

## 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	GGGCTGTGAG	GGCCAAAGCC	CATGTCCGAC	TCGTGGTCCC	CGAAGTCCGC	CTCGGGCCAG	GACCCAGGGG	GGCCGGGGAG	GGCTGTGGCC	GAGCTGTCTG	CGGGAAGGCT	120
121	CAAGAGGGAA	AAATATAATC	CTGAAAGGCC	ACAGAAATTA	AAGGAATCAG	CTGTGGCCCT	CCCTGAGGAC	CATCAGGACC	TGAAATGCCCT	TTTGGCTTGG	GTAGAAAGCT	AGCTGTGTTA	240
122	K R R E	R Y R N	P E R A	D K L K	E S E S	A V R L	L R S H	Q D	L N A L	L L E V	E G	P L C K	360
241	AAATATGCTC	CTCAGCAAG	TGATGACTG	TGACAGTCT	GAGGCTATG	CTATCATTC	TAGTTCATT	CTTCCAGTA	ATAGGCTCTG	CTTCCAGGA	TCGAGCTTCA	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	P V G	600
372	TATTTCTTCA	GCCAGGATG	TTGCGCTTAG	CTGTGGAGC	ATTCCTCAGG	CTCCAGCGGA	GACCACTCAC	CCTGTGCTGC	TGACTGTGGA	GCAGAGAAG	AAAGCTGTCT	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	L E	840
481	GTITGCTCAG	TATTTATGG	CACACAGAT	GTTCCTCCCT	CTTTCCTCTT	GTGAGAAT	ATGGAAATA	CAGAGTCTCT	TGTTGTTTGA	AGCGGTTGTC	CACTCTCAGC	TACAAGGAT	960
482	H S M	A H S M	A H S M	A H S M	A H S M	A H S M	A H S M	A H S M	A H S M	A H S M	A H S M	A H S M	1080
601	TTGAGCCCTG	CAAGAGCTGC	TGGAAAGCCA	TCCCGAATG	CATGCTGTGG	GATCCGTGCT	CTTCAGGAAT	CTGTGTGGCC	TTTGTGAAAC	GATGGAGACA	TGCTCCGAGC	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	L S L	1320
721	CGCAGAGCC	ATGCTTCTG	ATTTTGTCTA	AAATGTTGTT	TTGAGGGGAT	TTGAGAAAA	CTCAGATAA	AGAAAGAAC	AGAAATGGCC	CAGGTCCAGC	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 T	V E E	K H F	G Y T	V D V	960
841	CGAGAAATG	CTGATTTTTC	CACCTGACG	TTTGGCTTGT	GGATGACAG	AGGATGCTC	CACCTACAAG	ATCCGTGAGT	GGTGTTCGAG	AGCTTGTAGT	GGACACAGCC	TGGCACTGTT	1080
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	1200
961	AAATTCACCA	GATCCCTGTA	AGAGCTTCTT	CAGTCAATCC	CTGACACAGA	TACTACTCA	CAGCCCTGTC	CTGAAAGCAT	CTGATGCTG	TGAGAGTGC	ACAGAGTGA	CCCTTGGCCG	1320
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	1440
124	GALGACCCCT	CTGCTCACG	CAGCTTACCC	A S E P	GTGAGTCTGA	CTCGAGAGGA	ATGCTTGGCC	AGAGTCTGGA	AGCGCGAGG	TTTCACTGSC	AGAGAGTCT	1200	
351	T H P	L S L	S L Y R	R L F	V M L	S S E	L G	H Q	E V T	O E	L H W	TGCTCACTC	1800
1201	CTCTTGTGG	TTGCTCTGCT	TTGCTCTGCT	TCCAGAAAGG	CAGGACCTGC	TGSAAGCTC	AGGCTGGCC	TTGATGGCCC	AGGCTATCGA	GAGCTCGAGC	CTGACAGACA	TGCTCACTC	1920
391	S S L	V A L	V V C F	F E A	Q O L	L E D W	V A R	L M A	Q A F E	S C Q	L D S	M V T A	2040
1321	GTTCCTGGT	GTGGCCGAG	CAGCACTGGA	GGGCCCTGCT	GGCTTCTGTT	CATATCAGA	CTGTTCTCAAG	GCCTCTGTG	GGAGACACG	AGGCTACACT	GGCTGACGCA	AGAGAGGCT	2160
1441	GGTCTTCTG	TTTACTGCT	TGTCAGAAT	CGTGGCTTTT	GAGTCTCCCC	GGTACTGCA	GGTGGACTC	CTCCACCAC	CCCTGTCTTC	CAGCAAGACT	GGCTCTCTC	TGACAGACTA	2280
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 T D Y	2400
1561	CATCTTATTG	CGCAAGCAC	GGCTGCCGA	CCTCAAGSTT	TCTATAGAAA	ACATGGGACT	CTCAGAGACT	TTGTCTCAG	CTCGGGACT	TACTGAGCC	CACAGCCAG	CTTCTGAGA	2520
511	S L	R A K	R L A D	L K V	S E N	M W L	E D	L S A	G D I	T F P	H S Q	A L Q D	2640
1581	TGTTTAAAG	CCATCACTG	TGTTTGAACA	TCCGAGAC	ATCCGACTCA	CGTCTGSGA	GGCAGGATTA	TTCCAGAGCC	GTCCCTACT	GTCCCTACT	CTCCCTCCC	CTCTCACTC	2760
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	2880
1801	CTCAAGTCTC	CCCAAAATCC	CTCACTCCCG	TGTGGCGTTC	ATAGAGTCTC	TGAAGAGGC	AAATAAATC	CTGTACTCAC	CTACTGCTCC	GGCTGCTCTG	CTCTGTAGA	1920	
591	R D V	P K V	P 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	2040
1921	GAGCGAGAA	GATCCAGCC	TGGAGATGAG	GGGCAAGCC	AACTCTGCTG	AGGAGCCCTC	GGCAGCTGAC	ACAGCTGAC	TGGAGYAGT	GAGAGGCTTC	ATCAGAGCC	CGAGCCAGC	2160
531	K P E	D A R	L G V R	A S P	A F L	C S E	L G	H Q	E V T	O E	L H W	TGCTCACTC	2280
2041	TGATTTATA	TCCGACAGG	TGCGAGTAT	TTCTGAAAGA	CTGAGGCTGC	TCTGGGCCA	CAATGAGAT	GACAGCAGG	TGAGAAATC	AAAGATGAG	CTCAGACTA	ACAGCCGAG	2400
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	2520
2161	ACTGAGCCA	CGGAAAGCA	TTGCTGTGGA	CCCTCCCTGC	AGCTCTTCTT	GTGAGAAGC	GATGGCTGCC	TCCAGTGTG	CTCCCCGGA	GAGGAGAGG	TTCTGGGCTG	CCCTCTTCTG	2640
711	H E	F A G	T G T G	T G T G	T G T G	T G T G	T G T G	T G T G	T G T G	T G T G	T G T G	T G T G	2760
2281	GAGGACATG	TGAGACATG	TGCTCCCTGC	AGTGTCTCAC	GGCTCTGCTC	GTGAGAAGC	GATGGCTGCC	TCCAGTGTG	CTCCCCGGA	GAGGAGAGG	TTCTGGGCTG	CCCTCTTCTG	2880
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 A A	L A V H	2960
2401	CTCGGGTGC	TCCAGGCTGC	CGCTCCCGGA	GGTGGATGTC	GGTCTCTGCT	CAGCTGGTGC	TGGCCCTTCT	GTCCCTGCTC	CTCTCTGCTC	CTCTCTGCTC	TGTAGAGAGA	GGATTTCTT	2520
791	L G E	S R S	A L P E	V G P	G P P A	A P G A	A P G A	A P G A	A P G A	A P G A	A P G A	A P G A	2640
2521	GTTCCTGCT	CTGAAATTT	GTACAGAGC	AATTTCTTAC	CTCTCTGCA	AGTTTCTTCT	CGAGTACAGA	GATACTTCTG	GTGCTTCTG	ATCCTGCTC	CTTATTAAGA	AGTTTCTGAT	2760
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	2880
2641	CTCATGCTC	AGATTTGCT	CAGAGCCCG	ACAGCCCTT	TGAGGAGG	AGTACAGC	CCCTTCTG	AGACCTTGC	ACCTTCTCT	TGACAGCTC	CAGAGAGCT	CCCTTCTCT	2960
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	3080
2761	CTGACACAC	AGAACCTTC	GAGAGTGTG	GAAAGAGGA	GATGTTCACT	TAACTTACA	AGACTGGTTA	CAGCTGGAGC	TGGAATTTCA	AGCTGACTT	GATGCTTCTT	CAGATACTGA	2880
3111	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	3000
2881	ACGCGAAGC	TGCCAAGT	GGCGATCCG	TGAGCACTTT	CTCCCTGAAT	CTCCGCTTCT	AGGGGCTGCT	GAGCGAGACC	TGCAAGTCTC	GGCTGACTCT	CTGCTGAGC	CAGTATGGA	3080
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	3160
3001	TTTCCACCAA	AGCTCAAGGA	GTATGACCA	CTCAGAAAT	TCTGATTTG	TCTTTGGTGC	CCGCAAGAGA	AATGAGGATA	TATTTTTCAG	ATTGAGGAG	ATGGAGACTG	ACCTTGGAGT	3240
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	3320
3121	CAGCGAAGC	CCATATGTC	CTCTCCGCA	CAGCCCTTCC	CAGGAGCACT	CTCTTCTTGA	GATTTTCTGC	AGACGGCTC	AGGCTCTGAC	AGGCGGTTG	AGGCTGGCTG	CTGCTGAGC	3400
1031	Q O D	L I V	P L G H	T E S	O R H	F L F E	I F R	R L F	Q A L T	S W S	V A A	E L Q	3480
3241	GAGACAGG	GAGGCTGAA	TGTACAAGG	GATCCCTCTC	CGCCTGCTT	CGTCTGCTC	CTGGCGACT	AGCTTCCAG	CATCACTGCT	AGATGCTGAG	AGTTTCTTCA	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	3440
3361	CTTGGTCAAC	CTGTGATGA	GAAGCTTGTG	CTCCCCGGA	GGTGGCTGA	CACAGGACT	CATCGCCAC	TTCTTCAAGG	GCCTCTGTAA	CCGCTGTCTC	GGGAGACAG	ACCCCTCTC	3480
1111	L V N	F A T	G A N F	C A N F	C A N F	C A N F	C A N F	C A N F	C A N F	C A N F	C A N F	C A N F	3560
3481	GATGCTGAC	TTCATATGG	CCAAGTCCA	GAGCAAAATC	CCCTTGAATT	TGACTCTGTC	GAGTCTGTG	TGGCCGAGC	TGAGAGCTC	GGCTGACTCT	GGCTGAGACA	CAGACTTCCA	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	3680
3601	GAGCCCGTC	CGGAGAGAC	TGCAAGAGCT	ACAGAAGGTC	CGGCAATTTG	CCAGCGATTT	CCCTTCCOCT	GAGGCTGCTT	CCGCAAGCTC	CAACCCAGC	TGGCTCTCAG	CTCTGAGCT	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	3800
3721	CCACTTGGC	ATTCAACAG	TGAGGAGGA	AAACATGAG	AGGCACTTAA	AGAGCTGGA	CTCGAGAGA	GAGGAGCTA	TGCTTCTCTC	TTTCTGATG	TGCTTGTGTC	TGCTGATGTC	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 E L	L V P L	F F F	S L M	G A L S	3920
3841	GTCACATCTG	ACCTCAATA	GCACACAGA	CTGCCAAAG	GCCTTCCAG	TTTGTGAGC	AAATCTGAT	GGTTTAGAGA	AGAGAGAGT	ATGCTGGTCT	GCATCTTCTC	AGTTGACAGA	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	O L T E	4040
3961	GAGTGACTC	AGGCTGGGC	GGCTCTCTCT	CCCTTGGCT	CCGATGAGC	ACAGCGAGCT	GTCTGCTTTC	GCTTCTTACA	CTTCTTCTCT	CTATCTTCTC	GAGAGAGGCT	CTATCAGGTA	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	4160
4081	AGAGCCCTC	CTGATGCTG	CTGTGACAT	GTACTTGAAG	CTGTCTCAGC	TCTTCTGCTC	TGGGATACA	AGCACAATTT	CCGCTGAGC	TGCGAGAGC	CTCAGACTCA	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	L F L	K Q O	4280
4201	CAACCCCTG	GAACTGATA	CAAAAGCTG	TCTTTTCTG	CTGCACTTAA	TACCTGGCT	CCGAAAGAG	AGCTTCTCAC	ACGTGGAGCA	GGCTTGGCT	GATCTGGCC	CTTGGAGCC	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 S F	A L R	D L	4360
4321	AGAGTGAAG	CGCCCTCCCT	AGAGCAGCA	CCAGGCTGCT	CCGACCTGTC	ACCTTCCCA	GGAGCTACT	CTCTTCTGAC	GGGACCTCCC	ACTCCACAC	ACCCACCTC	CCCTCTAAT	4440
1411	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	TGGAGCTCTT	GTTCGACTAA	AAAGTGGATT	ACAAATCTCC	TGCACTGCTT	TAGTGGGAA	AGGAATCAAT	TATTTATAGA	CTGTCCGGCC	CCGACTCACT	4560
4561	CAGCTTTTC	GGGAATATA	ACCAAGTACT	CCAGAGCAGA	GGAGGCTTAC	TTGAGCCGTA	CAGCAGCCCT	CGCAAGGGG	GGAGCAGCTT	CGGCTGGCTG	CATCGGCTG	4680	
4681	CACAGAGC	GGGAATATA	ACCAAGTACT	CCAGAGCAGA	GGAGGCTTAC	TTGAGCCGTA	CAGCAGCCCT	CGCAAGGGG	GGAGCAGCTT	CGGCTGGCTG	CATCGGCTG	4800	
4801	GGCTGTGTC	CCAGAGTGT	GGTTCCTCT	CCAGGAGCAG	GGCCTTGTCT	TGGTGTGTC	TCAGCGGGTG	CACCAATGAC	ATGTGTACT	TSAGTGTGTC	GGCCTTCTCA	4920	
4921	AGAGCCCTC	CTGATGCTG	CTGTGACAT	GTACTTGAAG	CTGTCTCAGC	TCTTCTGCTC	TGGGATACA	AGCACAATTT	CCGCTGAGC	TGCGAGAGC	CTCAGACTCA	AGGCTCAAGC	5040
5041	GATTAAGAG	TGCGAGGCCA	CAGCCCTGGG	AGGGGGTCTC	GACTCACTAT	TACTGCAAG	GGTCTGCTCC	CGAATGTGCT	ATTTTGGTGA	CGAAGAAGTC	CTTCCGCTCC	TTGAGAGTCT	5160
5161	GTCCACATTC	GTCCACAGTA	TAGTTCGCA	CCTTGTAGAG	GGGAGAGTCC	AGTGAAGTCA	GGCCTCTGAT	GCTTCAACTC	CCGAGGGGA	CGAGATGAC	AATGTGAAG	CATCAGACT	5280
5281	GGGAGACT	TTCCTACAT	GGTTCACAT	CGATGGGCC	CAGCAGAGG	CGCTATGAG	GGCTCTGTTA	TTAGATGCT	TAACTCTCT	TATACAGCT	CAGTGGCTT	CATCAGACT	5400
5401	GTGCTTTCA	GGATGTTTT	TAAAGAAC	TCGAAAGCT	ATTTCTTAA	AAAAAAAAA	AAAAAAAAA	AAAAAAAAA	AAAAAAAAA	AAAAAAAAA	AAAAAAAAA	AAAAAAAAA	5503

## correction

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