SUPPLEMENTARY FIGURE LEGENDS

Supplemental Figure 1. Sanger sequence analysis of 27 FA families confirms mutations. The family ID, gene mutated and the Sanger sequence traces showing each mutation are presented. In each case, data is presented for the "proband", and when DNA is available, the father, mother, affected or normal siblings. The sequences are shown on top, and the base affected by the mutation is indicated (*). The primers used for finding FA gene mutations from genomic DNA were previously published in a large number of papers.

Supplemental Figure 2. Sequence coverage of all of FA genes by WES capture and sequencing for sample FA19. UCSC browser track for an FA gene are shown above, while sequences aligned to the reference sequence after WES capture and sequencing are shown below. The gaps in the rectangle indicate sequence not covered. Since this is only an exon capture strategy, sequences shown are for the exon and the immediate flanking region. All sequences are from sample FA19.

Supplemental Figure 3. Sequence coverage of all of FA genes by TruSeq capture and sequencing for sample FA26. UCSC browser track for an FA gene are shown above, while sequences aligned to the reference sequence after TruSeq capture and sequencing are shown below. The gaps in the rectangle indicate sequence not covered. All sequences are from sample FA26.

Supplemental Figure 4: Variant *FANCL* transcripts in FA26 are caused by the c.375-2033C>G mutation. The four different transcript variants (A-D) presumably caused by the

c.3745-2033C>G mutation are shown in detail *via* RT-PCR using primers located in exon 2 and exon 8. The Sanger sequence traces are shown below each variant, and the splicing patterns based on these sequences are shown above each variant.

Supplemental Figure 5. Variant *FANCI* transcripts in FA14 caused by the c.1583+142C>T mutation. A) RT-PCR analysis of for the *FANCI* mutation region in FA14 fibroblast cell line total RNA, using the primers for exons 14 and 18. The aberrant products are indicated by filled circles (•, ••).

B) The UCSC browser track for the *FANCI* gene is shown on top with the expansion of exons 14-18, analyzed by RT-PCR. The mutation (*), the WT sequence and the mutant sequence are indicated. The abnormal transcript that extends exon 16 by including the first 140 bases of intron 16 is indicated, along with the Sanger sequence trace.

Supplemental Table 1: **Targeted Gene Regions for MIP Capture and their Probe Design Coverage**

Supplemental Table 2: Ancestry and Other Information of FA Families

Supplemental Table 3: Coverage of the Targeted Region After MIP Capture, Enrichment and Sequencing

Supplemental Table 4: **Targeted Gene Regions for CGH Array Design**

Supplemental Table 1: **Targeted Gene Regions for MIP Capture and their Probe Design Coverage**

Target Gene	Alternate Name	Target Region (hg18)*	Design Coverage (%)†‡
BLM		chr15:89061583-89159690	94.21
BRCA1		chr17:38449839-38531026	86.02
CHEK1		chr11:125001883-125028952	95.42
CHEK2		chr22:27413731-27467822	86.65
FANCA		chr16:88330460-88411566	93.58
FANCB		chrX:14770450-14802105	99.06
FANCC		chr9:96900157-97120812	96.66
FANCD1	BRCA2	chr13:31786617-31872809	98.18
FANCD2		chr3:10042113-10117344	93.86
FANCE		chr6:35527116-35543859	93.94
FANCF		chr11:22599655-22604963	100
FANCG		chr9:35062835-35071013	96.53
FANCI		chr15:87587198-87662366	97.52
FANCJ	BRIP1	chr17:57113767-57296537	98.22
FANCL		chr2:58238882-58323019	96.53
FANCM		chr14:44673886-44740843	94.93
FANCN	PALB2	chr16:23520984-23561179	93.21
FOXC2		chr16:85158358-85160036	58.64
FOXF1		chr16:85101645-85105577	64.75
FOXL1		chr16:85169616-85172805	94.7
RAD51		chr15:38774651-38811648	90.66
RAD51AP1		chr12:4518317-4539475	97.45
USP1		chr1:62674563-62690063	96.22
WDR48		chr3:39068511-39112885	95.37

*The Target Region includes ~1kb beyond either side of the gene, a total of 1.36 Mb

†Percentage of the targetted region covered by the designed capture probes

‡ Probes could not be designed for exon 1 of FANCA and FANCE, as well as exons 18 and 19 of FANCD2

Supplemental Table 2. Clinical and Other Prior Information of FA Families

Sample ID	Complementation	Gender	Ancestry	DNA Source
		MIP Targeted Capture Method		
FA1	None	Male	N. European/Hispanic	PB
FA2	None**	Male	N. European	PB
FA3	None**	Male	European	Fib
FA4	None**	Male	Ashkenazi Jewish	PB
FA5	None**	Female	N. European	PB
FA6	None**	Female	Pakistani	LCL
FA7	None**	Male	Italian	LCL
FA8	None**	Male	N. European	PB
FA9	None	Male	N. European	PB
FA10	None	Female	Ashkenazi Jewish	Fib
FA11	FANCB	Male	N. European	Fib
FA12	FANCG	Female	N. European, Native American	FIb
FA13	FANCL	Female	N. European, Native American	Fib
FA14	nonACGEFL	Female	N. European	LCL
FA15	nonACD1D2EFGl	Male	N. European	Fib
FA16	nonACG	Female	N. European, Hispanic	Fib
FA17	nonACG	Female	Indian	Fib
FA18*	nonACG	Female	N. European	LCL
FA19*	nonACFG	Male	N. European	LCL
		TruSeq Targeted Capture Method		
FA20	None	Male	N. European	PB
FA21	None**	Female	European	LCL
FA22	None	Male	European	PB
FA23	None	Male	European	PB
FA24	None	Male	European	PB
FA25	None	Male	European	PB
FA26	None**	Female	N. European	Fib
FA27	None	Male	European, Native American	LCL

* Also performed whole exome sequencing (WES)

**Excluded FANCA by Sanger sequencing

PB=peripheral blood. Fib and LCL represent fibroblast and lymphoblastoid cell lines respectively

Supplemental Table 3: **Coverage of the Targeted Region After MIP Capture, Enrichment and Sequencing**

Sample ID	Post-Seq Coverage, Total (%)*	Post-Seq coverage, Exons (%)*
FA1	88.72	95.36
FA2	87.53	95.63
FA3	85.31	94.93
FA4	88.32	95.26
FA5	88.36	95.56
FA6	88.19	95.17
FA7	87.82	95.53
FA8	88.44	95.54
FA9	85.5	95.16
FA10	84.4	93.81
FA11	85.8	95.08
FA12	84.8	94.68
FA13	86.4	95.09
FA14	74	91.33
FA15	85.3	93.99
FA16	87.4	94.12
FA17	82.1	89.68
FA18	87.2	94.61
FA19	77.1	91.56

*Represents the genotype calls with a >10 mpg score after aligning the sequences to the reference genome

Supplemental Table 4: **Targeted Gene Regions for CGH Array Design**

Target Gene	Alternate Name(s)	Target Region (hg18)*
AP1TD1	MHF1 (CENP-S)	chr1:10402746-10435459
BLM		chr15:89051583-89169690
BRCA1		chr17:38439840-38540994
CDKN2A		chr9:21947751-21994490
FAAP100	C17orf70	chr17:77107387-77139868
FAAP24	C19orf40	chr19:38144988-38169800
FANCA		chr16:88131460-88610566†
FANCB		chrX:14671450-14901105‡
FANCC		chr9:96701501-97319867†
FANCD1	BRCA2	chr13:31777617-31881809
FANCD2		chr3:9993451-10166418§
FANCE		chr6:35518116-35552859
FANCF		chr11:22590655-22613963
FANCG		chr9:35013835-35120013§
FANCI		chr15:87578198-87671366
FANCJ	BRIP1	chr17:57104767-57306537
FANCL		chr2:58229882-58333019
FANCM		chr14:44664886-44749843
FANCN	PALB2	chr16:23511984-23570179
FANCO	RAD51C	chr17:54114962-54176691
FANCP	SLX4 (BTBD12)	chr16:3561184-3611586
MRE11		chr11:93780115-93876688
NBN		chr8:91004740-91076075
RAD50		chr5:131910529-132017494
RAD51AP1		chr12:4508317-4549475
SMAD4		chr18:46800581-46875407
STRA13	MHF2 (CENP-X)	chr17:77559868-77584062

*****Target regions extend 10kb beyond either side of the gene, unless noted otherwise.

†Extends 200kb beyond either side of the gene

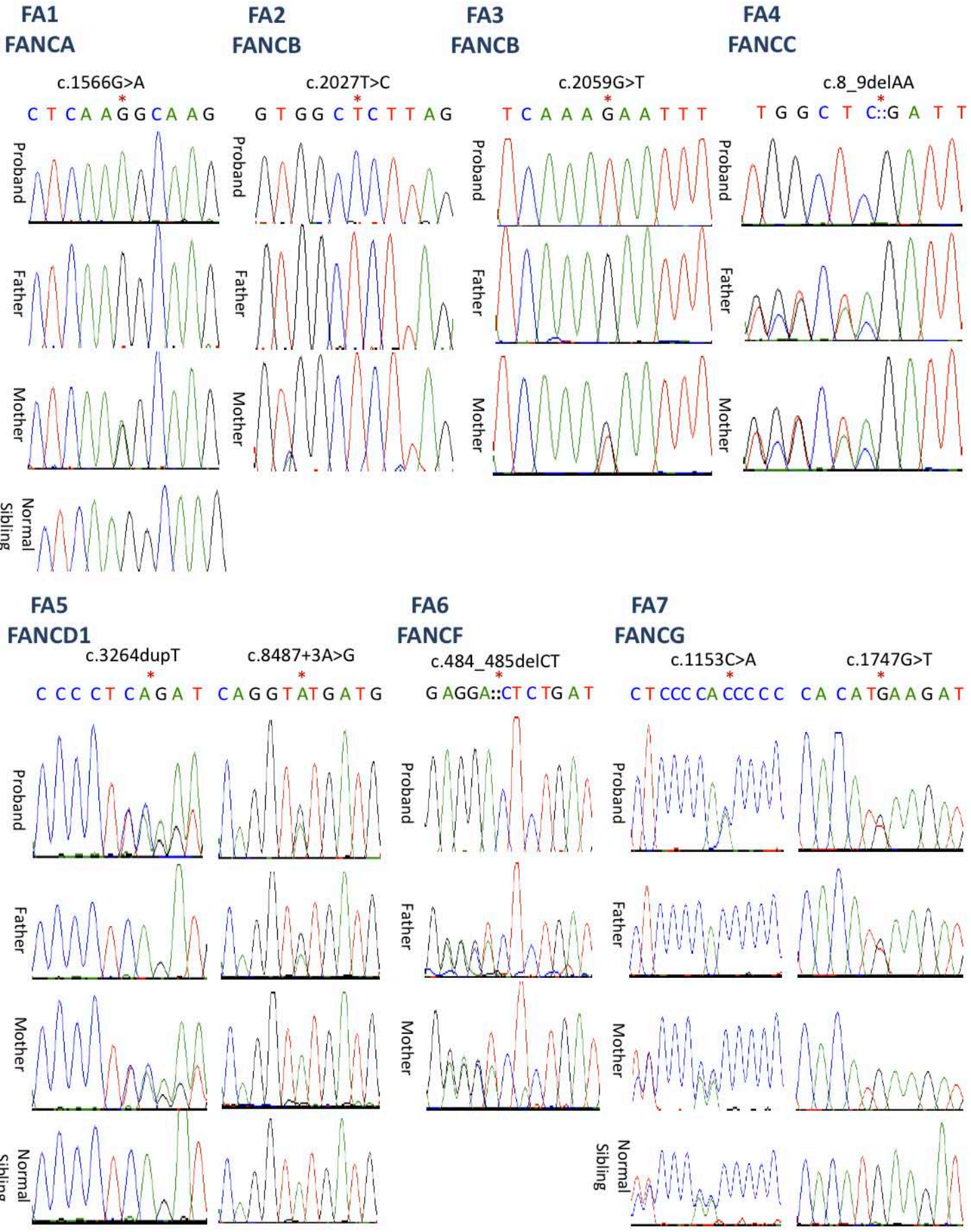
‡Extends 100kb beyond either side of the gene

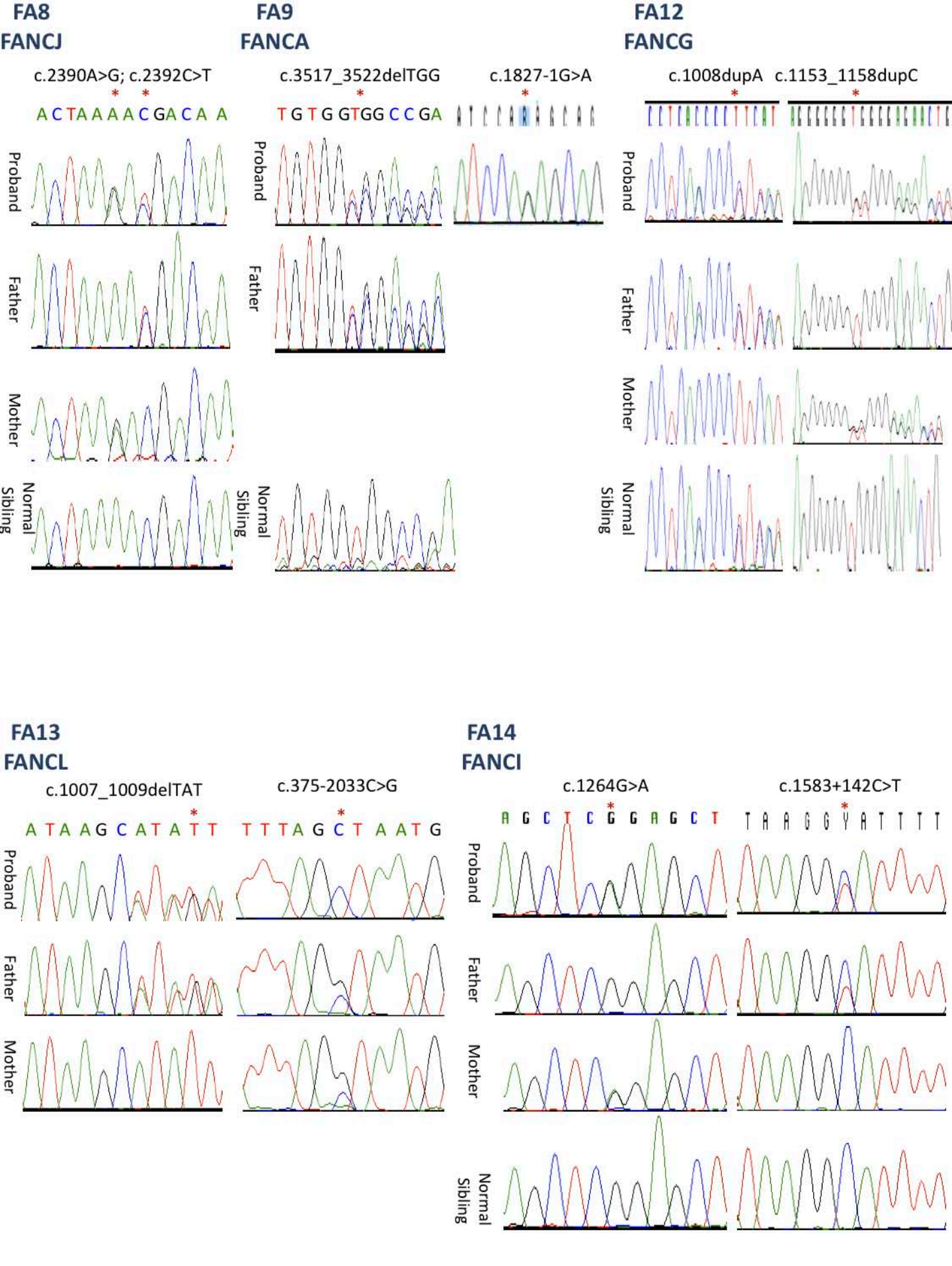
§Extends 50kb beyond either side of the gene

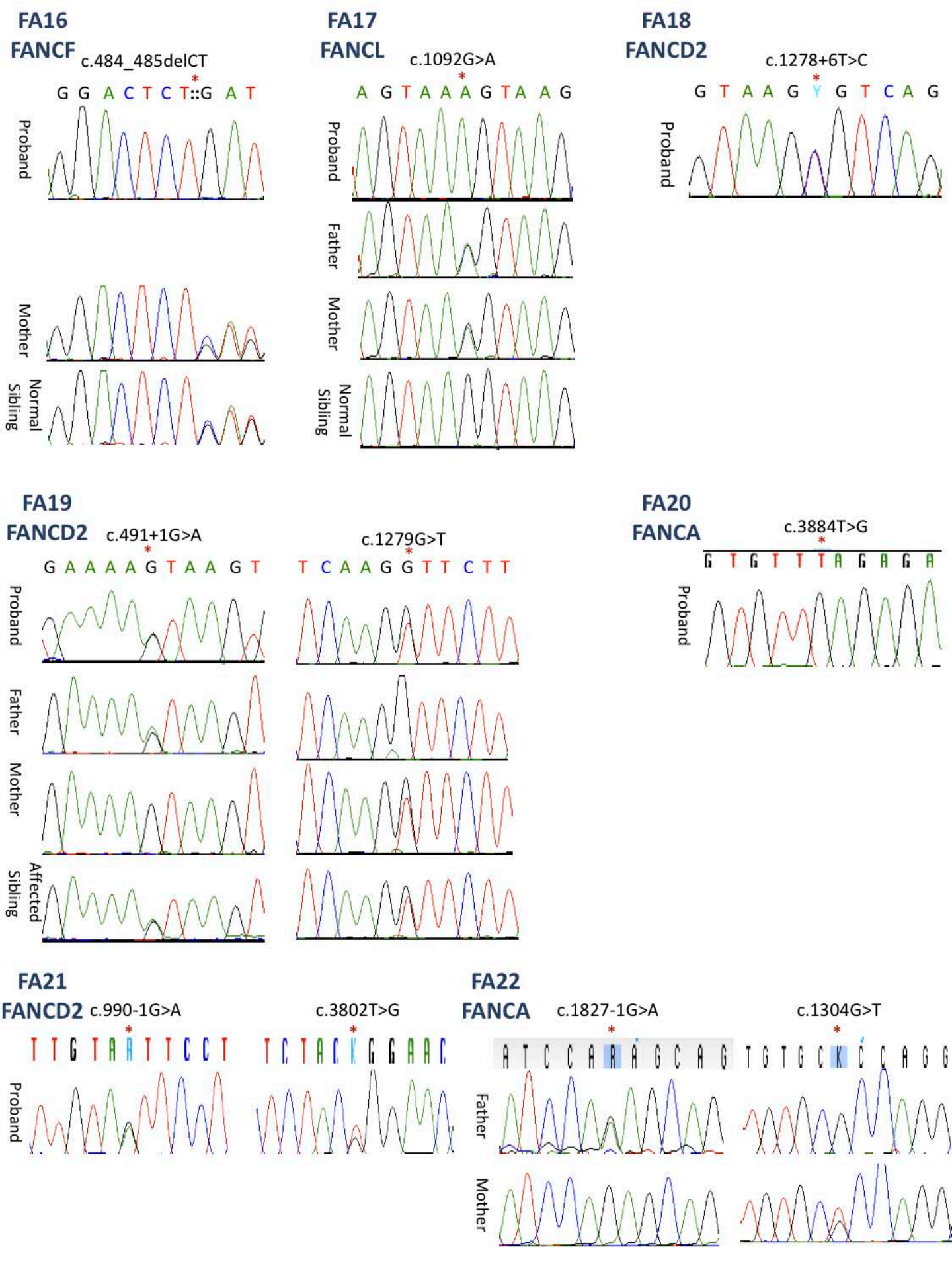
Median Spacing: 50bp

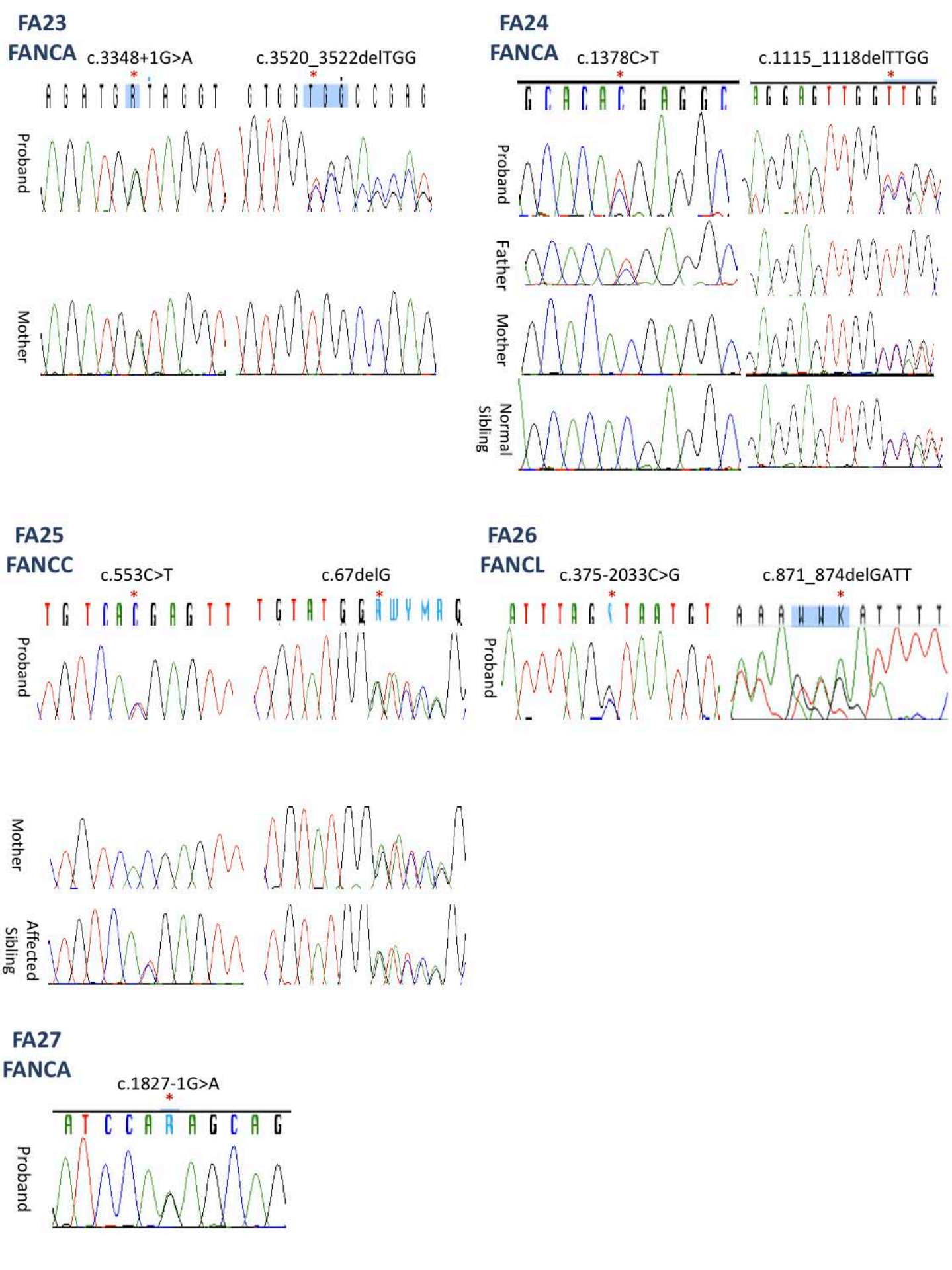
Multiple (3) Probes

Supplemental Figure 1. Validation of Sequence Variants by Sanger Sequencing

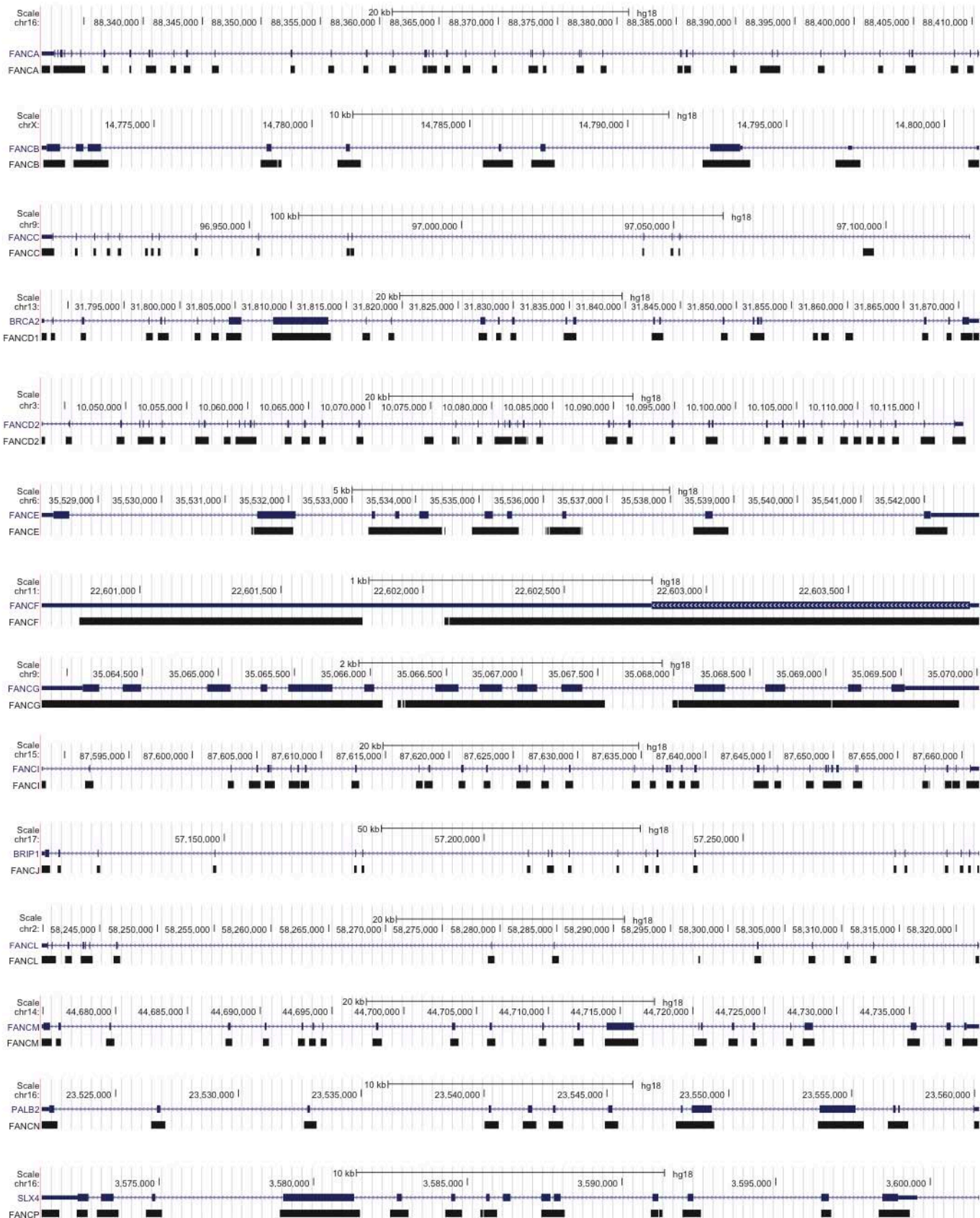



FA8
FANCJ
c.2390A>G; c.2392C>T
A C T A A A A C G A C A A
Proband
Father
Mother
Normal Sibling
FA9
FANCA
c.3517_3522delTGG
T G T G G T G G C C G A
c.1827-1G>A
Proband
Father
Normal Sibling
FA12
FANCG
c.1008dupA
c.1153_1158dupC
Proband
Father
Mother
Normal Sibling
FA13
FANCL
c.1007_1009delTAT
A T A A G C A T A T T
c.375-2033C>G
T T T A G C T A A T G
Proband
Father
Mother
FA14
FANCI
c.1264G>A
A G C T C G G A G C T
c.1583+142C>T
Proband
Father
Mother
Normal Sibling

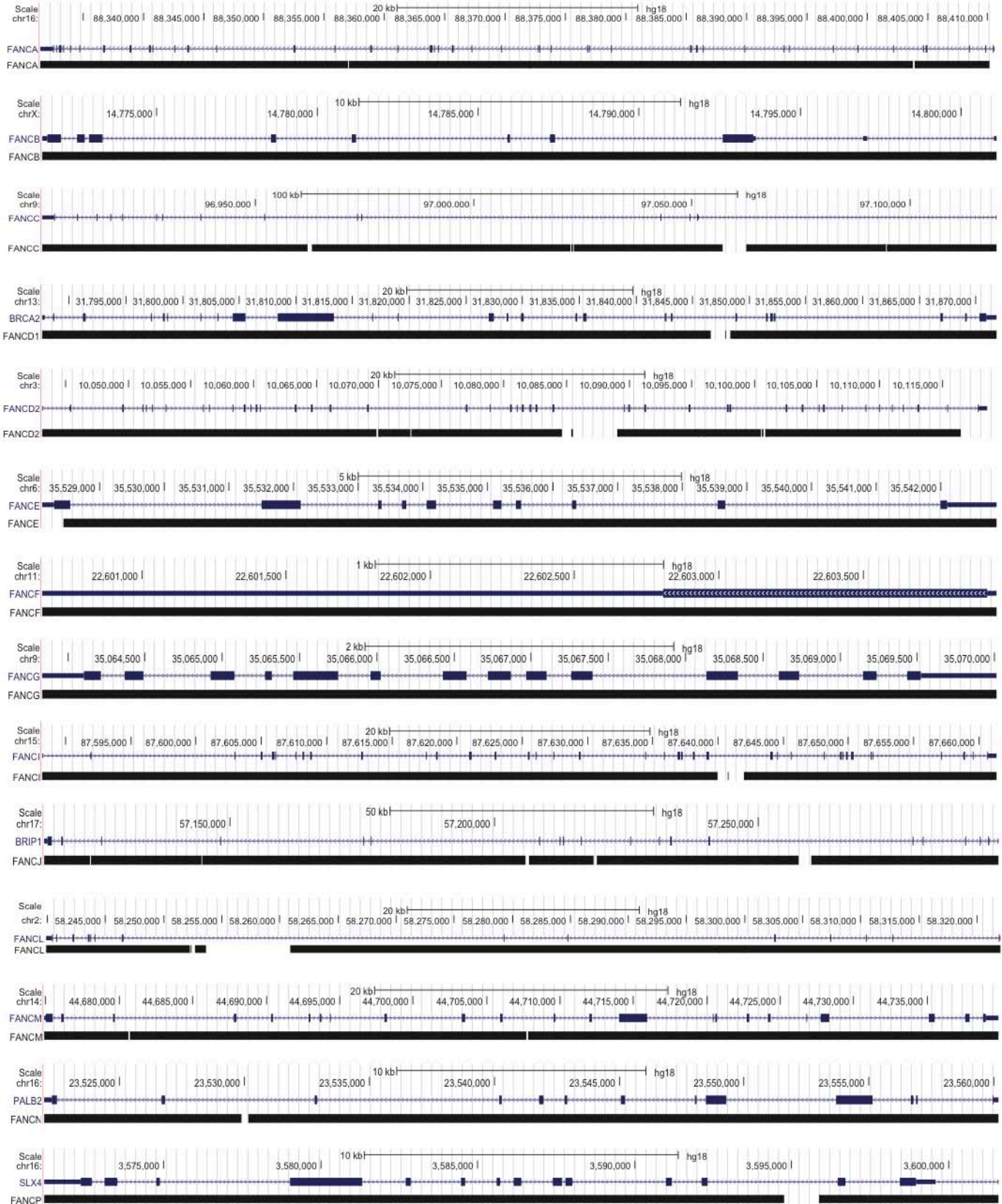

FA16
FANCF
c.484_485delCT
G G A C T C T::G A T
Proband
Mother
Normal Sibling
FA17
FANCL
c.1092G>A
A G T A A A G T A A G
Proband
Father
Mother
Normal Sibling
FA18
FANCD2
c.1278+6T>C
G T A A G Y G T C A G
Proband
FA19
FANCD2
c.491+1G>A
G A A A A G T A A G T
c.1279G>T
T C A A G G T T C T T
Proband
Father
Mother
Affected Sibling
FA20
FANCA
c.3884T>G
G T G T T T A G A G A
Proband
FA21
FANCD2
c.990-1G>A
c.3802T>G
Proband
FA22
FANCA
c.1827-1G>A
c.1304G>T
Father
Mother


FA23
FANCA
c.3348+1G>A
c.3520_3522delTGG
Proband
Mother
FA24
FANCA
c.1378C>T
c.1115_1118delTTGG
Proband
Father
Mother
Normal Sibling
FA25
FANCC
c.553C>T
c.67delG
Proband
Mother
Affected Sibling
FA26
FANCL
c.375-2033C>G
c.871_874delGATT
Proband
FA27
FANCA
c.1827-1G>A
Proband

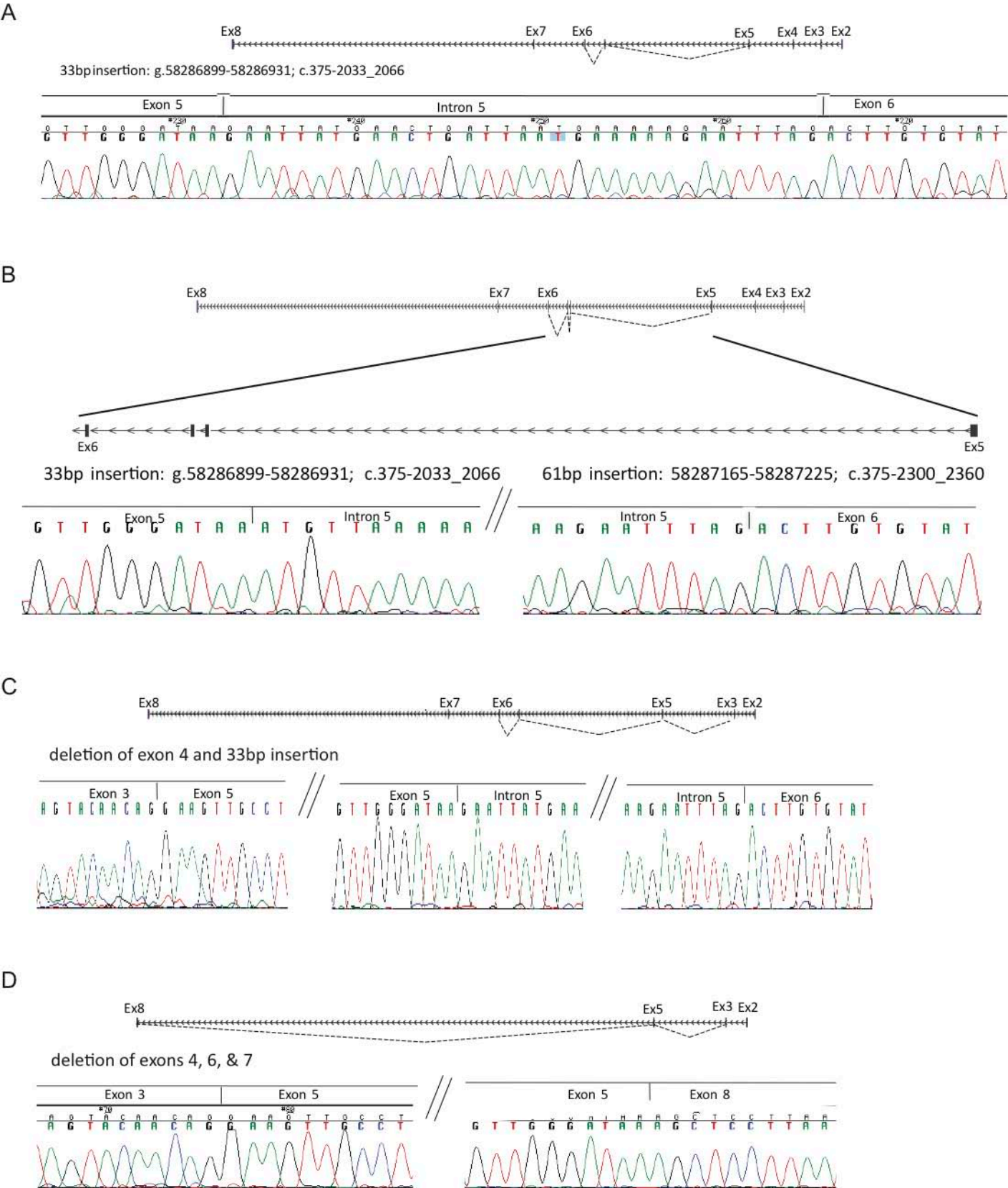
Supplemental Figure 2: Whole Exome Sequencing Coverage of FA Genes



Supplemental Figure 3: Tru-Seq Coverage of FA genes



Supplemental Figure 4: Transcript variants associated with c.375-2033C>G in FA26



Supplemental Figure 5. Transcript Variants in FA14