HHS Public Access

Author manuscript

Science. Author manuscript; available in PMC 2019 August 03.

Published in final edited form as:

Science. 2019 January 04; 363(6422): 81–84. doi:10.1126/science.aan1425.

DNA fragility in the parallel evolution of pelvic reduction in stickleback fish

Kathleen T. Xie[1,2], **Guliang Wang**[3], **Abbey C. Thompson**[1], **Julia I. Wucherpfennig**[1], **Thomas E. Reimchen**[4], **Andrew D. C. MacColl**[5], **Dolph Schluter**[6], **Michael A. Bell**[7], **Karen M. Vasquez**[3], **David M. Kingsley**[1,2,*]

[1]Stanford University School of Medicine, Stanford, CA, USA

[2]Howard Hughes Medical Institute, Stanford, CA, USA

[3]University of Texas at Austin, Austin, TX USA

[4]University of Victoria, Victoria, BC, Canada

[5]University of Nottingham, Nottingham, UK

[6]University of British Columbia, Vancouver, BC, Canada

[7]University of California Museum of Paleontology, Berkeley, CA, USA

Abstract

Evolution generates a remarkable breadth of living forms, but many traits evolve repeatedly, by still poorly understood mechanisms. A classic example of repeated evolution is the loss of pelvic hindfins in stickleback fish (*Gasterosteus aculeatus*). Repeated pelvic loss maps to recurrent deletions of a pelvic enhancer of the *Pitx1* gene. Here, we identify molecular features contributing to these recurrent deletions. *Pitx1* enhancer sequences form alternative DNA structures in vitro, and increase double-strand breaks and deletions in vivo. Enhancer mutability depends on DNA replication direction and is caused by (TG)-dinucleotide repeats. Modeling shows that elevated mutation rates can influence evolution under demographic conditions relevant for sticklebacks and humans. DNA fragility may thus help explain why the same loci are often used repeatedly during parallel adaptive evolution.

ONE SENTENCE SUMMARY:

DNA fragility and high mutation rates influence the genomic pathways of adaptive evolution.

*Correspondence: kingsley@stanford.edu.

Author contributions: K.T.X and D.M.K designed the study. K.T.X., G.W., A.C.T., and J.I.W. performed experiments. K.T.X, G.W., A.C.T., D.S., K.M.V., and D.M.K. analyzed data. T.E.R., A.D.C.M, D.S., and M.A.B. provided key populations and comments. K.T.X and D.M.K wrote the paper with input from all authors.

Competing interests: None.

Data and materials availability: GEO accession GSE121537.

MAIN TEXT

Many phenotypic traits evolve repeatedly in organisms adapting to similar environments, and studying these cases can reveal ecological and genetic factors shaping parallel evolution (1, 2). For example, loss of pelvic appendages has evolved repeatedly in mammals, amphibians, reptiles, and fishes. Marine stickleback fish (*Gasterosteus aculeatus*) develop a robust pelvic apparatus, whereas many freshwater populations have lost pelvic structures (3). Pelvic reduction is associated with particular ecological conditions, is likely adaptive, and maps to recurrent and independent deletions of a pelvic enhancer (*Pel*) upstream of the homeodomain transcription factor gene (*Pitx1*) that also show repeatable molecular signatures of positive selection (4–7). This unusual spectrum of regulatory deletions contrasts with the accumulation of single nucleotide changes in other studies (8, 9), hinting that special DNA features may shape adaptive variation at the *Pitx1* locus (6).

Pel enhancer sequences show high predicted helical twist flexibility (6), a DNA feature associated with delayed replication and fragile site instability (10). To examine whether *Pel* forms alternative DNA structures in vitro, we used 2-dimensional electrophoresis to analyze distributions of plasmid topoisomers (11) (Fig. 1A). A control stickleback genomic region showed smooth curves characteristic of B-DNA (Fig. 1B). In contrast, *Pel* sequences from marine populations showed mobility shifts characteristic of alternative DNA structure formation (Fig. 1B). Structural transitions started at a negative superhelical density of $-\sigma = 0.043$ and changed apparent linking numbers by 10–16 helical turns, similar to shifts produced by Z-DNA (left-handed DNA, starting $-\sigma = 0.046$) occupying ~105–170 bp (12, 13). *Pel* sequences from pelvic-reduced populations did not show unusual electrophoretic transitions (Fig. 1B), suggesting that natural mutations remove sequences forming alternative DNA structures.

To test the effect of *Pel* sequences on chromosome stability in vivo, we measured the rate of DNA double-strand breaks in yeast artificial chromosomes (Fig. 2A). Constructs without added test regions broke at background rates of 3.37 breaks per 10^6 divisions (Fig. 2B), consistent with previous reports (14). Chromosomes containing marine *Pel* broke ~25–50 times more frequently (Fig. 2B), a rate even higher than previously analyzed human fragile sites (14). *Pel* from freshwater pelvic-reduced populations (but not freshwater pelvic-complete populations (fig. S1)) broke at rates similar to the control (Fig. 2B), suggesting that natural *Pel* mutations remove breakage-prone regions.

Reverse complements of marine *Pel* broke ~10–20 times less frequently than identical sequences in the forward orientation (Fig. 2B). RNA transcription can influence fragile site breakage (15), but reversing transcription orientation of the nearby *URA3* marker did not significantly affect *Pel* fragility (Fig. 2C). In contrast, adding a replication origin on the opposite side of *Pel* did switch fragility, making the forward sequence stable, and the reverse complement fragile (Fig. 2C). Thus, *Pel* fragility is markedly dependent on DNA replication direction.

Pel contains abundant runs of alternating pyrimidine-purine repeats (Fig. 3A, file S1), which can adopt alternative structures like Z-DNA, previously associated with deletions in bacteria,

mice, and humans (16, 17). Three stretches of ~15, ~20, and ~50 (TG)-dinucleotide repeats in marine *Pel* total ~170 bp (consistent with linking number changes seen in topoisomer assays above). TG-repeats alone induced mobility shifts in topoisomer assays (Fig. 3B) (18) and elevated chromosome breakage in yeast, with longer repeats stimulating more breaks (Fig. 3C). In contrast, both long and short versions of the reverse complement sequence (CA-repeats) were stable (Fig. 3C), recapitulating the orientation dependence of *Pel* fragility.

We also tested the effect of TG- and CA-repeats in mammalian COS-7 cells (Fig. 3D) (19). Dinucleotide repeats elevated mutation frequencies, with TG-repeats being more mutagenic than CA-repeats of comparable lengths, and longer repeats being more mutagenic than shorter repeats (Fig. 3E), consistent with results from yeast assays. Mutations stimulated by the most mutagenic sequence, $(TG)_{41}$, were predominantly >100 bp deletions that removed part or all of the repeat and adjacent reporter gene (Fig. 3F, fig. S2A). Approximately 70% of deletion junctions contained microhomologies and insertions (Fig. 3F, fig. S2A–B), consistent with error-prone microhomology-mediated end-joining repair, and similar to junctions seen in stickleback pelvic-reduction alleles (6) (Fig. 3A). Ligation-mediated PCR suggested that breaks initiated near the dinucleotide repeats (fig. S2C). Taken together, our results indicate that TG-repeats form alternative DNA structures in vitro and can recapitulate the high mutation rates, orientation-dependence, and propensity to stimulate breaks and deletions of the full *Pel* region.

To determine the orientation of *Pel* sequences relative to DNA replication in sticklebacks (Fig. 4A, fig. S3), we sequenced S- and G-phase cells from developing embryos and calculated S:G read-depth ratios to determine replication timing (20). *Pel* is located in a timing transition region (Fig. 4B, fig. S4), consistent with unidirectional replication. The replication direction through *Pel* matches the fragile orientation (Fig. 4C), suggesting that *Pel* would form a TG-repeat-associated fragile site in vivo. Experimental CRISPR targeting confirmed that initiation of breaks in *Pel* were sufficient to trigger local DNA deletions and macroscopic loss of pelvic structures in genetic crosses (fig. S5).

Could elevated mutation rates contribute to reuse of *Pel* deletions in parallel evolution? Population genetic modeling indicates that new mutations occurring at the low rates of typical single nucleotide changes ($\sim10^{-9}$) would rarely arise at a particular locus in postglacial stickleback populations, while mutations occurring at elevated rates ($\sim10^{-5}$ for fragile sites) would arise often. When new mutations do occur, their subsequent fate is controlled by drift and selection (21). Neutral or small-effect point mutations will usually be lost or rise to fixation slowly, while deletions may cause larger phenotypic effects and can sweep if environmental conditions favor pelvic reduction (Fig. 4D, fig. S6, fig. S7). The combined effects on both the “arrival of the fittest” and the “survival of the fittest” may explain why recurrent *Pel* deletions are the predominant mechanism for evolving stickleback pelvic reduction. For other traits, ancient standing variants provide an alternative way to overcome the demographic constraints of waiting for *de novo* mutations in small populations, and can also lead to reuse of similar alleles in different populations (22, 23).

The demographic parameters typical of sticklebacks apply to many vertebrates evolving with small population sizes or facing rapid environmental changes. For example, migration of modern humans out of Africa occurred with relatively small populations adapting to new environments in 3,000 generations or less (24). Interestingly, nearly half of currently known mutations underlying adaptive traits in modern humans also appear to be produced by mechanisms with elevated mutation rates (table S1).

High mutation rates have been described at contingency loci in bacteria and other systems (25–30). Our studies add an important new example of DNA fragility contributing to repeated morphological evolution in vertebrates. Our data also highlight several mechanisms that could alter local mutation rates, including expansion/contraction of TG-repeats, changes in sequence orientation, or changes in DNA replication. Natural variation in such parameters may affect the evolvability of different loci and the particular genetic paths likely to be taken when ecological conditions favor a given phenotype. The sequence features associated with DNA fragility in the *Pel* region are also found in thousands of other positions in stickleback and human genomes (fig. S8). Notably, TG-repeats are enriched in other loci that have undergone recurrent ecotypic deletions during marine-freshwater stickleback evolution (31) (table S2, fig. S9), and near DNA breakage sites in humans (fig. S10). As causative changes are identified for more phenotypic traits, it will be interesting to see the extent to which DNA fragility has influenced the genes and mutations that underlie evolutionary change in nature.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGEMENTS

We thank V. Tien, J. Le, M. Yau, M. Thakur, A. Muralidharan, M. Whitlock, B. Belotserkovskii, R. Driscoll, K. Cimprich, J. Wang, S. Quake, and A. Casper for experimental assistance or advice; R. Daugherty, J. Rollins, B. Lohman, R. Mollenhauer, M. Reyes, and F. von Hippel for help with fieldwork; C. Freudenreich for yeast strains; Z. Weng and B. Carter for help with high-throughput sequencing and cell sorting.

Funding: NIH grants 5P50HG2568 (D.M.K.), CA093729 (K.M.V.), 2T32GM007790 (J.I.W); NSF grant DEB0919184 (M.A.B.); NSF and Stanford CEHG Graduate Fellowships (K.T.X.), NIH Predoctoral Fellowship (ACT); HHMI investigator (D.M.K.).

REFERENCES AND NOTES

1. Schluter D, Clifford EA, Nemethy M, McKinnon JS. Am Nat 163, 809–822 (2004). [PubMed: 15266380]
2. Stern DL, Orgogozo V. Science 323, 746–751 (2009). [PubMed: 19197055]
3. Bell MA. Biol J Linn Soc 31, 347–382 (1987).
4. Reimchen TE. Can J Zool 58, 1232–1244 (1980).
5. Bell MA, Orti G, Walker JA, Koenings JP. Evolution 47, 906–914 (1993). [PubMed: 28567888]
6. Chan YF et al. Science 327, 302–305 (2010). [PubMed: 20007865]
7. Karhunen M, Merila J, Leinonen T, Cano JM, Ovaskainen O. Mol Ecol Resour 13, 746–754 (2013). [PubMed: 23656704]
8. Prud'homme B et al. Nature 440, 1050–1053 (2006). [PubMed: 16625197]
9. Stern DL, Frankel N. Philos T R Soc B 368, (2013).

10. Thys RG, Lehman CE, Pierce LC, Wang YH. Curr Genomics 16, 60–70 (2015). [PubMed: 25937814]
11. Bowater R, Aboul-Ela F, Lilley DM. Methods Enzymol 212, 105–120 (1992). [PubMed: 1518443]
12. Nordheim A, Rich A. Proc Natl Acad Sci USA 80, 1821–1825 (1983). [PubMed: 6572943]
13. Rich A, Nordheim A, Wang AH. AnnuRevBiochem 53, 791–846 (1984).
14. Zhang H, Freudenreich CH. Mol Cell 27, 367–379 (2007). [PubMed: 17679088]
15. Helmrich A, Ballarino M, Tora L. Mol Cell 44, 966–977 (2011). [PubMed: 22195969]
16. Wang G, Christensen LA, Vasquez KM. Proc Natl Acad Sci USA 103, 2677–2682 (2006). [PubMed: 16473937]
17. Wang G, Carbajal S, Vijg J, DiGiovanni J, Vasquez KM. J Natl Cancer Inst 100, 1815–1817 (2008). [PubMed: 19066276]
18. Hamada H, Petrino MG, Kakunaga T, Seidman M, Stollar BD. Mol Cell Biol 4, 2610–2621 (1984). [PubMed: 6098814]
19. Wang G, Vasquez KM. Proc Natl Acad Sci USA 101, 13448–13453 (2004). [PubMed: 15342911]
20. Rhind N, Gilbert DM. Cold Spring Harb Perspect Biol 5, a010132 (2013).
21. Kimura M. Genetics 47, 713–719 (1962). [PubMed: 14456043]
22. Colosimo PF et al. Science 307, 1928–1933 (2005). [PubMed: 15790847]
23. Barrett RD, Schluter D. Trends Ecol Evol 23, 38–44 (2008). [PubMed: 18006185]
24. 1000_Genomes_Project_Consortium. Nature 526, 68–74 (2015). [PubMed: 26432245]
25. Moxon R, Bayliss C, Hood D. Annu Rev Genet 40, 307–333 (2006). [PubMed: 17094739]
26. Stoltzfus A, Yampolsky LY. J Hered 100, 637–647 (2009). [PubMed: 19625453]
27. Du X et al. Nucleic Acids Res 42, 12367–12379 (2014). [PubMed: 25336616]
28. Galen SC et al. Proc Natl Acad Sci USA 112, 13958–13963 (2015). [PubMed: 26460028]
29. Bacolla A, Tainer JA, Vasquez KM, Cooper DN. Nucleic Acids Res 44, 5673–5688 (2016). [PubMed: 27084947]
30. Hargreaves AD et al. Proc Natl Acad Sci USA 114, 7677–7682 (2017). [PubMed: 28674003]
31. Lowe CB et al. Genome Res 28, 256–265 (2018). [PubMed: 29229672]

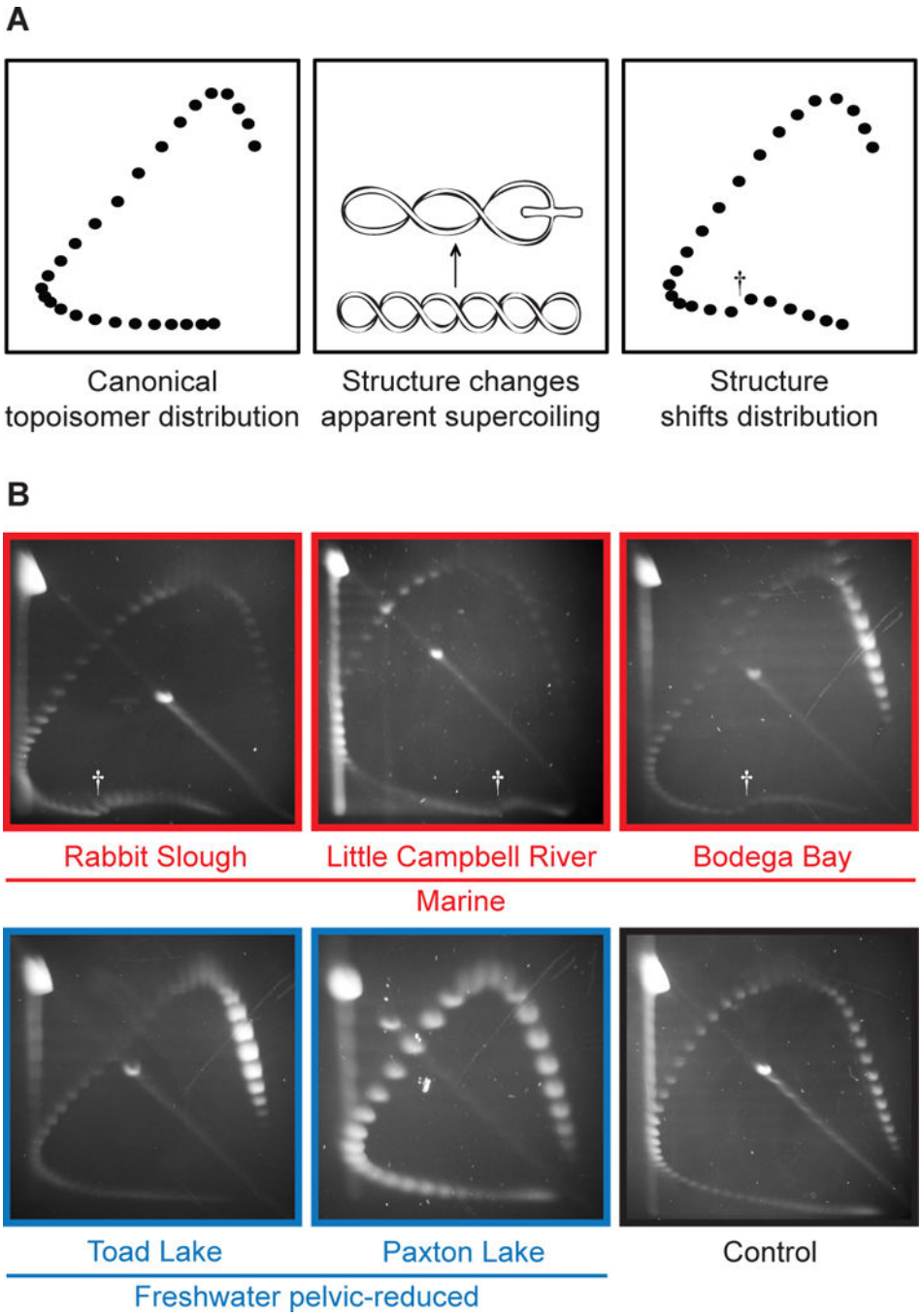

Figure 1. Marine but not freshwater *Pel* alleles form alternative structures in vitro.

(A) Two-dimensional electrophoresis of circular DNA topoisomers. A distribution of plasmid topoisomers is separated on an agarose gel; each topological class forms one spot. Canonical B-DNA forms a smooth distribution. Alternative structures cause mobility shifts. Distribution shifts at the linking number inducing alternative structure. Dagger, mobility shift. **(B)** *Pel* from marine and freshwater pelvic-reduced populations. Control, *AtpA1*.

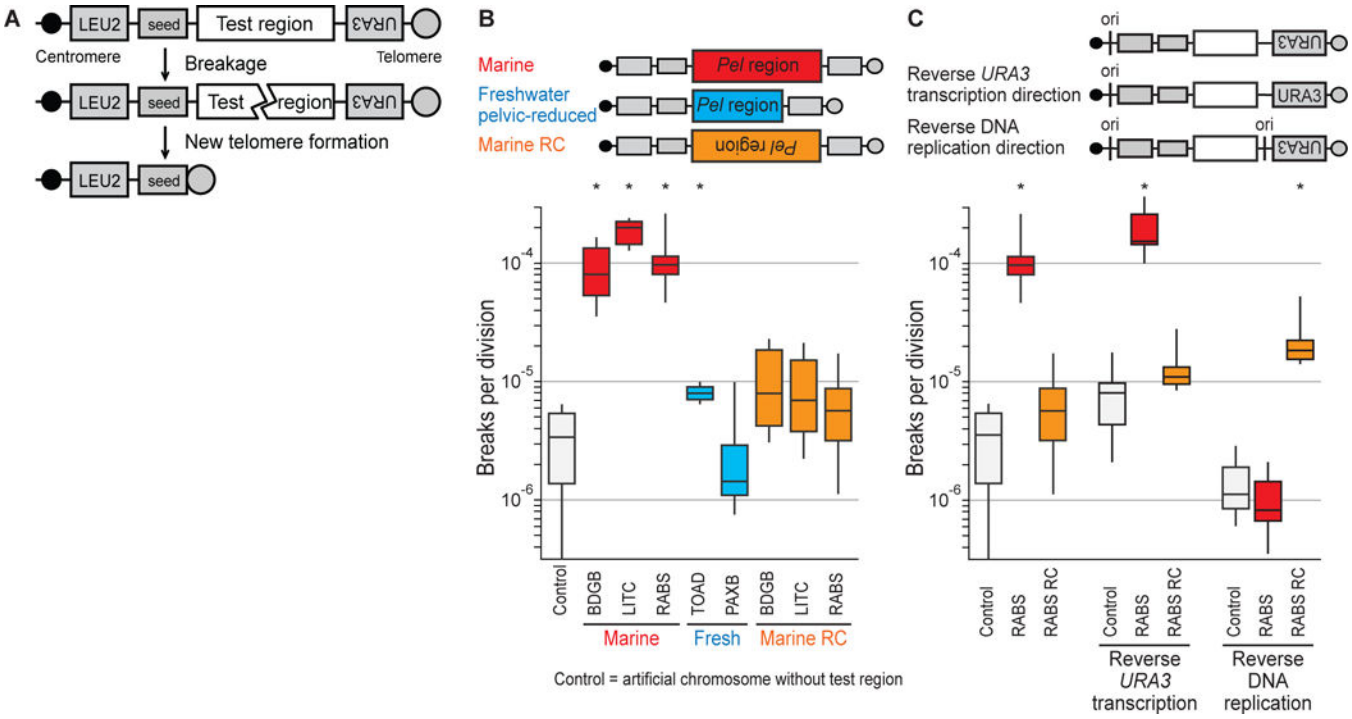

Figure 2. Marine but not freshwater *Pel* alleles break at high rates in yeast, in an orientation-dependent fashion.

(A) Test DNA is inserted in a yeast artificial chromosome between two selectable markers (*LEU2* and *URA3*) and downstream of a telomere seed site. Breakage results in loss of *URA3.* **(B)** Box-and-whisker plot of *Pel* breakage rates. Whisker ends indicate maximum and minimum of 6 fluctuation assays (10 cultures each). RC, reverse complement. *p < 0.01 (table S5). Population names, table S6. **(C)** Reversing replication direction through the test region, but not *URA3* transcription direction, reverses orientation of fragility. ori, DNA replication origin.

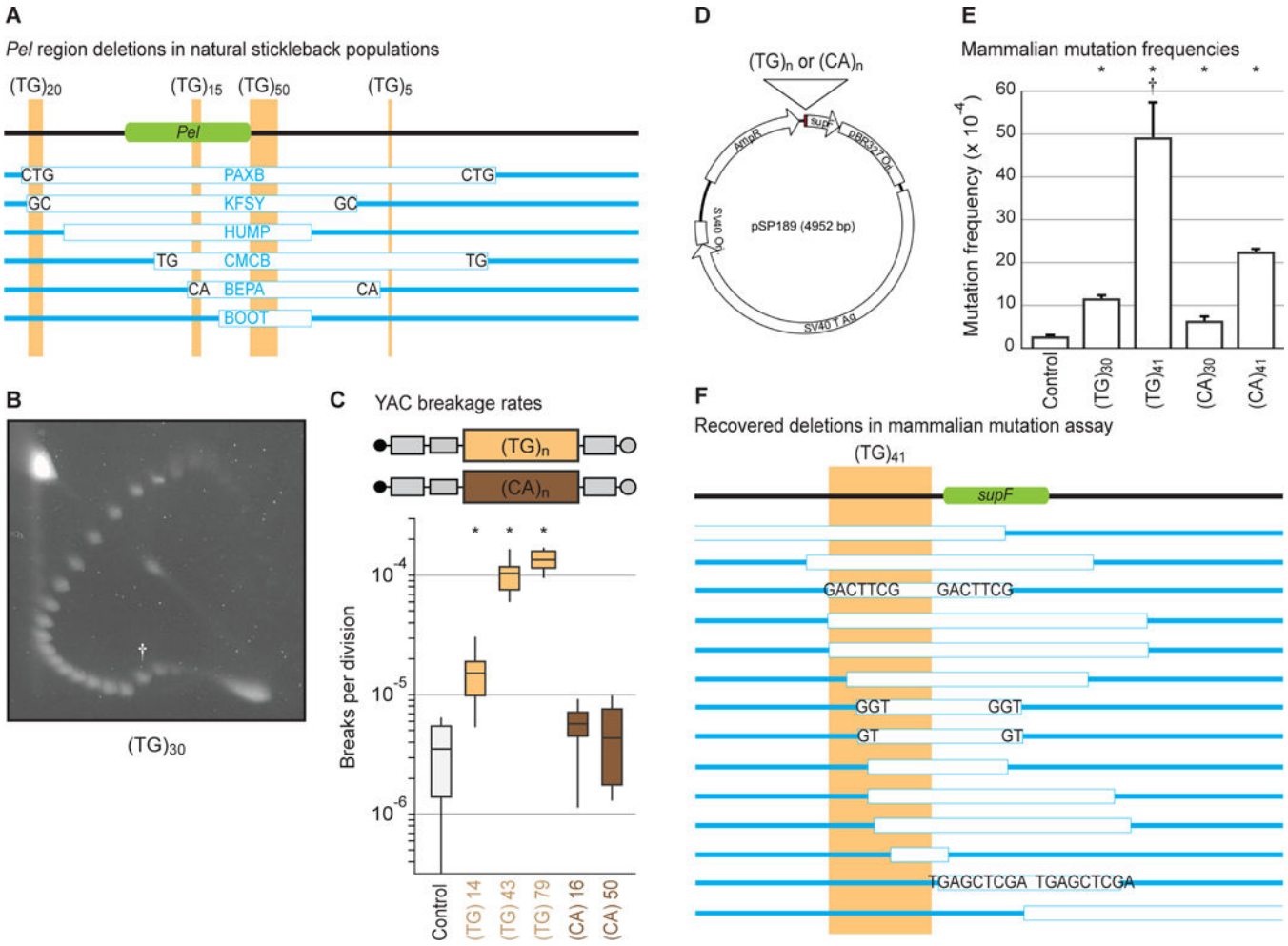

Figure 3. (TG)-dinucleotide repeats recapitulate structure formation, high breakage rate, orientation-dependence, and deletion spectrum.

(A) To-scale maps of *Pel* in different freshwater pelvic-reduced populations (table S6). Green, *Pel* sequence driving pelvis expression (6). Light-brown, TG-repeats. White boxes, DNA deletions in indicated populations. Blue, DNA remaining. Letters, microhomologies at deletion junctions. **(B)** Two-dimensional gel for $(TG)_{30}$. Dagger, mobility shift. **(C)** Yeast artificial chromosome breakage rates for TG- or CA-repeats of varying lengths. *$p < 0.01$ (table S5). **(D)** Reporter shuttle plasmid schematic. **(E)** Mammalian mutation frequencies. Error bars indicate SEM of 4–5 independent experiments. *$p < 0.05$ (Student's t-test). Dagger, deletions dominate mutation spectrum (fig. S2A). **(F)** To-scale map of $(TG)_{41}$-induced deletions in mammalian cells.

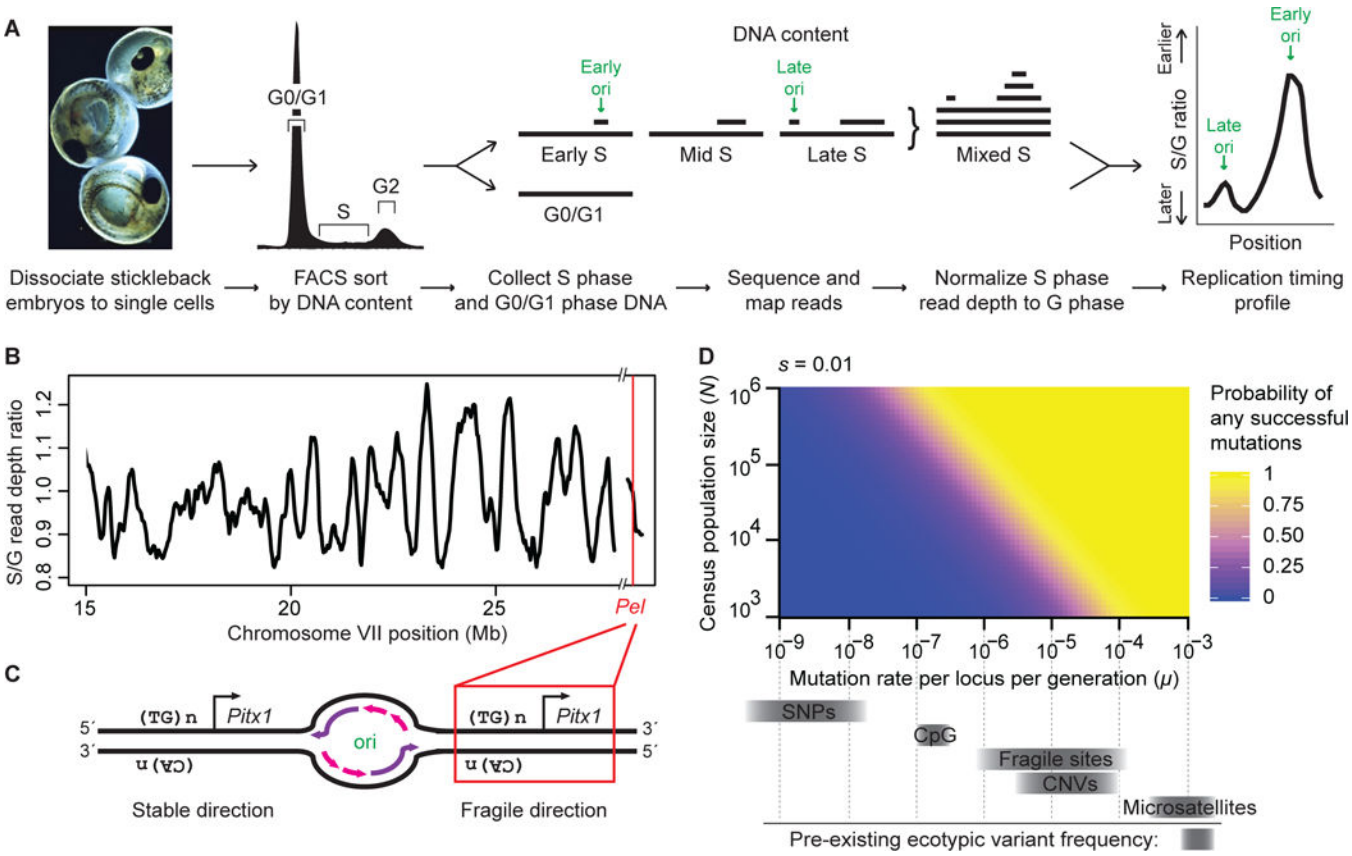

Figure 4. *Pel* is located in the breakage-prone orientation in sticklebacks, generating a fragile site likely to contribute to parallel evolution in natural populations.

(A) Workflow for profiling genome-wide replication timing. **(B)** Stickleback chromosome VII replication timing. Red line, *Pel* locus, which is subtelomeric. Hash marks, reference genome assembly gap. **(C)** Diagrams of stable and fragile replication orientations. ori, origin bubble. Purple, newly synthesized leading strand. Pink, newly synthesized lagging strand. **(D)** Probability of at least one de novo mutation arising at a particular locus in 10,000 generations and eventually fixing, as a function of typical stickleback population sizes (N) and mutation rates (μ, grey bars) for single nucleotides (SNPs), copy number variants (CNVs), and fragile sites. De novo point mutations are unlikely to occur and fix in small vertebrate populations, even when conferring a selective advantage (s=0.01, modeled here). In contrast, mutations occurring at fragile sites are likely to arise and contribute to repeated evolution when conferring a selective advantage. For additional parameters, including neutrality (s=0), see fig. S6 and fig. S7.