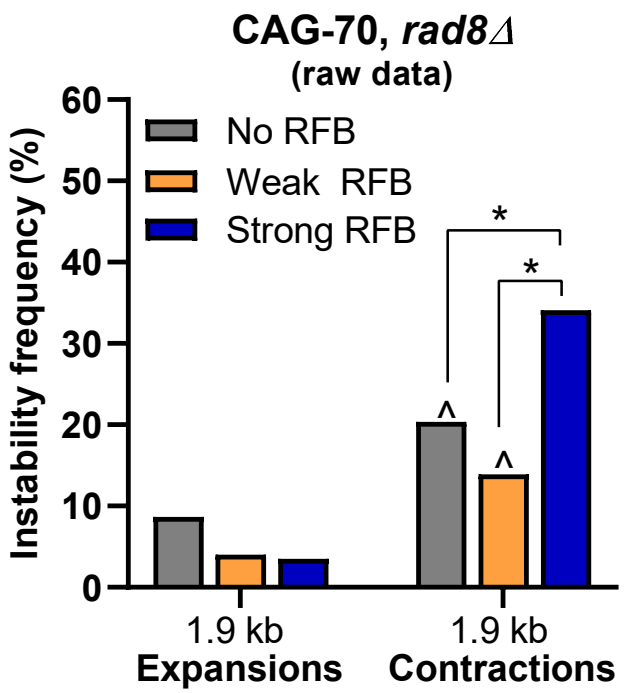
# A Rad8 (*sc*Rad5, hHTLF)-dependent process (template switch) causes some RFB-dependent CAG expansions and contractions



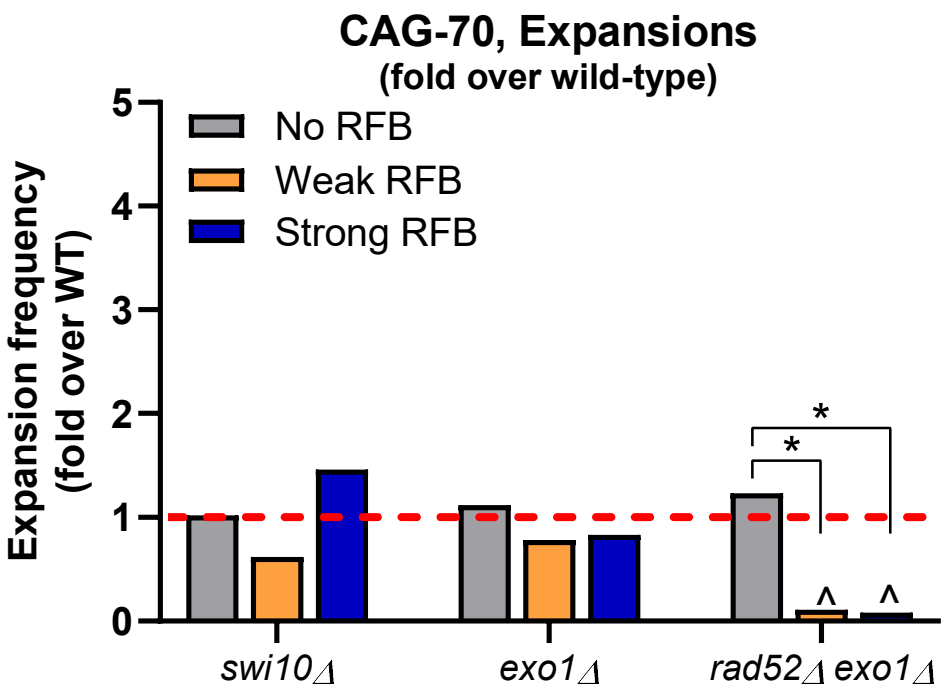



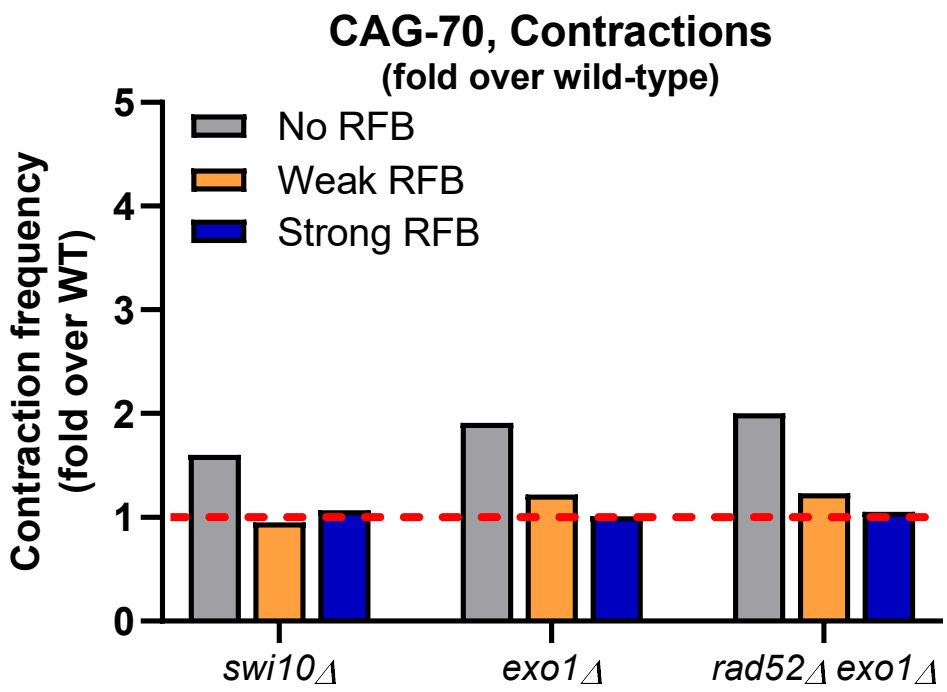

Some CAG expansions and contractions happening during template switch

# A Rad8 (scRad5, hHTLF)-dependent process (template switch) causes RFB-dependent CAG expansions and contractions

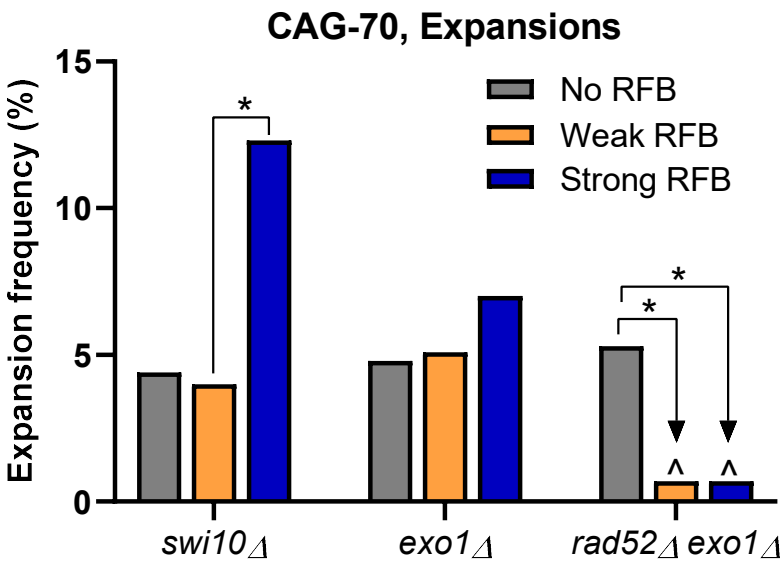


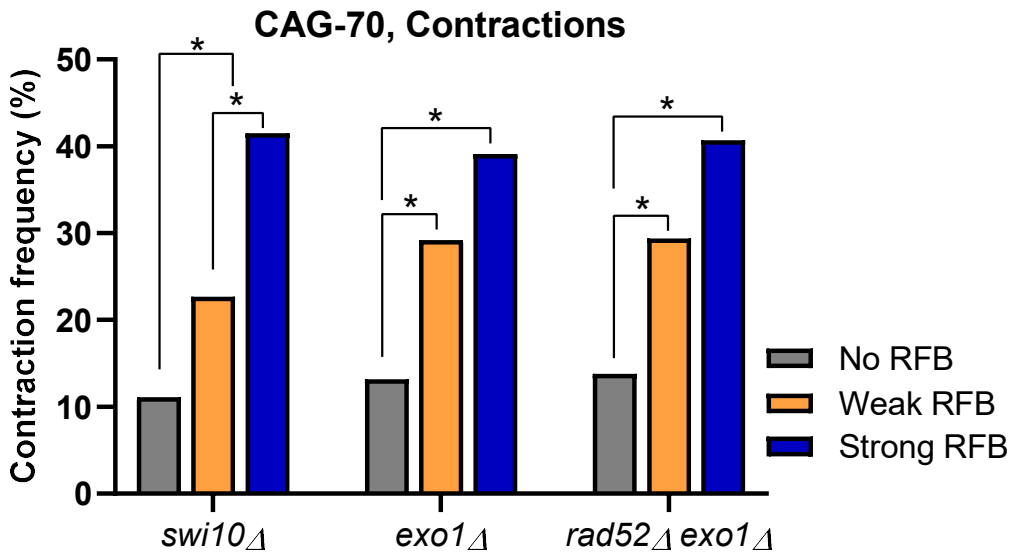
# Expansions and Contractions are not occurring during DSB repair
# -Not dependent on single-strand annealing (Swi10) or resection (Exo1)



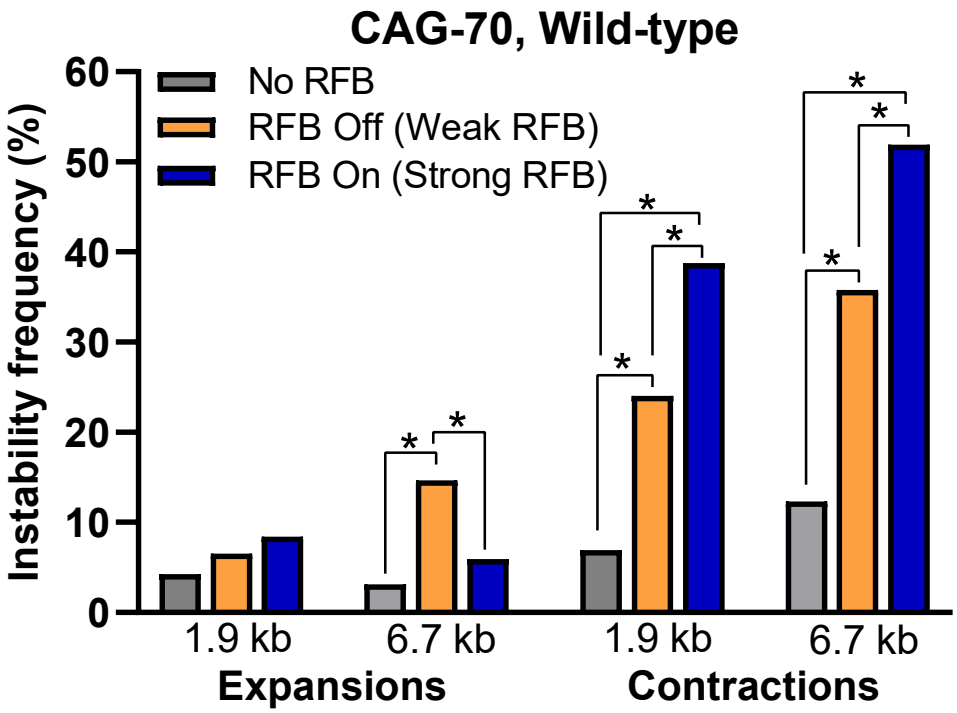



# Expansions and contractions are not occurring during DSB repair -not dependent on SSA (Swi10) or resection (Exo1)



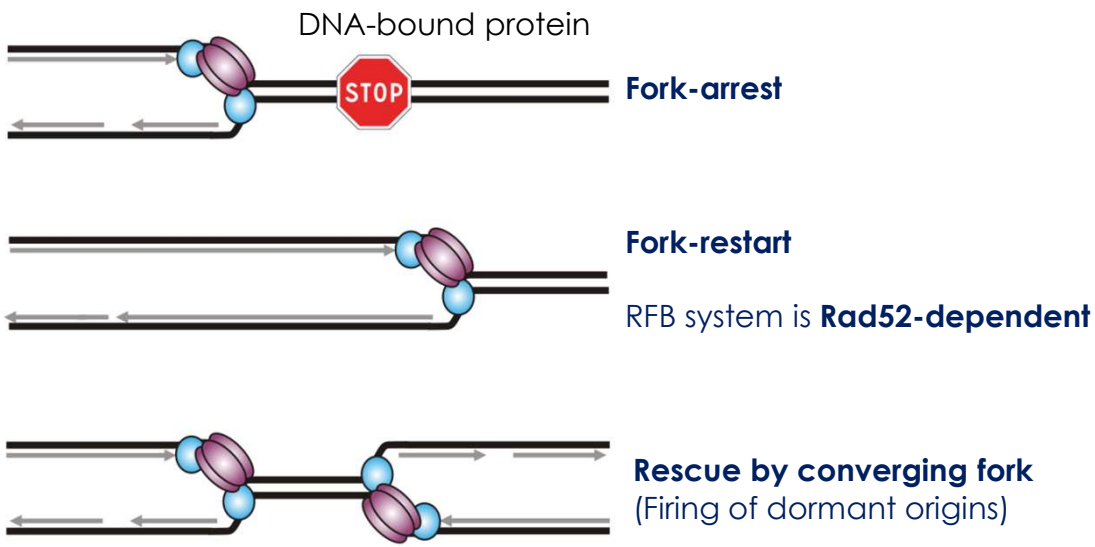
# Why does the RFB cause such a high frequency of contractions?
# -most are not dependent on Rad52-dependent restart



# Mechanisms of replication fork rescue

CAG expansions are Rad52-dependent,
Therefore, happening during fork restart

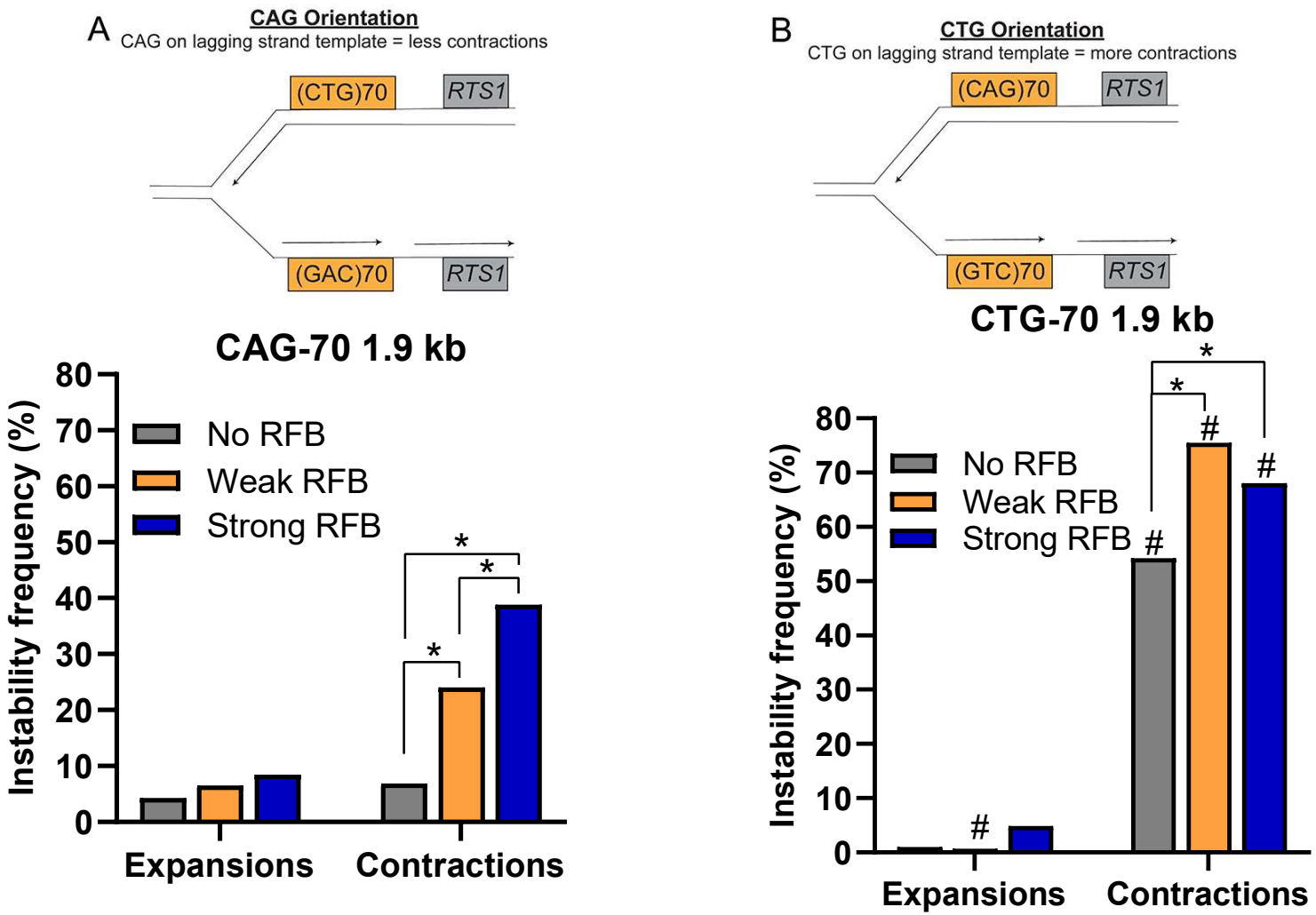
Are CAG contractions happening
during fork rescue?



# Why does the RFB cause such a high frequency of contractions?
# -most are not dependent on Rad52-dependent restart

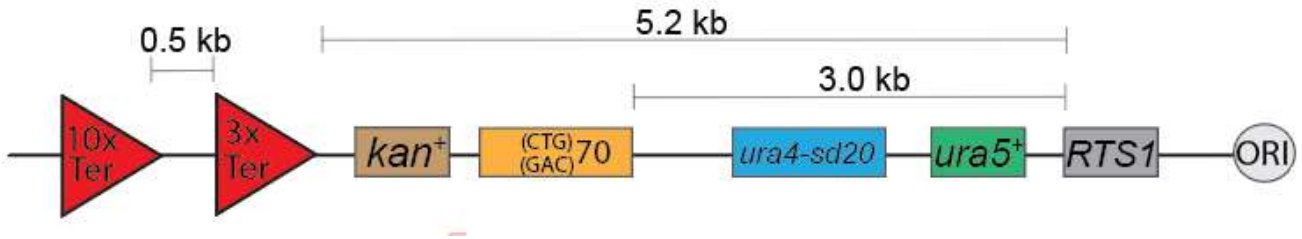

**Replication direction switch model:** CTG forms more stable hairpins than CAG
Therefore, CTG on the lagging strand template may form more structures, leading to contractions

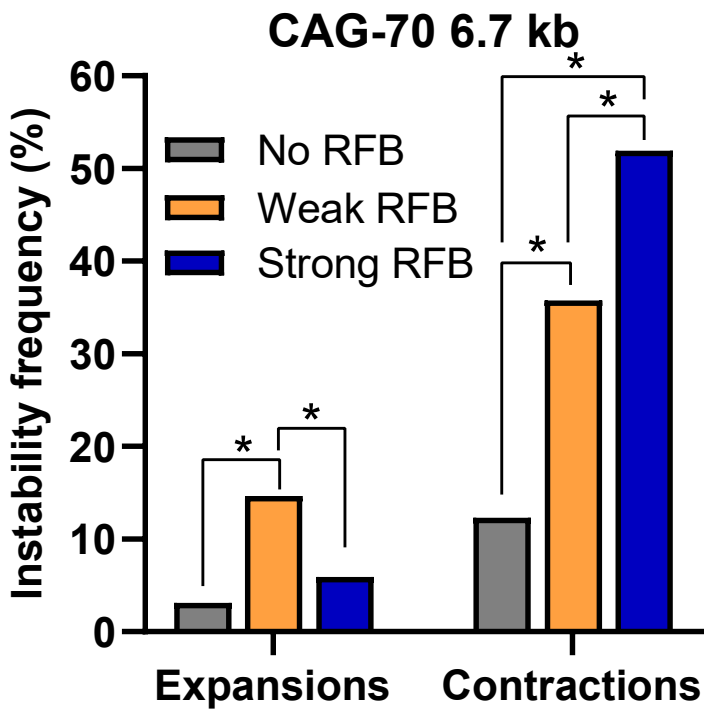
# CTG sequence on the lagging strand template causes a high frequency of contractions: consistent with replication direction switch model

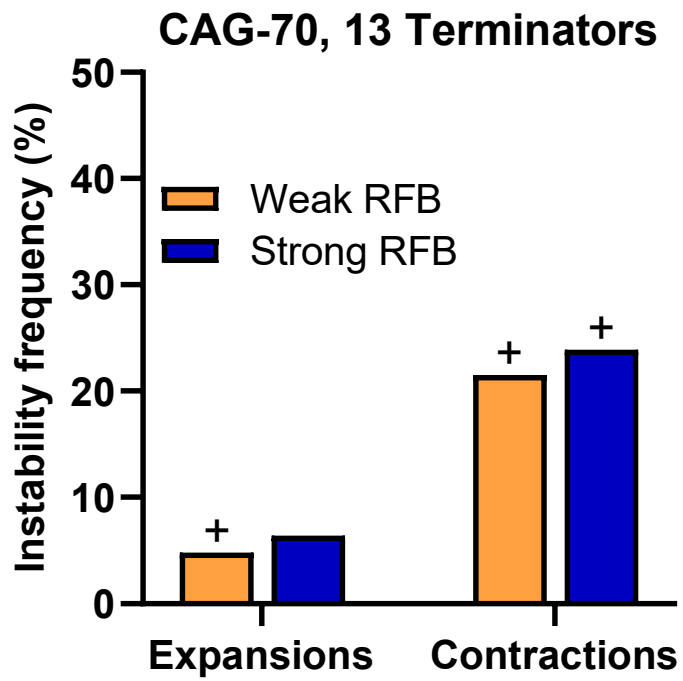


# Testing the replication direction switch model: blocking the converging fork reduces CAG contractions

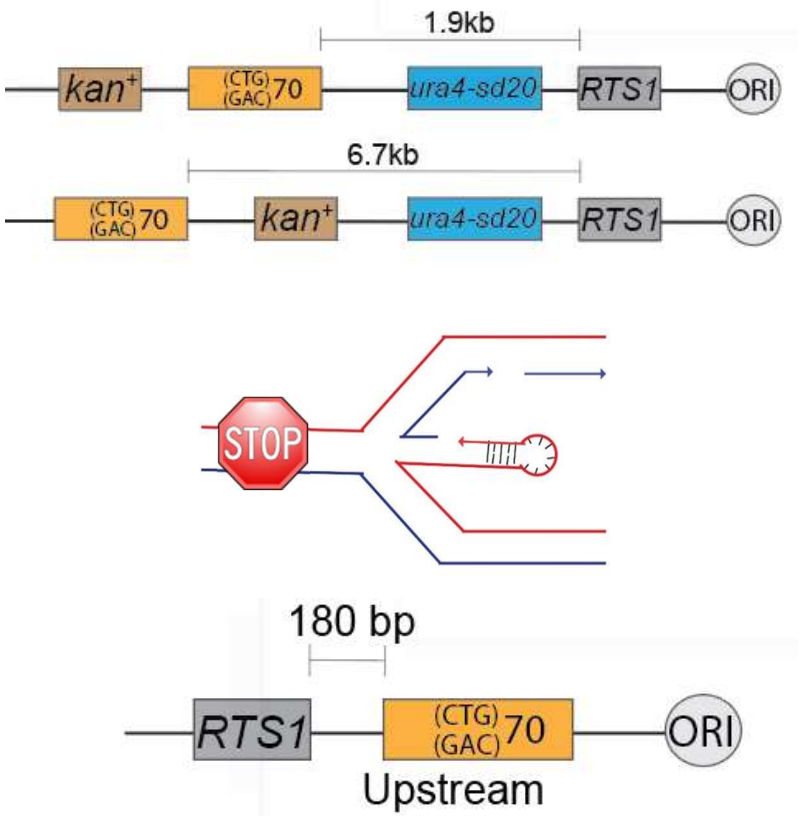
Contractions mostly occur due to a fork switch

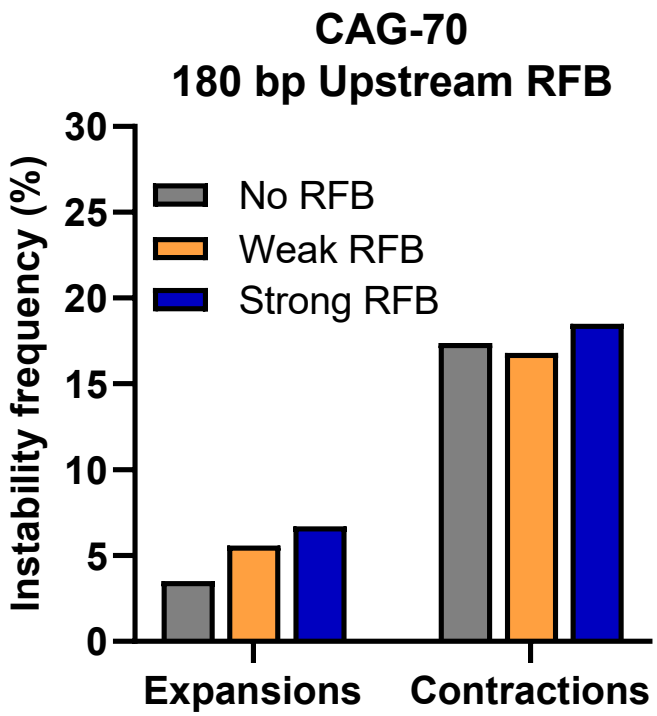


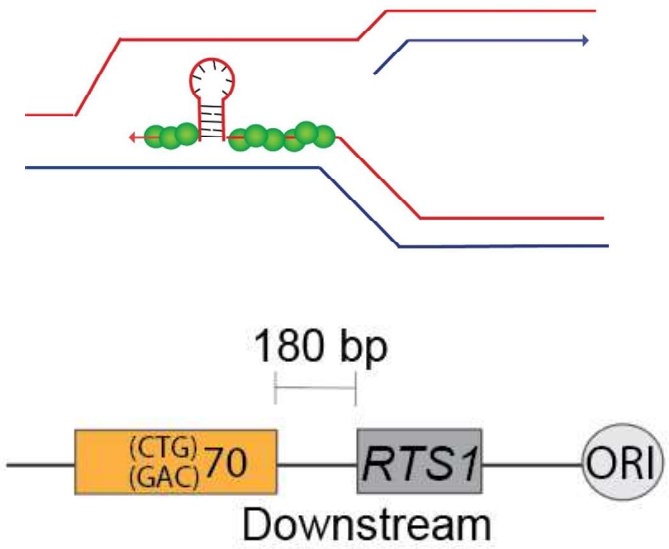
Are expansions dependent on fork reversal or fork resection through the repeat?

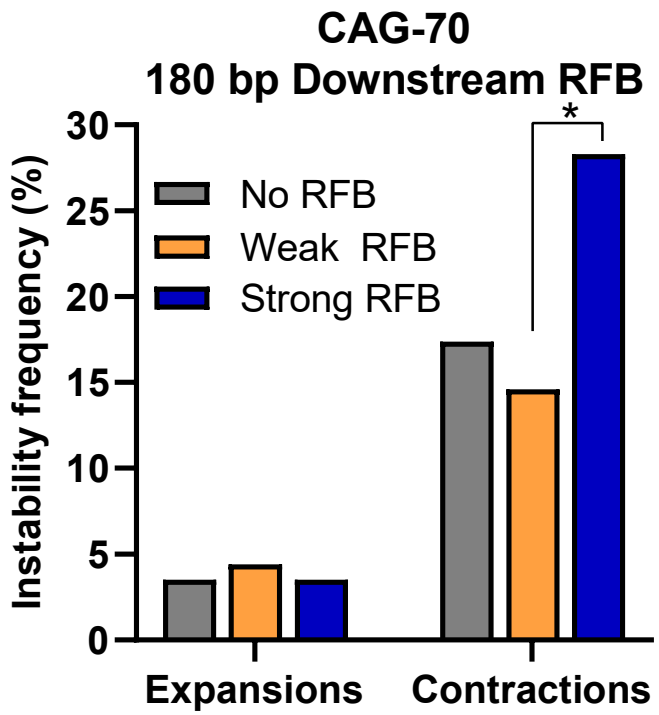



Placing the CAG tract just upstream of the barrier does not cause RFB-dependent instability

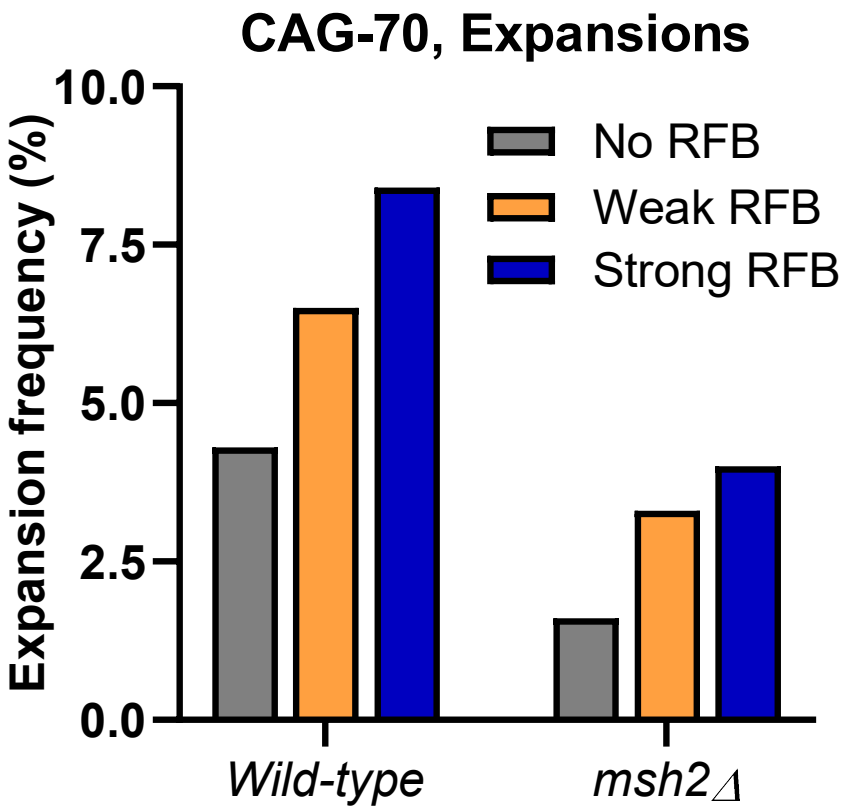
Are expansions occurring during the initial strand invasion phase of restart?

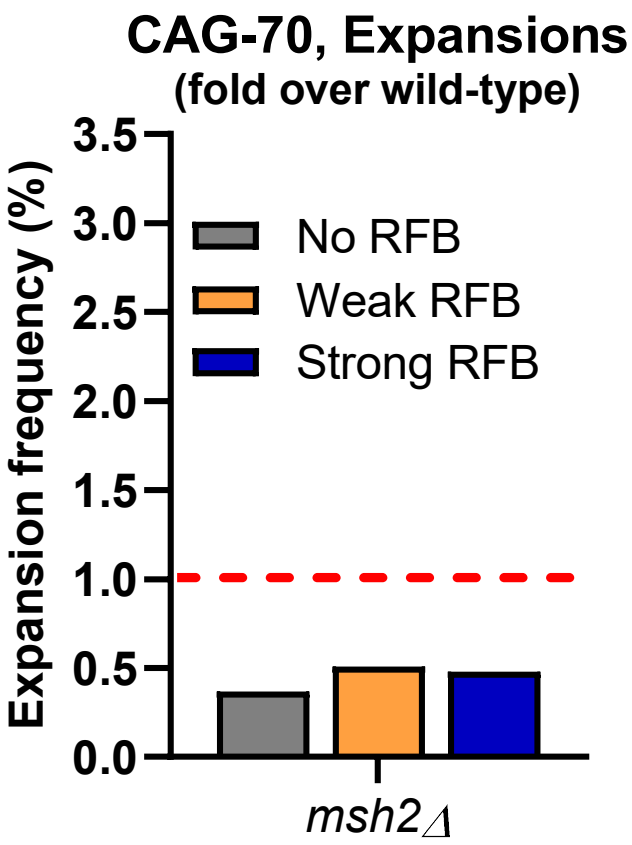


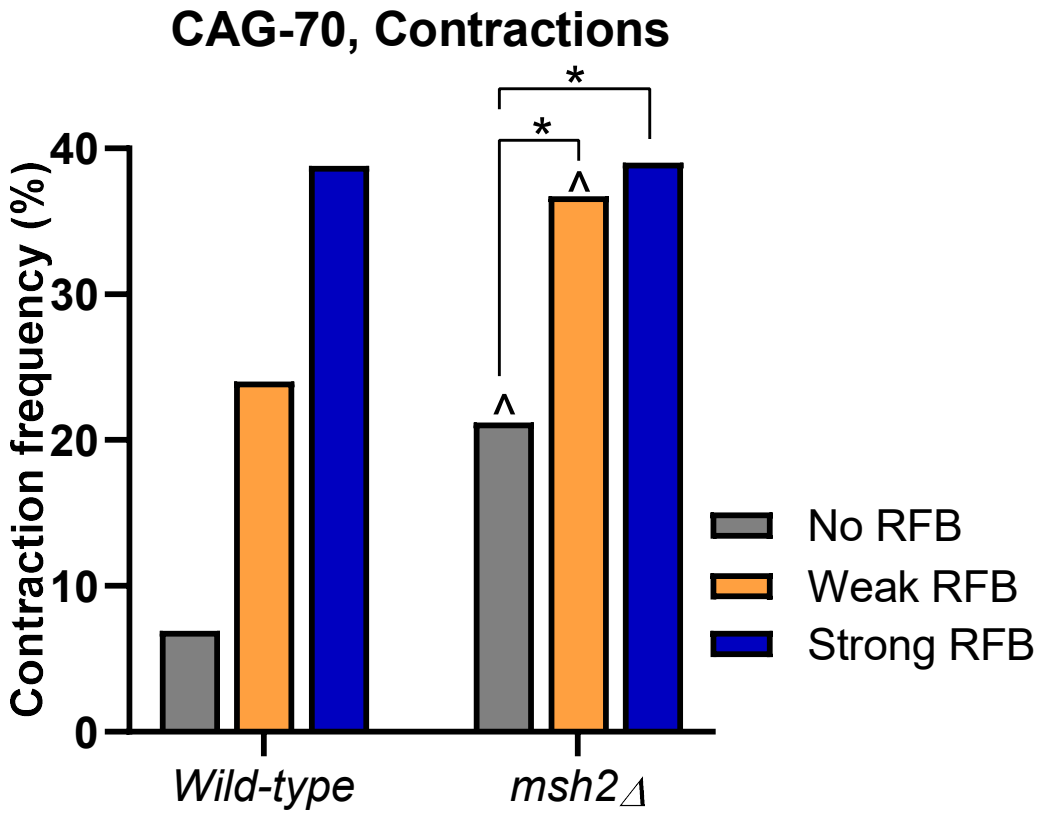

Placing the CAG tract just downstream of the barrier does not cause RFB-dependent instability

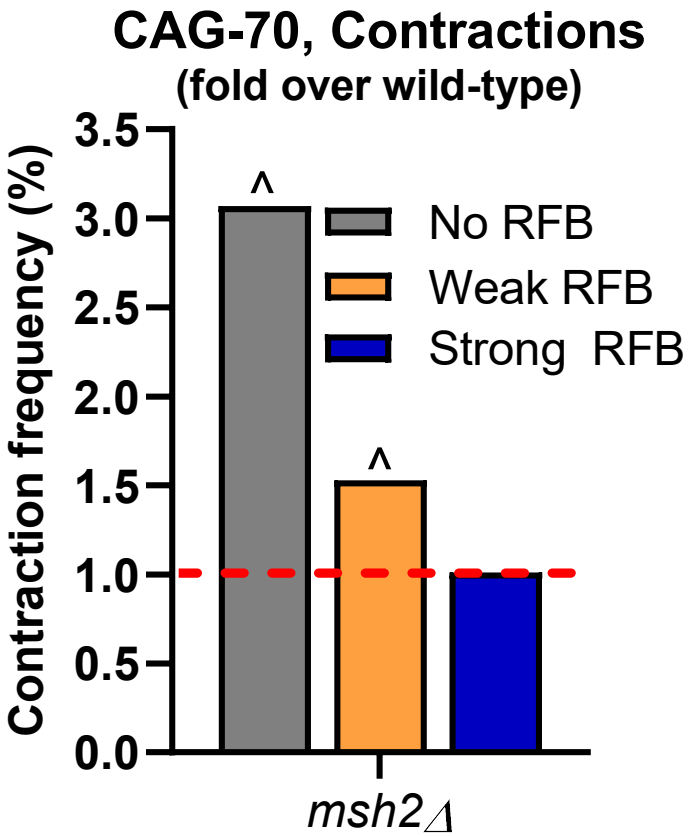
# MutS causes CAG expansions in *S. pombe* (as in humans and mouse models), but this is independent of the RFB



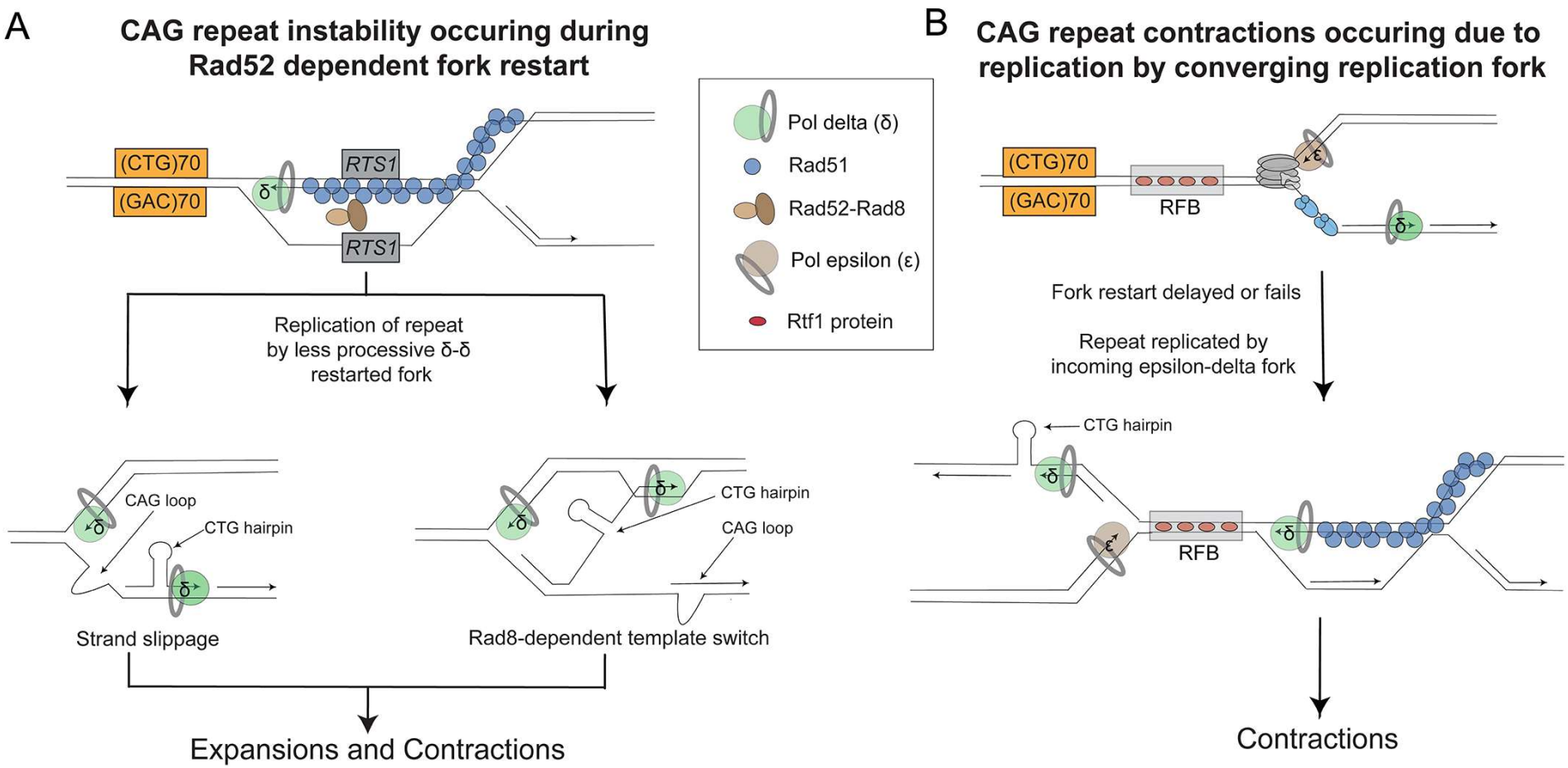



# MutS deletion causes CAG contractionss (as in humans and mouse models), but this is independent of the RFB



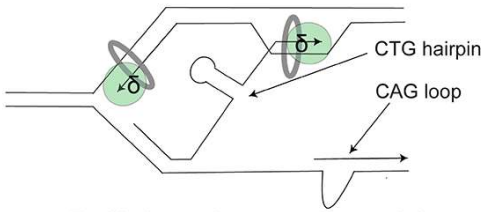



# Model

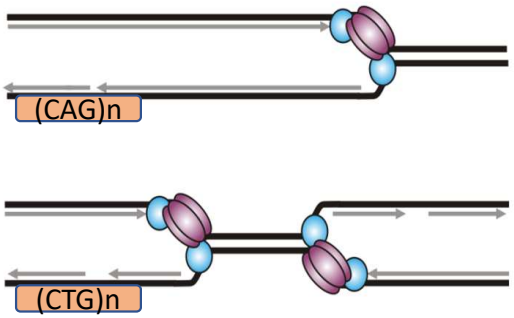

A
CAG repeat instability occuring during Rad52 dependent fork restart
(CTG)70
(GAC)70
RTS1
RTS1
Replication of repeat by less processive δ-δ restarted fork
CAG loop
CTG hairpin
Strand slippage
CTG hairpin
CAG loop
Rad8-dependent template switch
Expansions and Contractions
Pol delta (δ)
Rad51
Rad52-Rad8
Pol epsilon (ε)
Rtf1 protein
B
CAG repeat contractions occuring due to replication by converging replication fork
(CTG)70
(GAC)70
RFB
Fork restart delayed or fails
Repeat replicated by incoming epsilon-delta fork
CTG hairpin
RFB
Contractions

# Take Home Messages for Replication Fidelity:

- Restarted replication forks are especially prone to slippage through repetitive DNA tracts
- Increased template switching by the uncoupled fork can lead to repeat expansions
- Replication fork barriers in genomes can lead to a change in fork direction, and this has implications for the stability of structure-forming sequences
- This mechanism of repeat instability could be relevant to barriers caused either by the repeat itself, or neighboring barriers.
  - Cancer cells have altered replication programs and rely heavily on replication restart mechanisms



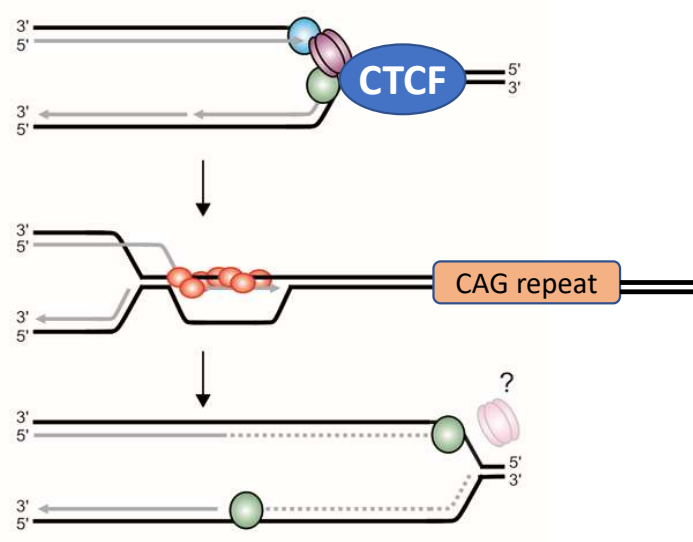
Rad8-dependent template switch



# Implications for Repeat Expansion Diseases:

- Transient replication fork barriers can drive repeat instability – could this explain the restricted developmental time window of intergenerational expansions?
  - Cell type or timing-specific barriers could lead to changes in the replication profile of the repeat locus

- Many expandable CAG/CTG tracts are near CTCF (chromatin insulator) binding sites that are associated with Topologically Associated Domains (TADs)
  - CTCF may cause fork slowing through the Myotonic Dystrophy locus (Cleary..Pearson, 2010)




Cleary JD, Tomé S, López Castel A, Panigrahi GB, Foiry L, Hagerman KA, Sroka H, Chitayat D, Gourdon G, Pearson CE. **Tissue- and age-specific DNA replication patterns at the CTG/CAG-expanded human myotonic dystrophy type 1 locus.** Nat Struct Mol Biol. 2010 Sep;17(9):1079-87.

# Acknowledgements

Thanks to Sarah Lambert for hosting me and being a terrific collaborator!

Catherine and Sarah at the Recombination meeting in Alicante, Spain

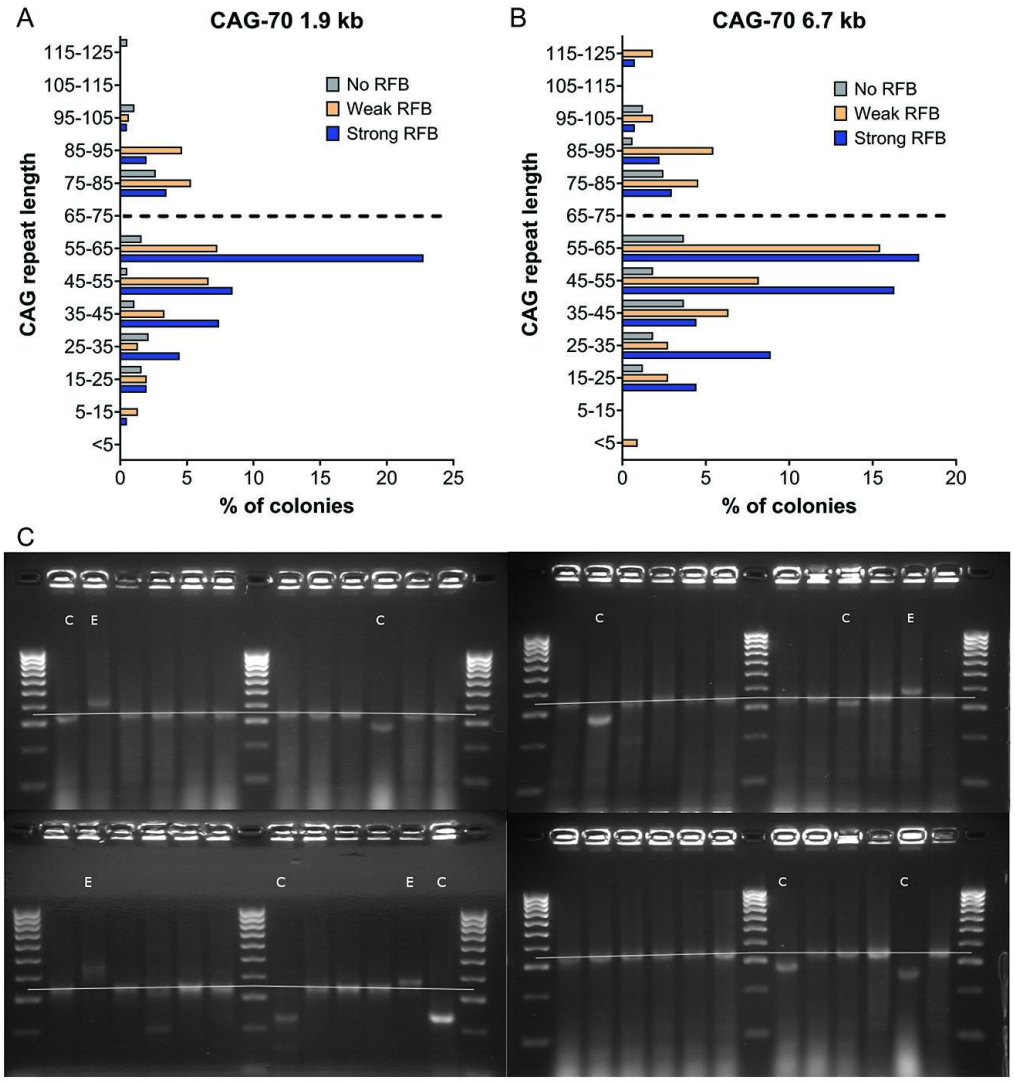
Freudenreich lab members:
Michaela Gold (undergrad at the time)
Zixin Hong (Tufts undergrad)
Jenna Whalen (Grad student at time)

Lambert lab members:
Karine Freon (Research Associate)
Ismail Iraqui (post-doc)

Funding:
**CHF**: NIH grant GM122880; Mayent Rothschild fellowship from Institut Curie

**SAL**: Ligue contre le cancer "Equipe Labellisée 2020 (EL2020LNCC/Sal)

**MAG**: Tufts Undergraduate Research Award; Tufts University Sandler International Research Program award to M.A.G., J.M.W and C.H.F.


A
CAG-70 1.9 kb
No RFB
Weak RFB
Strong RFB
CAG repeat length
115-125
105-115
95-105
85-95
75-85
65-75
55-65
45-55
35-45
25-35
15-25
5-15
<5
0
5
10
15
20
25
% of colonies
B
CAG-70 6.7 kb
No RFB
Weak RFB
Strong RFB
CAG repeat length
115-125
105-115
95-105
85-95
75-85
65-75
55-65
45-55
35-45
25-35
15-25
5-15
<5
0
5
10
15
20
% of colonies
C
C E
C
C
C E
E
C
E C
C
C

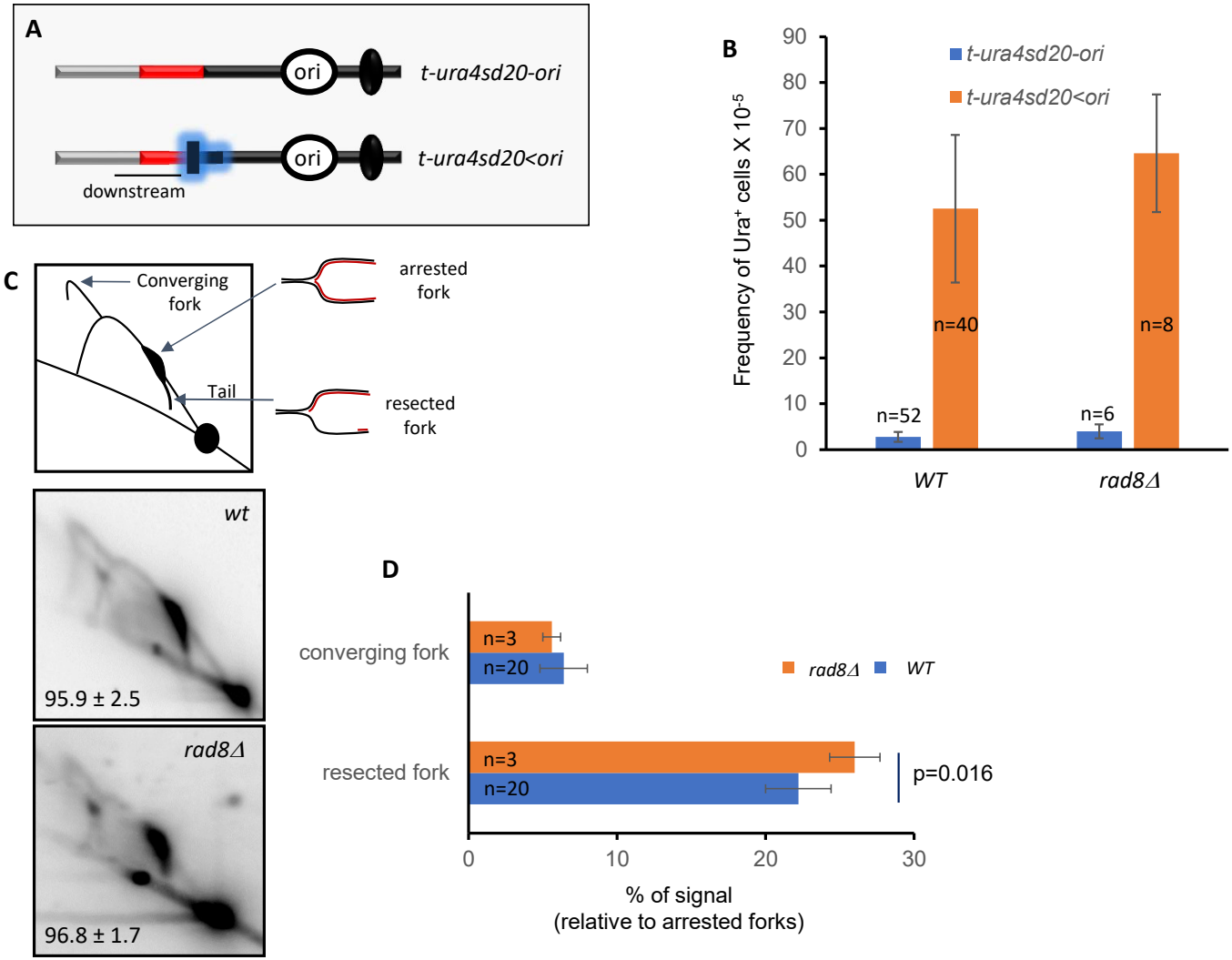

A
ori
t-ura4sd20-ori
ori
t-ura4sd20<ori
downstream
B
Frequency of Ura+ cells X 10-5
90
80
70
60
50
40
30
20
10
0
t-ura4sd20-ori
t-ura4sd20<ori
n=52
n=40
n=6
n=8
WT
rad8Δ
C
Converging fork
Tail
arrested fork
resected fork
wt
95.9 ± 2.5
rad8Δ
96.8 ± 1.7
D
converging fork
resected fork
n=3
n=20
n=3
n=20
rad8Δ
WT
p=0.016
0
10
20
30
% of signal
(relative to arrested forks)

Figure S3