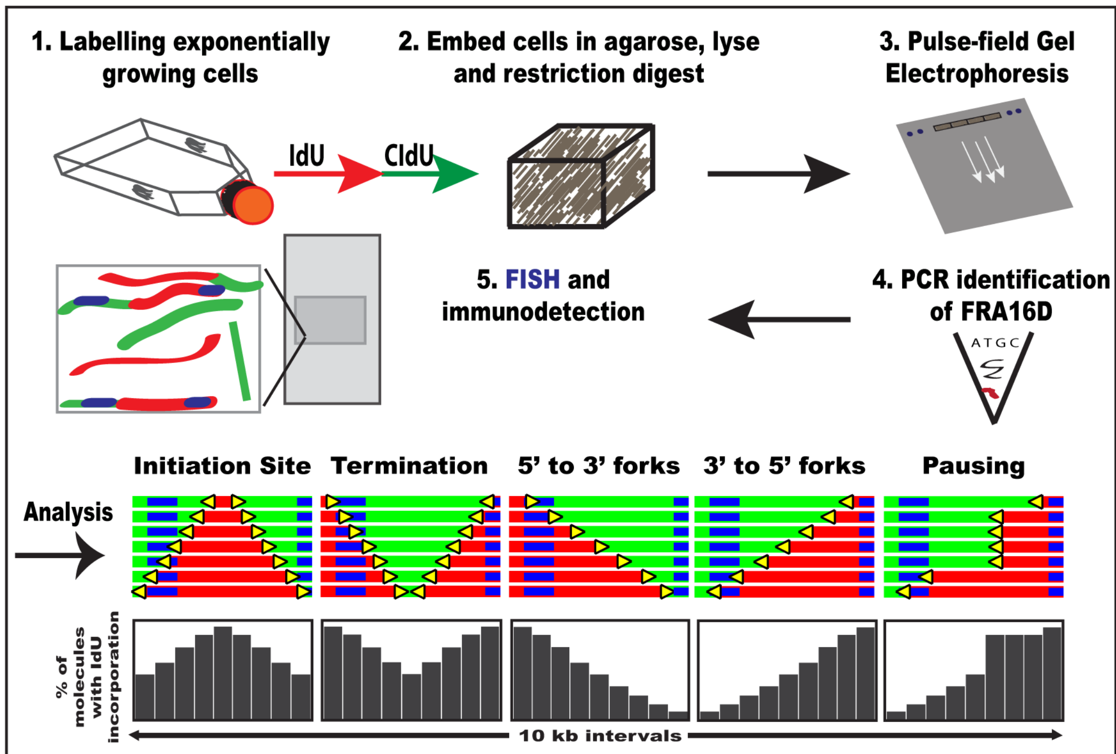
FIGURE S1

A



B

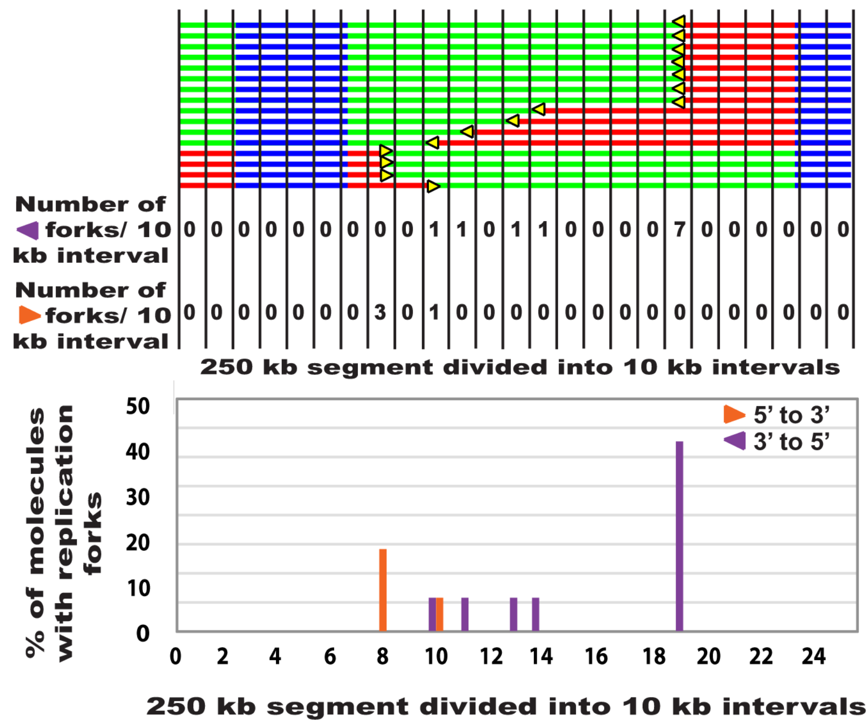
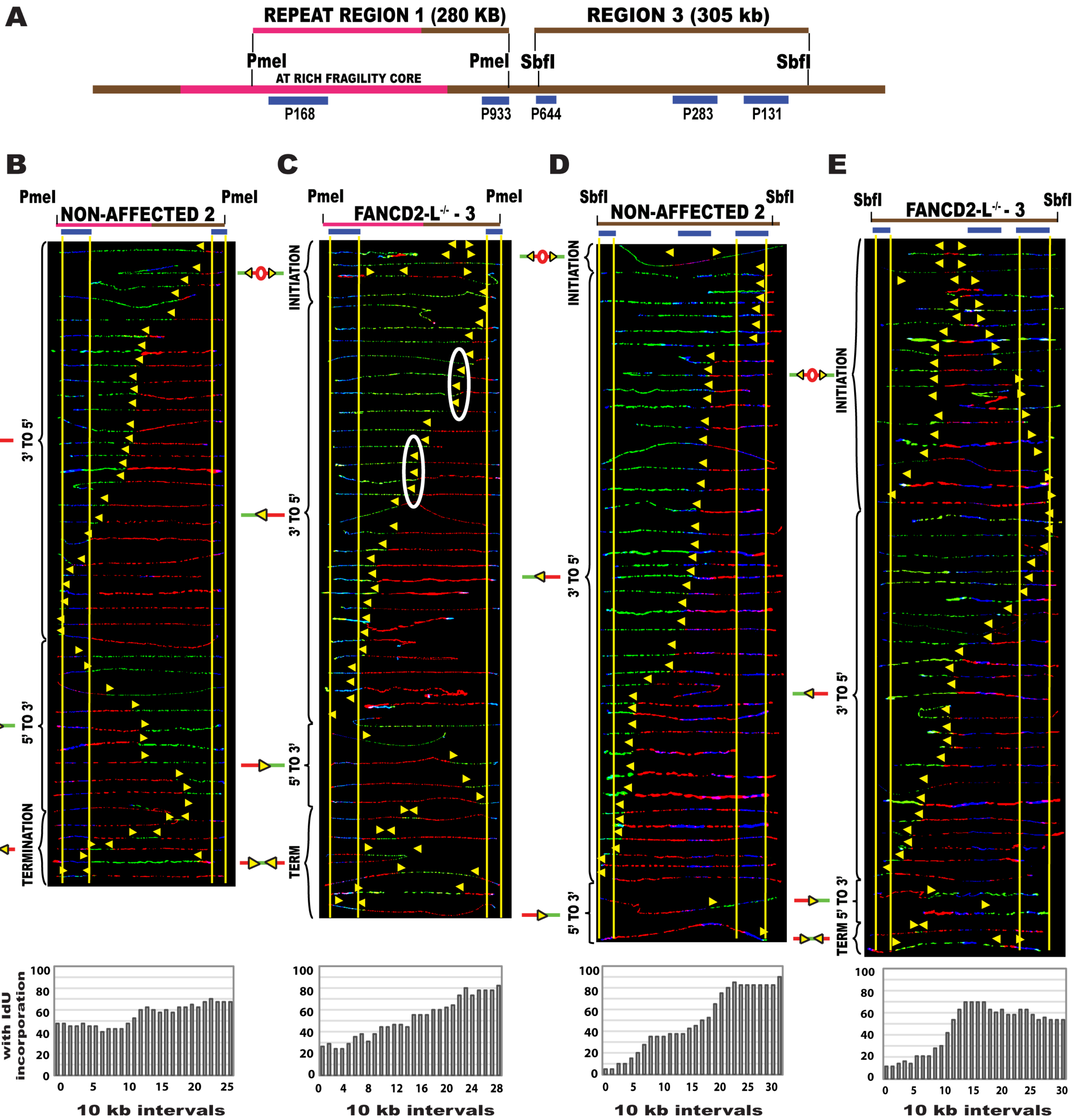


FIGURE S2



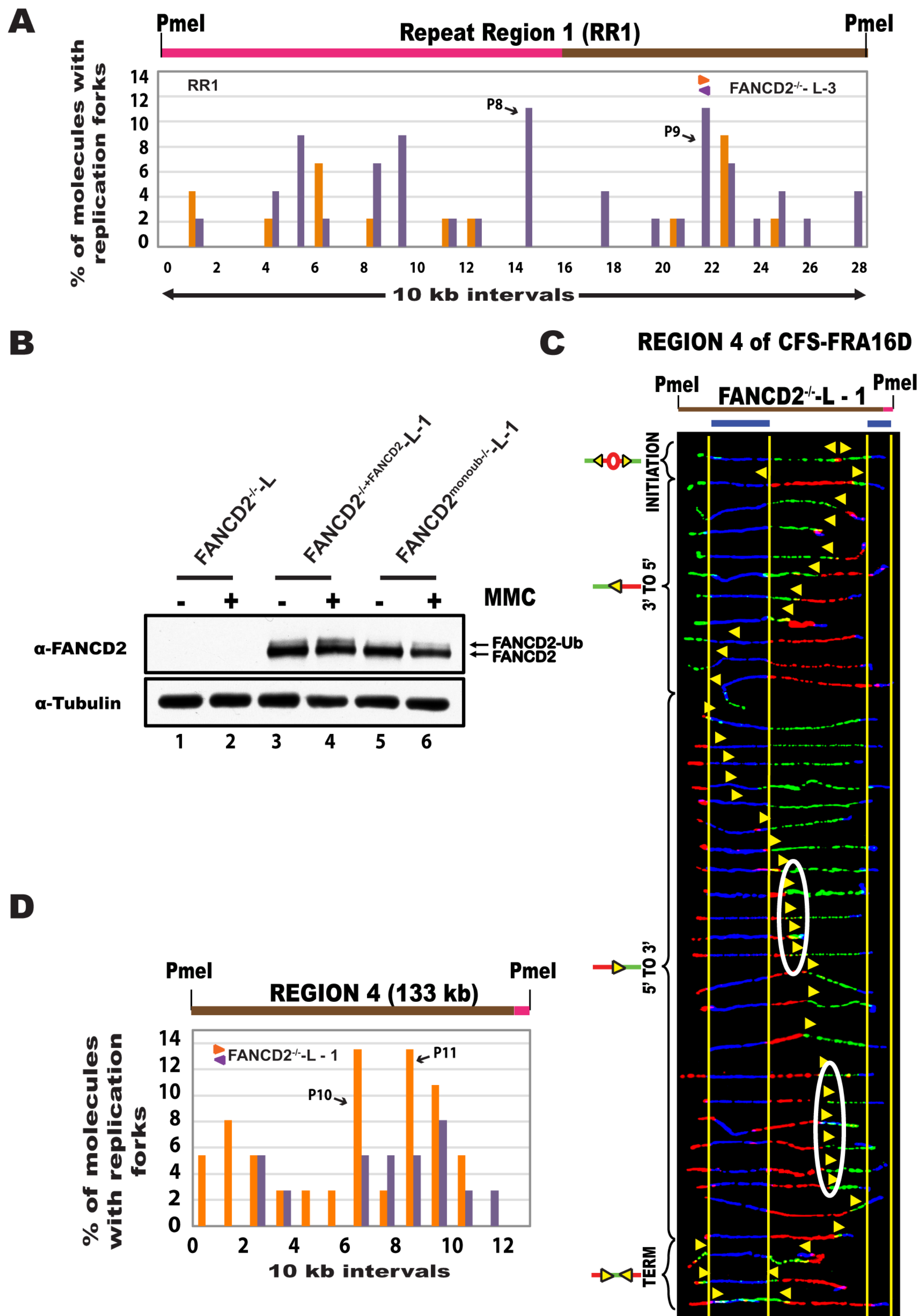
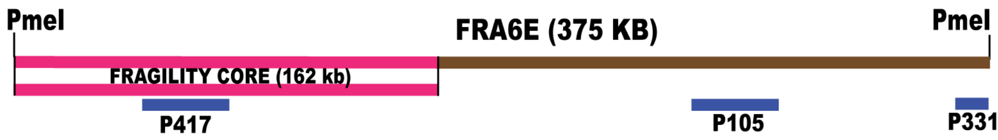


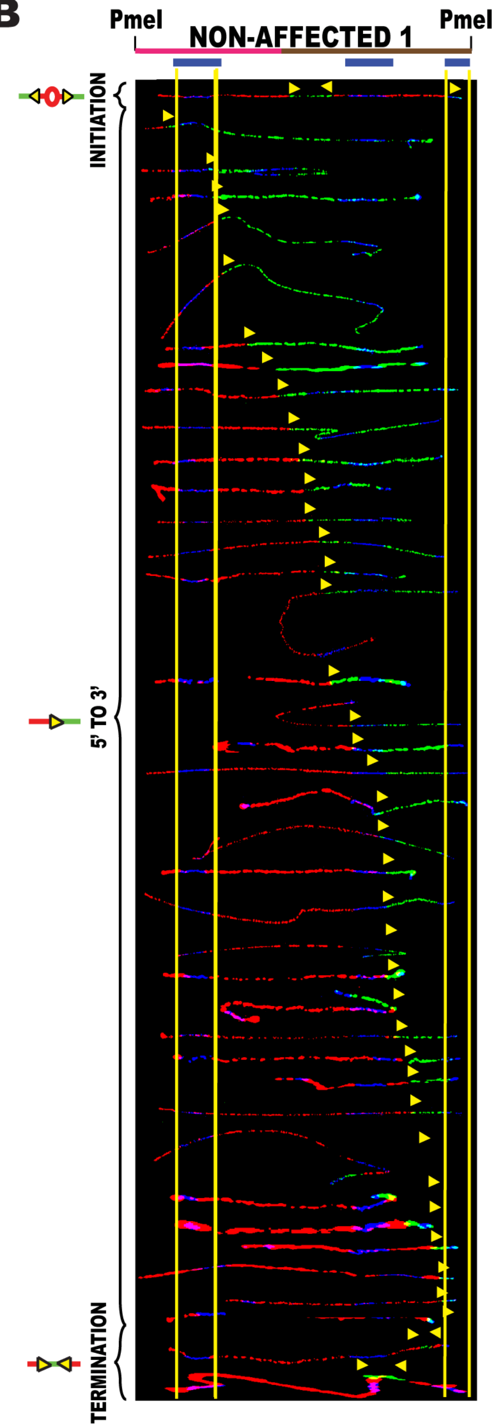
FIGURE S3
A
PmeI
Repeat Region 1 (RR1)
PmeI
RR1
FANCD2-/- L-3
P8
P9
% of molecules with replication forks
10 kb intervals
B
FANCD2-/--L
FANCD2-/-+FANCD2-L-1
FANCD2monoub-/--L-1
MMC
α-FANCD2
α-Tubulin
FANCD2-Ub
FANCD2
1 2 3 4 5 6
C
REGION 4 of CFS-FRA16D
PmeI
FANCD2-/--L - 1
PmeI
INITIATION
3' TO 5'
5' TO 3'
TERM
D
PmeI
REGION 4 (133 kb)
PmeI
FANCD2-/--L - 1
P10
P11
% of molecules with replication forks
10 kb intervals

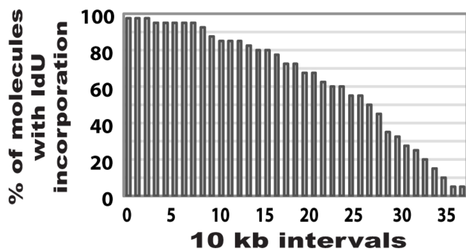
FIGURE S4

A

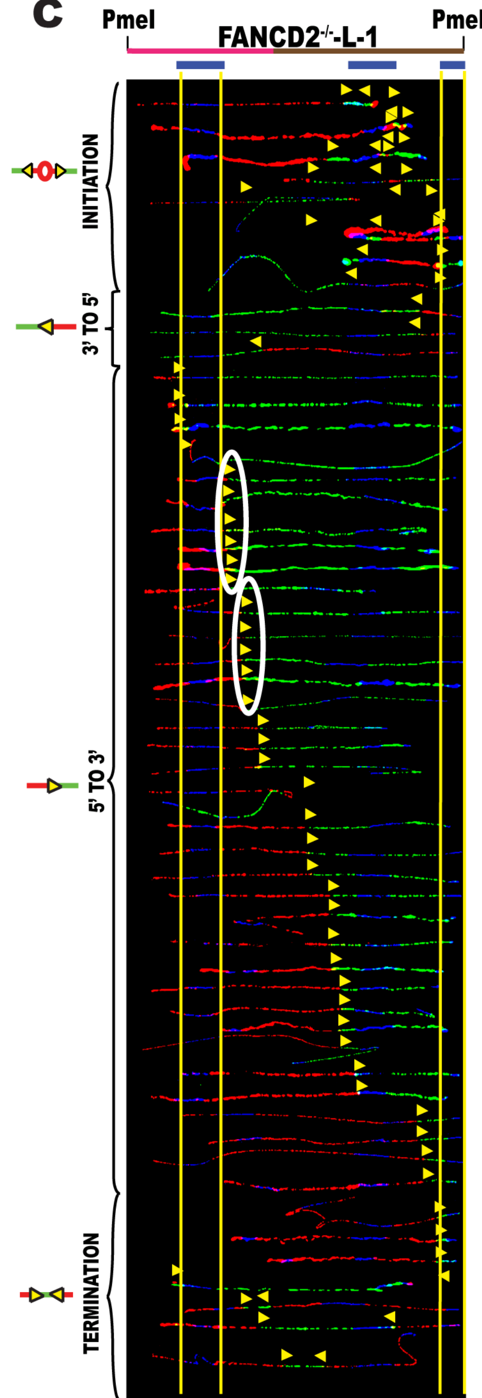


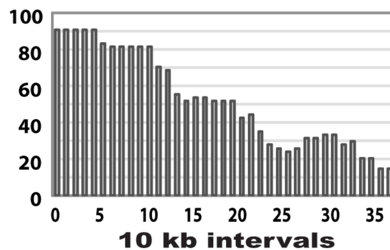
B



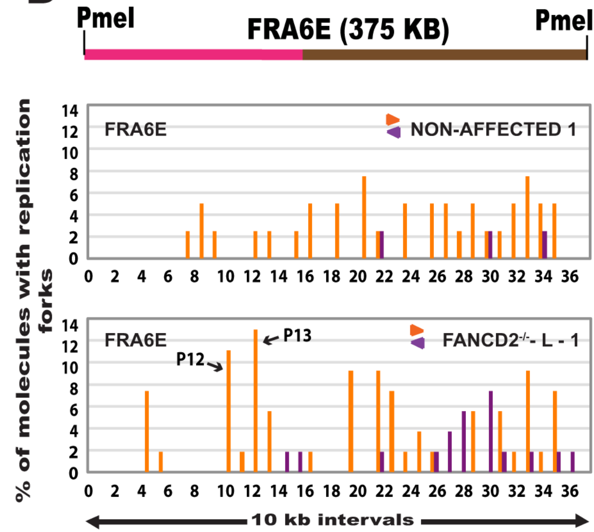


C

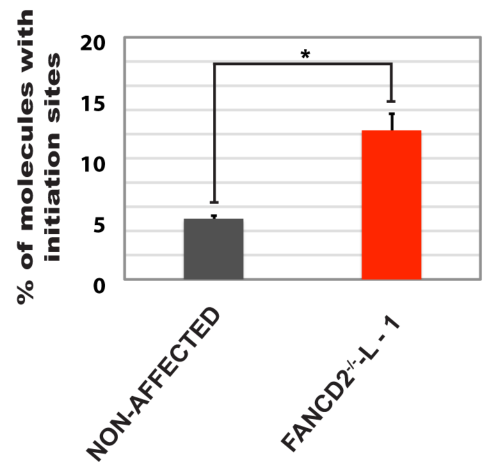




D



E



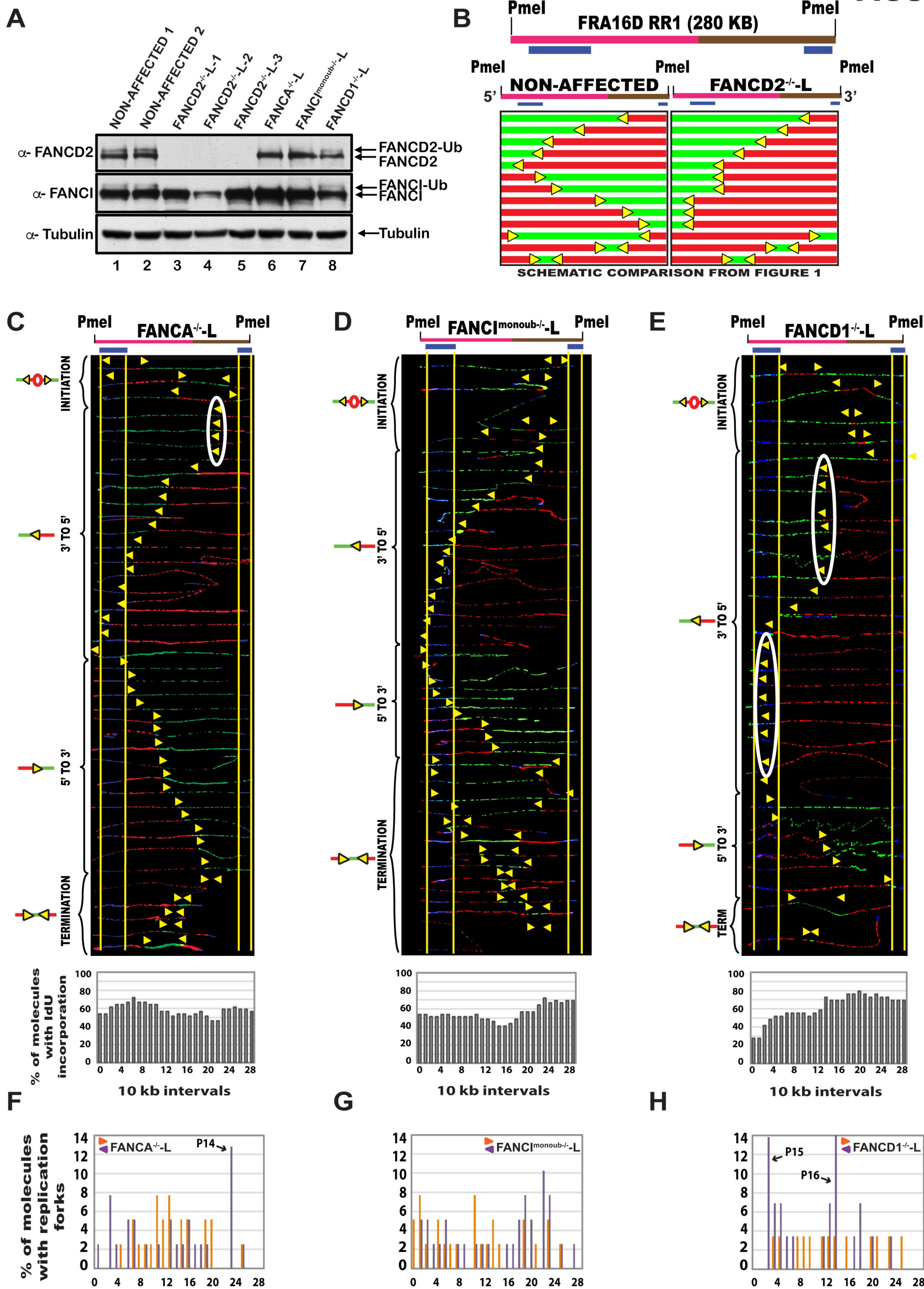


FIGURE S5
A
NON-AFFECTED 1
NON-AFFECTED 2
FANCD2-/--L-1
FANCD2-/--L-2
FANCD2-/--L-3
FANCA-/--L
FANCImonoub-/--L
FANCD1-/--L
α- FANCD2
α- FANCI
α- Tubulin
FANCD2-Ub
FANCD2
FANCI-Ub
FANCI
Tubulin
1 2 3 4 5 6 7 8
B
PmeI
FRA16D RR1 (280 KB)
NON-AFFECTED
FANCD2-/--L
5'
3'
SCHEMATIC COMPARISON FROM FIGURE 1
C
FANCA-/--L
INITIATION
3' TO 5'
5' TO 3'
TERMINATION
D
FANCImonoub-/--L
E
FANCD1-/--L
TERM
% of molecules with IdU incorporation
10 kb intervals
F
G
H
% of molecules with replication forks
P14
P15
P16

FIGURE S6

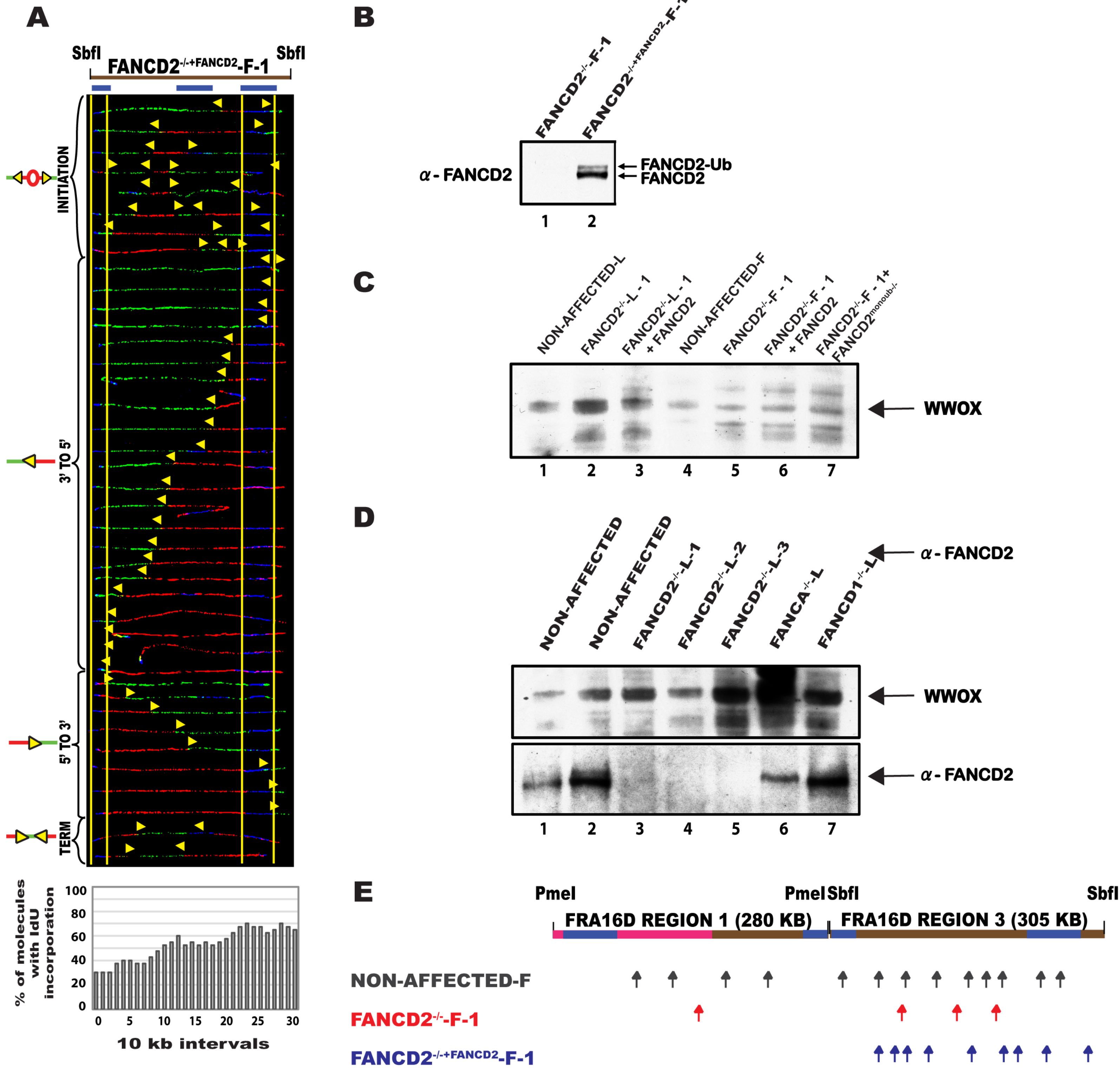
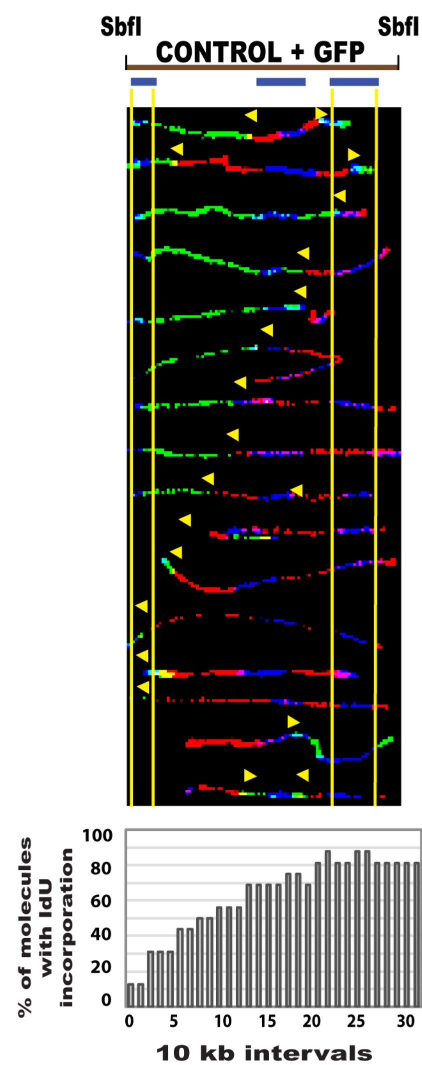
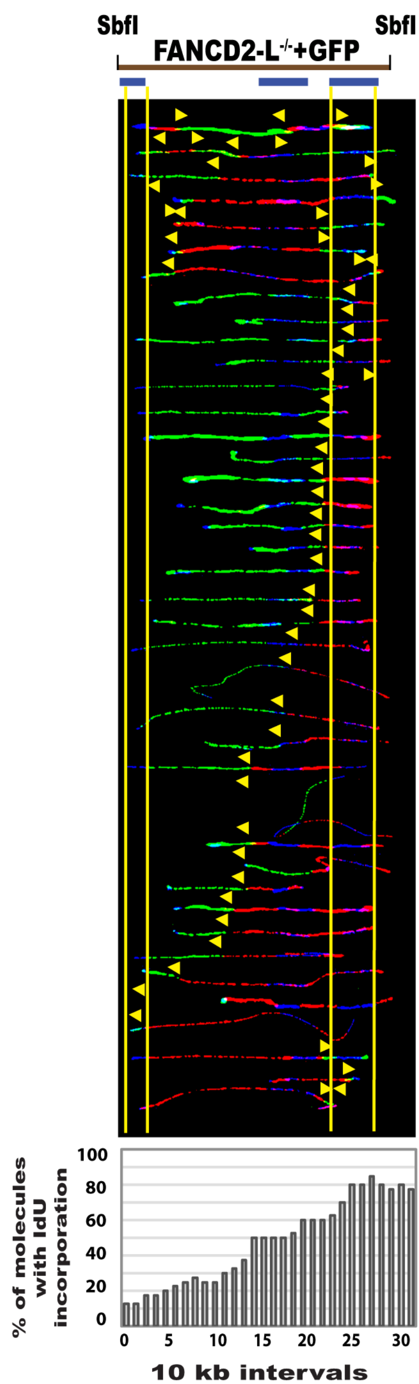


FIGURE S7





SUPPLEMENTARY DATA

Figure S1: Schematic of single molecule analysis of replicated DNA (SMARD). Related to Figures 1-6.

(A) Schematic representation of the various stages of SMARD. Cells are pulsed with nucleoside analogs (IdU-green; CIdU-red) and embedded in agarose plugs. The cells are first lysed; proteins are digested by proteinase K and then subjected to restriction digestion. The restriction digested DNA is resolved by pulse field gel electrophoresis. The slice containing the FRA16D locus is identified by PCR analysis. The agarose from the identified slice is melted and the DNA is stretched onto silanized glass slides. Biotinylated FISH probes are used for identification of the fragment and immunostaining is utilized to visualize the IdU tract in red, the CIdU tract in green and the FISH probes in blue. The resulting molecules are arranged to yield recognizable replication patterns (from the left): initiating molecules, terminating molecules, replication forks progressing in the 3' to 5' and 5' to 3' direction which are easily interpreted by the IdU incorporation histograms.

(B) Quantification of replication pausing: The SMARD image that contains the aligned DNA molecules is divided into 10 kb segments (black vertical lines), and the number of replication forks present in each 10 kb interval in either the 5' to 3' or the 3' to 5' direction are counted and these numbers are represented as the % of molecules with replication forks (y-axis) within the respective 10 kb interval of the segment (x-axis). A high percentage of molecules with replication forks in a particular 10 kb interval is indicative of fork pausing in that interval.

Figure S2: The replication program is altered at the endogenous CFS-FRA16D locus in the absence of FANCD2 protein. Results from additional non-affected and FANCD2-deficient lymphoblast lines - related to Figures 1 and 3

(A) Locus map of the RR1-*PmeI* and R3-*SbfI* segments, of CFS-FRA16D. The FISH probes that identify the segment are labeled in blue.

(B-C) Top; Locus map of *PmeI* digested RR1 segment. Middle; Aligned photomicrograph images of labeled DNA molecules from (B) non-affected 2 (GM03798) and (C) FANCD2$^{-/-}$-L-3 (2717) patient derived lymphoblasts. The yellow arrows indicate the sites along the molecules where the IdU transitioned to CldU. The molecules are arranged in the following order: molecules with initiation events, molecules with 3' to 5' progressing forks, molecules with 5' to 3' progressing forks and molecules with termination events. White ovals indicate regions of replication fork pausing and correspond to the pause peaks listed in Table S2. Bottom; The percentage of molecules incorporating IdU (red) is calculated from the replication program (middle) and are represented as a histogram.

(D-E) Top; Locus map of *SbfI* digested R3 segment. Middle; Aligned photomicrograph images of labeled DNA molecules from (D) Non-affected 2 (GM03798) and (E) FANCD2$^{-/-}$-L-3 (2717) patient derived lymphoblasts.

Figure S3: Replication forks moving into the AT rich fragility core of FRA16D stall in the absence of FANCD2. Related to Figure 2.

(A) Top; Locus map of the RR1 region. Bottom; The percentage of molecules at each 10 kb interval of RR1 in the FANCD2$^{-/-}$-L-3 (2717) patient derived lymphoblast. The replication forks moving in the 3' to 5' direction and the forks moving in the 5' to 3' direction are denoted by < and > colors respectively. A high percentage of molecules with replication forks in a particular 10 kb interval is indicative of fork pausing in that interval. Black arrows denote the most prominent pause peaks and correspond to the white ovals in the SMARD profile. Refer Table S2 for the coordinates of the 10 kb region corresponding to the pause peaks.

(B) Immunoblot analysis to detect the relative levels of the FANCD2 protein expression in the presence of absence of Mitomycin C (MMC) treatment in FANCD2$^{-/-}$-L-1 (PD20) (Lane1-2), FANCD2$^{-/-+FANCD2}$-L-1 (Lane3-4) and FANCD2$^{monoub-/-}$-L-1 (Lane5-6) lymphoblasts. Proteins from whole cell extract were separated and immunoblotted with anti-FANCD2 FI-17 and anti-tubulin (loading control) antibodies.

(C) Top; Locus map of the 133 kb *PmeI* digested R4 segment. Middle; Aligned photomicrograph images of labeled DNA molecules from FANCD2$^{-/-}$-L-1 (PD20) patient derived lymphoblast. The molecules are arranged as in Supplementary Fig. 1.

(D) The percentage of molecules at each 10 kb interval of R4 in the FANCD2$^{-/-}$-L-1 (PD20) patient derived lymphoblast. The replication forks moving in the 3' to 5' direction and the forks moving in the 5' to 3' direction are denoted by < and > colors respectively.

Figure S4: In the absence of FANCD2, replication pausing at the fragility core of CFS-FRA6E is accompanied by the activation of dormant origins. Analysis of the effect of FANCD2 deficiency at a second CFS locus, related to Figures 1-3.

(A): Locus map of a 375 kb region in the CFS-FRA6E obtained by *PmeI* digestion. The region includes the fragility core of CFS-FRA6E (pink line – 162 kb). The FISH probes that identify the segment are labeled in blue. The coordinates of FRA6E are summarized in Table S1, providing additional information about fosmids and primers used to identify the region.

(B-C) Top; Locus map of the *PmeI* digested FRA6E segment. Middle; Aligned photomicrograph images of labeled DNA molecules from non-affected 1 (GM02184), and FANCD2$^{-/-}$-L-1 (PD20) patient derived lymphoblast. The yellow arrows indicate the sites along the molecules where the IdU transitioned to CldU. White ovals indicate regions of replication fork pausing and correspond to the pause peaks listed in Table S2. The molecules are arranged as in Supplementary Fig. 1. Bottom; The percentage of molecules incorporating IdU (red) is calculated from the replication program (middle) and are represented as a histogram.

(D) The percentage of molecules at each 10 kb interval of Top; locus map of the 375 kb region of FRA6E in Middle; non-affected 1-L (GM02184), Bottom; FANCD2$^{-/-}$-L-1 (PD20) patient derived lymphoblast. The replication forks moving in the 3' to 5' direction and the forks moving in the 5' to 3' direction are denoted by < and > colors respectively. A high percentage of molecules with replication forks in a particular 10 kb interval is indicative of fork pausing in that interval. Black arrows denote the most prominent pause peaks

and correspond to the white ovals in the SMARD profile. Refer Table S2 for the coordinates of the 10 kb region corresponding to the pause peaks.

(E) Percentage of molecules with initiation sites in non-affected I (GM02184 - grey bar), FANCD2$^{-/-}$-L-1 (PD20) patient derived lymphoblast (red bars). Error bars represent mean ± s.d. from two independent experiments (*P<0.05).

Figure S5: The absence of other FA proteins leads to replication pausing at the endogenous fragility core of CFS-FRA16D. Related to Figures 1 and 6.

(A) Immunoblot analysis to detect the relative levels of the FANCD2 and FANCI protein expression, to demonstrate the absence of the FANCD2 protein in the three FANCD2$^{-/-}$-L patient derived cell lines.

(B) Schematic representation of results from Fig. 2 (*PmeI* segment), for comparison of replication programs at the RR1 segment of the endogenous CFS-FRA16D locus, beginning with Top; the locus map, Bottom left; the replication profile of the non-affected-L, Bottom right; replication profile of the FANCD2$^{-/-}$ lymphoblast.

(C-E) Top; Locus maps of *PmeI* digested RR1 segment. Middle; Aligned photomicrograph images of labeled DNA molecules from (C) FANCA$^{-/-}$-L, (D) FANCI$^{monoub-/-}$-L, and (E) FANCD1$^{-/-}$-L patient derived lymphoblast. The yellow arrows indicate the sites along the molecules where the IdU transitioned to CldU. White ovals indicate regions of replication fork pausing and correspond to the pause peaks listed in Table S2. The molecules are arranged as in Fig. 2. Bottom; The percentage of molecules incorporating IdU (red) is calculated from the replication program (middle) and is represented as a histogram.

(F-H) The percentage of molecules with replication forks at each 10 kb interval of RR1 in the FANCA$^{-/-}$-L, FANCI$^{monoub-/-}$-L and FANCD1$^{-/-}$-L lymphoblasts. The replication forks moving in the 3' to 5' direction and the forks moving in the 5' to 3' direction are denoted by < and > colors respectively. A high percentage of molecules with replication forks in a particular 10 kb interval is indicative of fork pausing in that interval. Black arrows

denote the most prominent pause peaks and correspond to the white ovals in the SMARD profile. Refer Table S2 for the coordinates of the 10 kb region corresponding to the pause peaks.

Figure S6: Reconstitution of the FANCD2 protein expression in PD20 fibroblasts suppresses initiation defect. Related to Figures 4 and 7.

(A) Top; Locus map of *SbfI* digested R3 segment. Middle; Aligned photomicrograph images of labeled DNA molecules from FANCD2$^{-/-+FANCD2}$-F-1 (PD20) patient fibroblasts.

(B) Immunoblot analysis to detect the relative levels of the FANCD2 protein expression in the FANCD2$^{-/-}$-F-1 (Lane1) and FANCD2$^{-/-+FANCD2}$-F-1 (Lane2) fibroblasts. Proteins from whole cell extracts were separated, immunoblotted and detected with anti-FANCD2 FI-17 antibody.

(C) Immunoblot analysis to compare the relative levels of the WWOX protein expression in lymphoblasts (Lanes 1-3) and fibroblasts (Lanes 4-7). Proteins from whole cell extracts were separated, immunoblotted and detected with anti-WWOX antibody.

(D) Immunoblot analysis to compare the relative levels of the WWOX protein expression in Non-affected (Lane1-2), FANCD2-deficient (Lane3-5), FANCA-deficient (Lane 6) and BRCA2/FANCD1-deficient (Lane7) lymphoblasts. Proteins from whole cell extracts were separated, immunoblotted and detected with anti-WWOX and anti-FANCD2 antibodies.

(E) Top; Locus map of CFS-FRA16D spanning the RR1 and R3 segments. Bottom; Observed sites of initiation in non-affected-F (grey arrows), FANCD2$^{-/-}$-F-1 (red arrow), and FANCD2$^{-/-+FANCD2}$ -F-1 (blue arrows).

Figure S7: Replication programs of non-affected and PD20 lymphoblasts transfected with GFP control vectors are similar to non-transfected lymphoblasts. Controls for Figure 5E-F.

(A-B) Top; Locus map of the *SbfI* digested Region 3 (R3) segment. Middle; Aligned photomicrograph images of labeled DNA molecules from (A) Non-affected I or control + GFP, (B) FANCD2$^{-/-}$-L-1 + GFP lymphoblasts. The yellow arrows indicate the sites along the molecules where the IdU transitioned to CldU. Bottom; The percentage of molecules incorporating IdU (red) is calculated from the replication program (middle) and are represented as a histogram.

Table S1: Summary of CFS segments analyzed by SMARD, genomic coordinates, corresponding PCR primers, and Fosmids (GRCh38/hg38) used for identification of the regions. Table S1, related to Figures 1-6.

CFS Name-Segment	Coordinates (HG38)		PCR Primer Sequence	Fosmids
	Start	End		
FRA16D-RR1-SbfI 120 kb	78353779	78472997	5'ACCCATAAACCACTGCGAGG 5'AGAAGCCTACGGAGACTATGA	WI2-3279A4 WI2-1842I12
FRA16D-RR2-PmeI 280 kb	78535654	78816150	5'GAGGCCTGGTGTATGCACTT 5'CTACAGACAGGCAGGCACAA	WI2-1680B6 WI2-933I24
FRA16D-R3-SbfI 305 kb	78815000	79120060	5'CCGATGCAACTGTCTGTCCT 5'TCCAACAACGGTCTCACCAG	WI2-644P6 WI2-2833E8 WI2-1310G20
FRA16D-R4-PmeI-133 kb	78259650	78393086	NONE	WI2-2262H15 WI2-1513I10
FRA6E-PmeI 375 kb	161386120	161761066	5'TTTGGGACTCTGGCAACACA 5'TGACTCAGTCCAACACCCAC	WI2-417M13 WI2-1053K10 WI2-3314B10

Table S2: Coordinates of each 10 kb interval corresponding to the pausing sites. Table S2, related to Figures 2, 4, S3-5.

CFS Name-Segment	Cell Line	Pausing Peak Name	Coordinates Of Pause Sites (10 kb)	
			Start	End
FRA16D-RR1-SbfI 120 kb	FANCD2$^{-/-}$-L-1	P1	78393518	78403453
	FANCD2$^{-/-}$-L-2	P4	78403453	78413388
	FANCD2$^{-/-}$-L-2	P5	78433258	78443193
FRA16D-RR2-PmeI 280 kb	FANCD2$^{-/-}$-L-1	P2	78605778	78615796
	FANCD2$^{-/-}$-L-1	P3	78695937	78705955
	FANCD2$^{-/-}$-L-2	P6	78665884	78675902
	FANCD2$^{-/-}$-F-1	P7	78685920	78695937
	FANCD2$^{-/-}$-L-3	P8	78675902	78685920
	FANCD2$^{-/-}$-L-3	P9	78746026	78756044
	FANCA$^{-/-}$-L	P14	78756044	78766061
	FANCD1$^{-/-}$-L	P15	78555689	78565707
	FANCD1$^{-/-}$-L	P16	78665884	78675902
FRA16D-R4-PmeI-133 kb	FANCD2$^{-/-}$-L-1	P10	78321236	78331500
	FANCD2$^{-/-}$-L-1	P11	78341764	78352029
FRA6E-PmeI 375 kb	FANCD2$^{-/-}$-L-1	P12	161487457	161497590
	FANCD2$^{-/-}$-L-1	P13	161507724	161517858

Table S3: Genomic coordinates of regions analyzed by DRIP and the corresponding qPCR primers. Table S3, related to Figure 5.

Region Name	Coordinates	Primer Sequence
WWOX - H1	78239233-78433198	5’ CCCCGTGAGGAAGGAAGAG
		5’ CCCTGGACCCAGTTCTCCAT
WWOX - H2	78432814-78433411	5’TTGTCCAGCCCATCATAAAGG
		5’GCATTTACCCAAACAGACTTAACCTAA
WWOX - H3	78498070-78498503	5’ TGCAGTGAGCCGAGATTGTG
		5’ TGAGATGGAGTCTTGCTCTGTTG

Supplemental Experimental Procedures

Antibodies and immunoblotting

For immunoblotting analysis, cell pellets were washed in PBS and lysed in 2% w/v SDS, 50 mM Tris-HCl, 10 mM EDTA, to prepare whole cell extracts. Proteins were resolved on NuPage 3-8% w/v Tris-Acetate or 4-12% w/v Bis-Tris gels (Invitrogen) and transferred to polyvinylidene difluoride (PVDF) membranes. The following antibodies were used: rabbit polyclonal antisera against FANCD2 (NB100-182; Novus Biologicals), FANCI (A301-254A, Bethyl), and mouse monoclonal sera against α-tubulin (MS-581-PO; Neomarkers), anti-WWOX antibody (N-19 Santa Cruz sc-20528), and anti-RNaseH1 antibody (15606-1-AP, Proteintech).

Transfection for RNaseH1 over expression

Exponentially growing cells were transfected with an empty plasmid (pEGFP-C1) or with an RNH1-overexpressing plasmid (pM27-GFP). Transfections were performed by electroporation using the Cell Line Nucleofector® Kit V (Lonza) following the manufactured instructions using the transfection program X-001 in Nucleofector (Amaxa). 24 hours after transfection, cells were pulsed for 4 hours with 30 μM IdU following by 4 hours with 30 μM CldU, washed with ice-cold PBS and embedded in agarose plugs and SMARD was performed as described above.

Fluorescent in situ hybridization (FISH) methods

Cells were maintained in Dulbeccos' Modified Eagle Medium (Gibco) supplemented with FBS and gentamycin. 24 hr prior to harvest, flasks were treated with 0.2 μM Aphidicolin. Four hours prior to harvest, flasks were treated with colcemid (0.05 μg/ml) to arrest cells at metaphase. Cells were trypsinized, pelleted, and re-suspended in a hypotonic solution (0.075 M KCl, 5% FBS) prior to being fixed with 3:1 methanol:acetic acid. Cells were affixed to slides and baked for 5 minutes at 95ºC before being incubated in 2xSSC (0.3M NaCl, 0.03M $Na_3C_6H_5O_7*2H_2O$, pH7) at 37ºC for 30 minutes. Slides were run through alcohols (70%, 80% and 90% for 2 minutes each) and dried before probing with the Abbott IGH/MAF probe set. Hybridization was on a HYBrite (Vysis) running a hybridization program of ramping up to 72ºC and holding for 2 minutes and then cooling and holding at 37ºC. Slides were incubated overnight at 37ºC in a humidified chamber and post-washed the following morning: 0.4XSSC w/ 0.3% NP40 heated to 72ºC for 2 minutes and 2XSSC w/ 0.15% NP-40 at room temperature for 30 seconds. Slides were mounted in Antifade with DAP (VECTASHIELD)I. Imaging was done using a Nikon E900 fluorescence microscope with an Applied Imaging camera, and analyzed using Cytovision Software. 50 metaphase cells each from treated and untreated cultures were scored for disruption of the MAF (red) signal. Fragility was confirmed on the DAPI-stained chromosome.