

Expression cloning of a cDNA for the major Fanconi anaemia gene, *FAA*

J R L T Foe, M A Rooimans, L Bosnoyan-Collins, N Alon, M Wijker, J Lightfoot, M Carreau, D F Callen, A Savoia, N C Cheng, C G M van Berkel, M H P Strunk, J J P Gille, G Pals, F A E Kruyt, J C Pronk, F Arwert, M Buchwald & H Joenje

Nature Genetics 14, 320–323 (1996).

The nucleotide sequence used in Fig. 3 was inadvertently cropped, resulting in improper alignment. The correct version is shown below.

Fig. 3 Nucleotide sequence of the *FAA* cDNA insert of the complementing clone D. An open reading frame starts at the first ATG (position 32) and ends at position 4399 with a TGA stop codon, predicting a protein of 1,455 amino acids. Amino acids are indicated by single letter codes. EMBL nucleotide sequence database accession number: X99226.

1	AGCCGCCCC	GGGCGTGTG	GCGCCAAAGC	CATGTCCGAC	TCGTGGGTCC	CGAAGTCCGC	CTCGGGCCAG	GACCCAGGGG	GCGCCGGGAG	GCCCTGGGCG	GAGCTGCTCG	CGGGAAGGCT	120
121	CAAGAGGGAA	AAATATAATC	CTGAAAGGCC	ACAGAAATTA	AAGGAATCAG	CTGTGGCCCT	CCCTGAGGAC	CATCAGGACC	TGAAATGCCCT	TTTGGCTTGG	GTAGAAGGTC	AGCTGTGTTA	240
122	K R R E	R Y N R	P E R A	D K L K	E S E	A V R L	L R S H	Q D L	N A L A	L L E V	E G F	P L C K	360
241	AAATATGCTC	CTCAGCAAG	TGATGACTG	TGACAGTCT	GAGGCTATG	CTATCATTC	TAGTTCATT	CTTCCAGTA	ATAGGCTCG	CTTCCAGGA	TCGAGCTCA	AGGCTGGGG	480
371	K L S L	L S K V	I D C	D S S E	A Y A	N H S	A I Q D	Q A S S	H L G	M V	F V	G	600
372	TATTTCTTCA	GCCGAGATG	TTGCGCTTAG	CTGTGGAGC	ATTCCTCAGG	CTCCAGCGGA	GACCACTCAC	CCTGTGCTCG	TGACTGTGGA	GCAGAGAAG	AAAGCTGCTT	CCCTTTAGA	720
361	I L S L	A G M V	V A S S	V G Q	I C T	A P A E	T S H	P V L	T S H	Q R K K	L S L	S	840
481	GTITGCTCAG	TATTTATGG	CACACAGAT	GTTCCTCCGT	CTTTCCTCTT	GTGAGAAT	ATGGAAATA	CAGAGTCTT	TGTTGTTTGA	AGCGGTTGTC	CACTCTCAG	TACAAGGAT	960
482	H S M	A S L M	H S M	P S S	L R G	F Q K H	S D L R	R E T	G S L R	G Y H	H L H	Q	1080
601	TTGAGCCCTG	CAAGAGCTG	TGGAAAGCCA	TCCCGAATG	CATGCTGTGG	GATCCGCTCT	CTTCAGGAAT	CTGTGCTGCC	TTTGTGAACA	GATGGAAGCA	TGCTCCGAG	ATGCTGAGCT	1200
591	L S L	Q E L	L E S H	P D M	H A V	G S W L	F R N	L C C	O M E A	S C Q	H A D V	230	
721	CGCAGAGCC	ATGCTTCTG	ATTTTGTCTA	AAATGTTGTT	TTGAGGGGAT	TTGAGAAAA	CTCAGATTA	AGAAAGAAC	AGAAATGGCC	CAGGTCCAGG	TTGATGACT	840	
722	A R A	H L S	D F V	H F V	L R G	F Q K H	S D L R	R E T	G S L R	G Y H	H L H	Q	960
841	CGAGAAATG	CTGATTTTGT	CACCTGACG	TTTGGCTTGT	GGATGACAG	AGGATGCTC	CATCCACAG	ATCCCTGAGG	AGTGTGTGAT	GGACACAGC	TGGCAGTGT	960	
271	O R M	L I F A	L D A	L A A	G V Q	E E S S	T H K	I V R	C W F G	V F S	G H T	L G S V	310
961	AAATTCACCA	GATCCCTGTA	AGAGCTTCT	CATGCAATCC	CTGACACAGA	TACTACTCA	CAGCCCTGTC	CTGAAAGCAT	CTGATGCTT	TGAGAGTGC	ACAGAGTGA	CCCTTGGCCG	1080
311	I S T	D P L	K R F P	S H T	L T Q	I L T H	S P V	L K A	S D A V	O M O R	E F W	A R V	350
124	GALGACCCCT	CTGCTCACG	CACCTTACCC	CTGATGTCTA	GTCCAGAGGA	CTGCTGCTCA	ATGTTGGCA	AGAGTGTGA	AGCCGAGAG	TTTCACTGSC	AGAGAGTCT	1200	
351	T H P L	D A R L	L G V R	L S A	A F L	N M G L	H Q	E V L	T O E	L H W	TGCTCACTC	1800	
1201	CTCTTGTGG	TTGCTCTG	TTGCTCTG	TCCAGAGGG	CAGGACCTGC	TGSAAGCTC	AGGCTGGCT	TTGATGGCCC	AGGCTGGCC	GAGCTGGCC	CTGACAGCA	1320	
391	S S L	V A L	V V C F	F E A	Q L L	L E D W	V A R	L M A	Q A F E	S Q L D	S M V T	430	
1321	GTTCCTGGT	GTGGCCGAG	CAGCACTGGA	GGGCCCTCT	GGCTTCTGTT	CATATCAGA	CTGTTCTCAG	GCCTCTGTG	GGAGACACG	AGGCTACAT	GGCTGACGA	AGAGAGGCT	1440
1322	A R A	H L S	D F V	H F V	L R G	F Q K H	S D L R	R E T	G S L R	G Y H	H L H	Q	1560
441	GGTCTTCTG	TTTACTGCT	TGTCAGAAT	CGTGGCTTTT	GAGTCTCCCC	GGTACTGCA	GGTGGACTC	CTCCACCAC	CCCTGCTTCC	CACCAAGCAT	GGCTCTCTC	1560	
471	V F L	F T F	L S E L	V P F	E P T	R Y L Q	V H I	L H P	P L V P	S K Y K	S L L	L T D Y	510
1561	CATCTATTG	CGCAAGCAC	GGCTGCCGA	CCTCAAGST	TCTATAGAAA	ACATGGGACT	CTCAGAGCAT	TTGTCTCAG	CTCGGGACT	TACTGAGCC	CACAGCCAG	CTTCTGAGA	1680
511	S L	R A K	R L A D	L K V	S E E	N M G L	V E D	L S A	G D I	T F P	H S Q	A L Q D	550
1561	TGTTTAAAG	CCATCAATG	TGTTTGAACA	TCCGAGGAC	ATCCGACTCA	CGTCTGSGA	GGCCAGGAT	TTCCAGGAG	GTCCCTACT	GTCCCTACT	CTCCCTCCC	CTCTCACTC	1800
551	V E K A	I M V	F E H	T G N	I P V	T V M E	A S I	F R R	P V Y V	S H F I	P A L	L T F	590
1801	CTCAAGTCTC	CCCAAAATCC	CTCACTCCCG	TGTGGCGTGT	ATAGAGTCTC	TGAAGAGGC	AAATAAATC	CTGTACTCAC	CTACTGCTCC	GGCTGCTCTG	CTCTGTAGA	1920	
591	R D V	L P K V	D S R	V A F	I E S	L K R A	D K I	P P S	L Y S T	V C O A	C S A	A E E	630
1921	GAGCGAGAA	GATCCAGCC	TGGAGATGAG	GGGCAAGCC	AACTCTGCTG	AGGAGCCCT	GGCAGGCTC	ACAGCTGAC	TGGAGYGT	GAGAGGCTCC	ATCAGAGCC	CGAGCCAGC	7040
591	K P E	D A R	L G V R	L S A	A F L	N M G L	H Q	E V L	T O E	L H W	TGCTCACTC	840	
2041	TGATGTTATA	TCCGACAGG	TGCGAGTAT	TTCTGAAAG	CTGAGGCTGC	TCTGGGCCA	CAATGAGAT	GACAGCAGG	TGAGAAATC	AAAGATGAG	CTCAGACTG	ACAGCCGAG	2160
671	D V I	S A Q	V A V I	S E R	L R A	V L G H	N E D	D S S	V E I S	K I Q	L S I	N T P R	710
2161	ACTGAGCCA	CGGGAAGCA	TTGCTGTGGA	CCCTCCCTG	AGCTCTTCT	GTGAGAACC	GATGGCTGCC	TCCAGTGTG	CTCCCCGGA	GAGGAGAGG	CTTGGGGTCT	CCCTCTTCT	2280
711	L E D	F L A	D K V	S E E	N M G L	V E D	L S A	G D I	T F P	H S Q	A L Q D	750	
2281	GAGGACATG	TGAGACATG	TGCTCCCTGC	AGTGTCCACC	GGCTCTGCT	GTGAGAACC	GATGGCTGCC	TCCAGTGTG	CTCCCCGGA	GAGGAGAGG	CTTGGGGTCT	CCCTCTTCT	2400
751	R T M	C G R	V L P A	V L T	R L C	Q L L R	H Q G	P S I	S A P H	V L G L	L A A	L A V	790
2401	CTCGGTTGC	TCCAGGCTG	CGCTCCCGA	GGTGATGAG	GGTCTCTGCT	CGACTGGTG	TGGCCTTCT	GTCCCTGCG	CTCTCTGCG	CTCTCTGCG	TGTAGAGA	GGATTTCTT	2520
791	L V G	S R S	A L P E	V G P	G P P A	A P G A	A P G A	G L P	V P A	L F P	V P A	L F P	830
2521	GTCTCTGCT	CTGAAATTT	GTACAGAGC	AATTTCTTC	CTCTCTGCA	AGTTTCTCT	CGACTGAGA	GATCTTTGT	GTCTCTGCT	ATCTCTGCT	CTTTTAAAA	AGTTTCTGCT	2640
811	F P E	T A S	T A S	T A S	T A S	T A S	T A S	T A S	T A S	T A S	T A S	T A S	850
2641	CTCATGCTC	AGATTTGCT	CAGAGCCCG	ACAGCCCTT	TCGAGGAGG	ACGTAGCCG	CCCTTCTG	AGACCCTTG	ACCTTCTCT	TCCAGACTG	CAGAGAGCT	CCCTTCTCT	2760
871	L M F	R L F	S E A R	Q P L	S E E	D V A S	L S W R	H I L	P S A	A D W	Q R A	A L S L	910
2761	CTGACACAC	AGAACCTTC	GAGAGTGTG	GAAAGAGGA	GATGTTCACT	TACTTACCA	AGACTGGTTA	CACCTGGAG	TGGAAATTA	AGCTGACTC	GATGCTTCT	CAGATACGA	2880
311	W T H	R T P	R E V L	K E E	D V H	L T Y Q	D W L	H L E	L E I Q	F E A	D A L	S D T E	950
2881	ACCGCAGAC	TCCACCAAT	GGCGATCTA	TGACACTTT	CTCCCTGACT	CTCCCTGCT	AGGGGCTGT	GACGAGAGC	TGCAAGTCT	GGTACACT	TTTGTACCT	CAGTATGGA	3000
951	R Q D	F H Q	W A T H	E H F	L P E	S S A S	G G C	D G D	L Q A A	C T I L	V N A	A L M D	990
3001	TTTCCACCA	AGCTCAAGA	GTATGACCA	CTCAGAAAT	TCTGATTTG	TCTTTGGTG	CCGACAGGA	AATGAGGATA	TATTTTCAG	ATTGAGGAG	ATGGAGACT	ACCTTGGACT	3120
991	F H Q	S S R	S Y D H	S E N	S E N	V F G G	R T G	N E D	I S K	L Q E	M V A	D L E L	1030
3121	CAGCAGAGC	CCATATGTC	CTCTCCGCA	CACCCCTTC	CAGGAGGAT	CTCTCTTGT	GATTTTCTG	AGACGGCTC	AGGCTGTGAC	AGGCGGTTG	AGGCGGTTG	CTCCAGGCT	3240
1031	Q O D	L I V	P L G H	T E S	Q E H	F L F E	I F R	R L S	Q A L T	S W S	V A	E L Q	1070
3241	GAGACAGG	GAGGCTGAA	TGTACAAGC	GATCCCTCT	CGCCTGCTT	CGTCTGCT	CTGGCCGCT	AGCTTCCAG	CATCACTCC	AGATGCTGAG	AGTTTCCA	3360	
1071	R O R	E L L	M Y K R	I L L	R L P	S V I L	C G S	S F Q	A E Q P	I T A R	C E	T F H	1110
3361	CTTGTGCA	CTGTGATGA	GAAGCTTGT	CTCCCCGGA	GGTCCCTGA	CACAGGACT	CATCGCCAC	TTCTTCAAG	GCCTCTGAA	CCCTCTGTC	GGAGAGCAG	ACCCCTCTC	3480
1111	L V N	F V L	G A T F	C A A F	G A T F	C A A F	G A T F	C A A F	G A T F	C A A F	G A T F	C A A F	3520
3481	GATGCTGAC	TTCATATGG	CCAAGTCCA	GACGAAATG	CCCTTGAAT	TGACTCTGC	GCTGCTGTG	TGGCCGAGC	TGAGAGTCT	GGTGTGCT	GGTGTGCT	CAGCATCCA	3600
1151	M V D	F I L	A K C O	T K C	P L I	L T S A	L V W	W P S	L E P V	L L C C	K N R	R H C	3190
3601	GAGCCGCTG	CGGAGAGAC	TGCAGAAGCT	ACRAGAAGC	CGGCAATTT	CCCTGCOCT	CCCTGCOCT	CCCTGCOCT	CCCTGCOCT	CCCTGCOCT	CCCTGCOCT	CCCTGCOCT	3720
1191	S P L	R F L	D O F L	D E F	D O F A	A G D S	D S P	E A S	P A P	N P F	W L S	A A A L	1230
3721	CCACTTGGC	ATTCAACAG	TCCAGGAGA	AAACATGAG	AGGACTTAA	AGAGCTGGA	CTCCGAGGA	GAGGACTAT	TGTTTCTCT	TTTCTGATG	TGTTTCTCT	TTTCTGATG	3840
1231	H F A	I Q O	V R E E	N I R	K O L	K K L D	C E R	E L	V P L	F F F	S L M	G A L S	1270
3841	GTCATCTG	ACCTCAATA	GCACCACAGA	CTCGCAAAG	GTCTTCCAG	TTTGTGAG	AAITCTGAT	GGTTTAGGA	AGAGGAGAT	ATGCTGGT	GCATCTTTC	AGTTGACGA	3960
1271	S H L	T S N	S T T D	L P K	A P H	V C A A	I L E	C L E	K R K I	L E A	L F P	Q L T E	1310
3961	GAGTGCCTC	AGGCTGGGC	GGCTCTCTC	CCCTTGGCC	CCGATGAGC	ACACGAGCT	CTGTCTTTC	GCTTCTTTC	GCTTCTTTC	GCTTCTTTC	GCTTCTTTC	GCTTCTTTC	4080
1311	S D L	R L C	R L A	R V A	P D Q	H T R L	L P P	A F Y	S L L S	Y F F	E D A	A I R	1350
4081	AGAGCCCTC	CTGCTGTTG	CTGTGACAT	GTACTTGA	CTGTCCAGC	TCTTCTGCG	TGGGATACA	AGCCAGATTT	CTCCAGAGC	TCCAGAGC	CTCCAGAGC	AGGCTCAAGC	4200
1351	E A F	L H V	A V D M	V I K	L V Q	L F V A	G D T	S T V	P P A	G R S	I F L	K Q O	1390
4201	CACCCCTGG	GAACTGATA	CAAAAGCTG	TCTTTTCTG	CTGCAATTA	TACCTGGTC	CCGAAAAG	AGCTTCTAC	ACGTGGAGC	CGCTGGCT	GATCTGGCC	CTCTGGACC	4320
1391	N P V	E L I	T K A R	L P L	L Q L	I F R C	C P K K	S F S	H V A	H W	S R	G	1430
4321	AGAGTGCAT	CGCCCTCC	AGAGCAGCA	CCAGCTCCG	CCGACCTGT	ACCTGTCCA	GGAGCTCAT	CTCTTCTAC	GGAGCTCC	ACTCCACAC	ACCCACCTC	CCCTCTAAT	4440
1431	E V S	A A L	Q S R O	Q A A	P D A	D L S O	E P H	I F					1456
4441	AAATTTATAC	AGCAATAACA	TGGAGCTTCT	GTTCACATA	AAAGTGGATT	ACAAATCTC	TGCACTGCTT	TAGTGGGAA	AGGAATCA	TATTTATGA	CTGTCCGGC	CCGACTCACT	4560
4561	CAGCTTTGG	GGGAATATA	ACCAAGTACT	CCAGAGCAGA	GGAGGCTTAC	TTGAGCCGTA	CACCAAGCC	CCCTCAAGA	CCCAAGGCG	GGAGCAGCT	CTGCGGCTG	CATCGGCTG	4680
4681	CACAGAGC	GGGAATATA	ACCAAGTACT	CCAGAGCAGA	GGAGGCTTAC	TTGAGCCGTA	CACCAAGCC	CCCTCAAGA	CCCAAGGCG	GGAGCAGCT	CTGCGGCTG	CATCGGCTG	4800
4801	GGCTGGTGG	CCAGAGTGT	GGTTCCTCT	CCAGGAGCAG	GGCCTTGTCT	TGGTGTGTC	GGCTGGTGG	CACCATGAC	ATGTGTACT	TSAGTGTGT	GGCCTTCTA	ACCCCTCCG	4920
4921	AGAGCCCTC	CCAGGAGTGT	TCCAGCTCAG	CAGCCCTGG	AGGGGGTCT	GACTCACACT	TACTGCAAG	GGCTGGTCC	CGAATGTCC	ATTTGGTGA	CGAAGAGTG	CTTCCGCTC	5040
5041	GATTAAGAG	TGCGAGGCCA	CGCCCTGGG	AGGGGGTCT	GACTCACACT	TACTGCAAG	GGCTGGTCC	CGAATGTCC	ATTTGGTGA	CGAAGAGTG	CTTCCGCTC	TTGAGAGTT	5160
5161	GTCCACATTC	GTCCACATTA	TAGTTCGCA	CCTTGTGAG	GGGAGGATCC	AGTGGATCCA	GGCCTCTGAT	GCTTCAACT	CCGAGGCGA	CGAGATGAC	AATGTGAAG	CATCAGACT	5280
5281	GGGAGACT	TTCTGACAT	GGTTCACAT	CGATGGGCC	CAGCAGAGG	CGCTATGAG	GGCCTCTGTA	TTAGATGCT	TAACTCTG	TATACAGT	CAGTGGCTT	CATCAGACT	5400
5401	GTGCTTTCA	GGATGTTTT	TAAAGAAC	TCGAAAGCT	ATTTCTTAA	AAAAAAAA	AAAAAAAA	AAAAAAAA	AAAAAAAA	AAAAAAAA	AAAAAAAA	AAAAAAAA	5503

correction

The insulin-like growth factor II receptor gene is a target of microsatellite instability in human gastrointestinal tumours

Nature Genetics 14, 255–257 (1996).