

Mouse and Human A2BP1 sequence alignments

MNCEREQLRGNQEAAAAPDTMAQPYASAQFAPPQNGIPAEYTAP
 HPHPAPEYTGQTTVPDHTLNLYPPTQTHSEQSADTSAQTVSGTATQTDDAAPTGDGQFPQ
 TQPSENTESKSSQPKRLHVSNIFFRFRDPDLRQMFGQFGKILDVEIIFNERGSKGFGFV
 TFENSADADRAREKLHGTVEGRKIEVN NATARVMTNKKTVNPTYNGWKLNPVVGAVY
 SPDFYAGTVLLCQANQEGSSMYSGPSSLVYTSAMPGFYPYPAATAAAAYRGAHLRGRGR
 TVYNTFRAAAPPPPIPAYGGVVYQDGFYGADIYGGYAAARYAQPTPATAAAYSDSYGR
 VYAADPYHHTLAPAPTYGVGAMNAFAPLTDKTRSHADDVGLVLSSLQASIYRGGYNR
 FAPY

Isoform gamma

MNCEREQLRGNQEAAAAPDTMAQPYASAQFAPPQNGIPAEYTAPHPHPAPEYTGQTTVPDHTLNLYPPTQ
 THSEQSADTSAQTVSGTATQTDDAAPTGDGQFPQTQPSENTESKSSQPKRLHVSNIFFRFRDPDLRQMFGQFG
 KILDVEIIFNERGSKGFGFVTFENSADADRAREKLHGTVEGRKIEVN NATARVMTNKKTVNPTYNGWKL
 NPVVGAVYSPDFYAGTVLLCQANQEGSSMYSGPSSLVYTSAMPGFYPYPAATAAAAYRGAHLRGRGR
 TVYNTFRAAAPPPPIPAYGGVVYQDGFYGADIYGGYAAARYAQPTPATAAAYSDSYGRVYAADPYHHTLAPAP
 TYGVGAMNAFAPLTDKTRSHADDVGLVLSSLQASIYRGGYNRFAPY

CLUSTAL W (1.83) multiple sequence alignment for human A2BP1 isoforms1-5 and
 mouse isoform A2BP1 (equivalent to AY659954 published in paper Tissue dependent
 isoforms of mammalian Fox-1 homologs are associated with tissue-specific
 splicing activities Nakahata S and Kawamoto S in Nucliec Acid Research, 2005)

iso-1	MLASQGVLLHPYGVPMIVPAAPYLPGLIQGNQEAAAAPDTMAQPYASAQF
iso-2	MLASQGVLLHPYGVPMIVPAAPYLPGLIQGNQEAAAAPDTMAQPYASAQF
iso-3	MLASQGVLLHPYGVPMIVPAAPYLPGLIQGNQEAAAAPDTMAQPYASAQF
iso-4	MNCERE-----QLRGNQEAAAAPDTMAQPYASAQF
iso-5	MNCERE-----QLRGNQEAAAAPDTMAQPYASAQF
AY659954	MNCERE-----QLRGNQEAAAAPDTMAQPYASAQF
	* ..: ::*****

iso-1	APPQNGIPAEYTAPHPHPAPEYTGQTTVPEHTLNLYPPAQTHSEQSPADT
iso-2	APPQNGIPAEYTAPHPHPAPEYTGQTTVPEHTLNLYPPAQTHSEQSPADT
iso-3	APPQNGIPAEYTAPHPHPAPEYTGQTTVPEHTLNLYPPAQTHSEQSPADT
iso-4	APPQNGIPAEYTAPHPHPAPEYTGQTTVPEHTLNLYPPAQTHSEQSPADT
iso-5	APPQNGIPAEYTAPHPHPAPEYTGQTTVPEHTLNLYPPAQTHSEQSPADT
mus	APPQNGIPAEYTAPHPHPAPEYTGQTTVPDHTLNLYPPTQTHSEQS-ADT
	*****:*****:***** **

iso-1	SAQTVSGTATQTDDAAPTGDGQFPQTQPSENTENKSQPKRLHVSNIFFRFRD
iso-2	SAQTVSGTATQTDDAAPTGDGQFPQTQPSENTENKSQPKRLHVSNIFFRFRD
iso-3	SAQTVSGTATQTDDAAPTGDGQFPQTQPSENTENKSQPKRLHVSNIFFRFRD
iso-4	SAQTVSGTATQTDDAAPTGDGQFPQTQPSENTENKSQPKRLHVSNIFFRFRD
iso-5	SAQTVSGTATQTDDAAPTGDGQFPQTQPSENTENKSQPKRLHVSNIFFRFRD
mus	SAQTVSGTATQTDDAAPTGDGQFPQTQPSENTESKQPKRLHVSNIFFRFRD
	*****.*****

iso-1	PDLRQMFGQFGKILDVEIIFNERGSKGFGFVTFENSADADRAREKLHGTV
iso-2	PDLRQMFGQFGKILDVEIIFNERGSKGFGFVTFENSADADRAREKLHGTV
iso-3	PDLRQMFGQFGKILDVEIIFNERGSKGFGFVTFENSADADRAREKLHGTV
iso-4	PDLRQMFGQFGKILDVEIIFNERGSKGFGFVTFENSADADRAREKLHGTV
iso-5	PDLRQMFGQFGKILDVEIIFNERGSKGFGFVTFENSADADRAREKLHGTV
mus	PDLRQMFGQFGKILDVEIIFNERGSKGFGFVTFENSADADRAREKLHGTV

iso-1	VEGRKIEVN NATARVMTNKKTVNPTYNGWKLNPVVGAVYSPEFYAGTVLL
iso-2	VEGRKIEVN NATARVMTNKKTVNPTYNGWKLNPVVGAVYSPEFYAGTVLL
iso-3	VEGRKIEVN NATARVMTNKKTVNPTYNGWKLNPVVGAVYSPEFYAGTVLL
iso-4	VEGRKIEVN NATARVMTNKKTVNPTYNGWKLNPVVGAVYSPEFYAGTVLL
iso-5	VEGRKIEVN NATARVMTNKKTVNPTYNGWKLNPVVGAVYSPEFYA----
mus	VEGRKIEVN NATARVMTNKKTVNPTYNGWKLNPVVGAVYSPDFYAGTVLL
	*****:***

```
iso-1      CQANQEGSSMYSAPSSLVYTSAMPGFFPYPAATAAAAYRGAHLRGRGRTVY
iso-2      CQANQEGSSMYSAPSSLVYTSAMPGFFPYPAATAAAAYRGAHLRGRGRTVY
iso-3      CQANQEGSSMYSAPSSLVYTSAMPGFFPYPAATAAAAYRGAHLRGRGRTVY
iso-4      CQANQEGSSMYSAPSSLVYTSAMPGFFPYPAATAAAAYRGAHLRGRGRTVY
iso-5      -----VPGFFPYPAATAAAAYRGAHLRGRGRTVY
mus        CQANQEGSSMYSGPSSLVYTSAMPGFFPYPAATAAAAYRGAHLRGRGRTVY
                :*****
```

```
iso-1      NTFRAAAPPPPIPAYGGVVYQEPVYGNKLLQGGYAAAYRYAQPTPATAAAAY
iso-2      NTFRAAAPPPPIPAYGGVVYQEPVYGNKLLQGGYAAAYRYAQPTPATAAAAY
iso-3      NTFRAAAPPPPIPAYGGVVYQEPVYGNKLLQGGYAAAYRYAQPTPATAAAAY
iso-4      NTFRAAAPPPPIPAYGGVVYQDGFYGADI-YGGYAAAYRYAQPTPATAAAAY
iso-5      NTFRAAAPPPPIPAYGGVVYQDGFYGADI-YGGYAAAYRYAQPTPATAAAAY
mus        NTFRAAAPPPPIPAYGGVVYQDGFYGADI-YGGYAAAYRYAQPTPATAAAAY
                *****: .** .: *****
```

```
iso-1      SDSYGRVYAA-D-----PYHHALAPAPTYGVGAMNAF
iso-2      SDSYGRVYAA-D-----PYHHALAPAPTYGVGAM---
iso-3      SDRNQFVFVAADEISCNTSAVTDEFMLPTPTTTHLLQPPPTALV-----
iso-4      SDSYGRVYAA-D-----PYHHALAPAPTYGVGAMNAF
iso-5      SDSYGRVYAA-D-----PYHHALAPAPTYGVGAMNAF
mus        SDSYGRVYAA-D-----PYHHTLAPAPTYGVGAMNAF
                **      *:.* *          .  * * *.** *
```

```
iso-1      APLTDAKTRSHADDVGLVLSSLQASIYRGGYNRFAPY
iso-2      -----ASIYRGGYNRFAPY
iso-3      -----P
iso-4      APLTDAKTRSHADDVGLVLSSLQASIYRGGYNRFAPY
iso-5      APLTDAKTRSHADDVGLVLSSLQASIYRGGYNRFAPY
mus        APLTDAKTRSHADDVGLVLSSLQASIYRGGYNRFAPY
```

CLUSTAL W (1.83) multiple sequence alignment (comparison of human isoform 4 and mus A2BP1 gamma isoform)

```
iso-4      MNCEREQLRGNQEAAAAPDTMAQPYASAQFAPPQNGIPAEYTAPHPHPAP
AY659954    MNCEREQLRGNQEAAAAPDTMAQPYASAQFAPPQNGIPAEYTAPHPHPAP
                *****
```

```
iso-4      EYTGQTTVPEHTLNLYPQAQTHSEQSPADTSAQTVSGTATQTDDAAPTDG
mus        EYTGQTTVPDHTLNLYPPTQTHSEQS-ADTSAQTVSGTATQTDDAAPTDG
                *****:*****:***** *****
```

```
iso-4      QPQTQPSSENTENKSQPKRLHVSNIFFRFRDPDLRQMFGQFGKILDVEIIF
mus        QPQTQPSSENTESKSQPKRLHVSNIFFRFRDPDLRQMFGQFGKILDVEIIF
                *****.******
```

```
iso-4      NERGSKGFGFVTFENSADADRAREKLHGTVVEGRKIEVNNATARVMTNKK
mus        NERGSKGFGFVTFENSADADRAREKLHGTVVEGRKIEVNNATARVMTNKK
                *****
```

```
iso-4      TVNPTYNGWKLNPVVGAVYSPEFYAGTVLLCQANQEGSSMYSAPSSLVYT
mus        TVNPTYNGWKLNPVVGAVYSPDFYAGTVLLCQANQEGSSMYSGPSSLVYT
                *****:*****.******
```

```
iso-4      SAMPGFFPYPAATAAAAYRGAHLRGRGRTVYNTFRAAAPPPPIPAYGGVVY
mus        SAMPGFFPYPAATAAAAYRGAHLRGRGRTVYNTFRAAAPPPPIPAYGGVVY
                *****
```

```
iso-4      QDGFYGADIYGGYAAAYRYAQPTPATAAAAYSDSYGRVYAADPYHHALAPAP
```

mus QDGFYGADIYGGYAAARYAQPTPATAAAYSDSYGRVYAADPYHHTLAPAP
*****:*****

iso-4 TYGVGAMNAFAPLTDAKTRSHADDVGLVLSSLQASIYRGGYNRFAPY
mus TYGVGAMNAFAPLTDAKTRSHADDVGLVLSSLQASIYRGGYNRFAPY

CLUSTAL W (1.83) multiple sequence alignment for human A2BP1 isoforms1-5 and mouse isoform A2BP1 (equivalent to AY659955 published in paper Tissue dependent isoforms of mammalian Fox-1 homologs are associated with tissue-specific splicing activities Nakahata S and Kawamoto S in Nucliec Acid Research, 2005)

AY659955 MNCERE-----QLRGNQEAAAAPDTMAQPYASAQF
iso-1 MLASQGVLHPYGVPMIVPAAPYLPGLIQGNQEAAAAPDTMAQPYASAQF
iso-2 MLASQGVLHPYGVPMIVPAAPYLPGLIQGNQEAAAAPDTMAQPYASAQF
iso-3 MLASQGVLHPYGVPMIVPAAPYLPGLIQGNQEAAAAPDTMAQPYASAQF
iso-4 MNCERE-----QLRGNQEAAAAPDTMAQPYASAQF
iso-5 MNCERE-----QLRGNQEAAAAPDTMAQPYASAQF
* ..: :*****

AY659955 APPQNGIPAEYTAPHPHPAPEYTGQTTVPDHTLNLYPPTQTHSEQS-ADT
iso-1 APPQNGIPAEYTAPHPHPAPEYTGQTTVPEHTLNLYPPAQTHSEQSPADT
iso-2 APPQNGIPAEYTAPHPHPAPEYTGQTTVPEHTLNLYPPAQTHSEQSPADT
iso-3 APPQNGIPAEYTAPHPHPAPEYTGQTTVPEHTLNLYPPAQTHSEQSPADT
iso-4 APPQNGIPAEYTAPHPHPAPEYTGQTTVPEHTLNLYPPAQTHSEQSPADT
iso-5 APPQNGIPAEYTAPHPHPAPEYTGQTTVPEHTLNLYPPAQTHSEQSPADT
*****:*****:***** **

AY659955 SAQTVSGTATQTDDAAPTGDGPQTQPSSENTESKSPKRLHVSNIFFRFRD
iso-1 SAQTVSGTATQTDDAAPTGDGPQTQPSSENTENKSPKRLHVSNIFFRFRD
iso-2 SAQTVSGTATQTDDAAPTGDGPQTQPSSENTENKSPKRLHVSNIFFRFRD
iso-3 SAQTVSGTATQTDDAAPTGDGPQTQPSSENTENKSPKRLHVSNIFFRFRD
iso-4 SAQTVSGTATQTDDAAPTGDGPQTQPSSENTENKSPKRLHVSNIFFRFRD
iso-5 SAQTVSGTATQTDDAAPTGDGPQTQPSSENTENKSPKRLHVSNIFFRFRD
*****.*****

AY659955 PDLRQMFGQFGKILDVEIIFNERGSKGFGFVTFENSADADRAREKLHGTV
iso-1 PDLRQMFGQFGKILDVEIIFNERGSKGFGFVTFENSADADRAREKLHGTV
iso-2 PDLRQMFGQFGKILDVEIIFNERGSKGFGFVTFENSADADRAREKLHGTV
iso-3 PDLRQMFGQFGKILDVEIIFNERGSKGFGFVTFENSADADRAREKLHGTV
iso-4 PDLRQMFGQFGKILDVEIIFNERGSKGFGFVTFENSADADRAREKLHGTV
iso-5 PDLRQMFGQFGKILDVEIIFNERGSKGFGFVTFENSADADRAREKLHGTV

AY659955 VEGRKIEVNNATARVMTNKKTVNPTYNGWKLNPVVGAVYSPDFYAGTVLL
iso-1 VEGRKIEVNNATARVMTNKKTVNPTYNGWKLNPVVGAVYSPEFYAGTVLL
iso-2 VEGRKIEVNNATARVMTNKKTVNPTYNGWKLNPVVGAVYSPEFYAGTVLL
iso-3 VEGRKIEVNNATARVMTNKKTVNPTYNGWKLNPVVGAVYSPEFYAGTVLL
iso-4 VEGRKIEVNNATARVMTNKKTVNPTYNGWKLNPVVGAVYSPEFYAGTVLL
iso-5 VEGRKIEVNNATARVMTNKKTVNPTYNGWKLNPVVGAVYSPEFYA-----
*****:***

AY659955 CQANQEGSSMYSGPSSLVYTSAMPGFPPYPAATAAAAYRGAHLRGRGRTVY
iso-1 CQANQEGSSMYSA PSSLVYTSAMPGFPPYPAATAAAAYRGAHLRGRGRTVY
iso-2 CQANQEGSSMYSA PSSLVYTSAMPGFPPYPAATAAAAYRGAHLRGRGRTVY
iso-3 CQANQEGSSMYSA PSSLVYTSAMPGFPPYPAATAAAAYRGAHLRGRGRTVY
iso-4 CQANQEGSSMYSA PSSLVYTSAMPGFPPYPAATAAAAYRGAHLRGRGRTVY
iso-5 -----VPGFPYPAATAAAAYRGAHLRGRGRTVY
:*****

AY659955 NTFRAAAPPPPIPAYGGVYQDGFYGADI-YGGYAAARYAQPTPATAAAY

iso-1 NTFRAAAPPPPIPAYGGVVYQEPVYGNKLLQGGYAAARYAQPTPATAAAAY
iso-2 NTFRAAAPPPPIPAYGGVVYQEPVYGNKLLQGGYAAARYAQPTPATAAAAY
iso-3 NTFRAAAPPPPIPAYGGVVYQEPVYGNKLLQGGYAAARYAQPTPATAAAAY
iso-4 NTFRAAAPPPPIPAYGGVVYQDGFYGADI-YGGYAAARYAQPTPATAAAAY
iso-5 NTFRAAAPPPPIPAYGGVVYQDGFYGADI-YGGYAAARYAQPTPATAAAAY
*****: .** .: *****

AY659955 SDRNQFVFVATDEISCNTSAVTDEFMLPTPTTHLLQPPPTALV-----
iso-1 SDSYGRVYAAD-----PYHHALAPAPTYGVGAMNAF
iso-2 SDSYGRVYAAD-----PYHHALAPAPTYGVGAMNAF
iso-3 SDRNQFVFVAADEISCNTSAVTDEFMLPTPTTHLLQPPPTALV-----
iso-4 SDSYGRVYAAD-----PYHHALAPAPTYGVGAMNAF
iso-5 SDSYGRVYAAD-----PYHHALAPAPTYGVGAMNAF
** *:.*

AY659955 -----P
iso-1 APLTDAKTRSHADDVGLVLSSLQASIYRGGYNRFAPY
iso-2 -----ASIYRGGYNRFAPY
iso-3 -----P
iso-4 APLTDAKTRSHADDVGLVLSSLQASIYRGGYNRFAPY
iso-5 APLTDAKTRSHADDVGLVLSSLQASIYRGGYNRFAPY

Mus musculus ataxin-2 binding protein 1 variant 2 (A2bp1)

AY659955
MNCEREQLRGNQEAAAAPDTMAQPYASAQFAPPQNGIPAEYTAPHPHPAPEYTGQTTVPDHTLNLYPPTQT
HSEQSADTSAQTVSGTATQTDDAAPTQDGPQTQPSSENTESKSPKRLHVSNIPIFRFRDPDLRQMFQFGKI
LDVEIIFNERGSKGFGFVTFENSADADRAREKLHGTVEGRKIEVNNATARVMTNKKTVNPNYTNQWKLNPV
VGAVYSPDFYAGTVLLCQANQEGSSMYSGPSSLVYTSAMPGFYPAPATAAAAYRGAHLRGRGRTVYNTFRA
AAPPPPIPAYGGVVYQDGFYGADIYGGYAAARYAQPTPATAAAAYSDRNQFVFVATDEISCNTSAVTDEFML
PTPTTHLLQPPPTALVP

Mus musculus ataxin-2 binding protein 1 variant 1 (A2bp1)

AY659954
MNCEREQLRGNQEAAAAPDTMAQPYASAQFAPPQNGIPAEYTAP
HPHPAPEYTGQTTVPDHTLNLYPPTQTHSEQSADTSAQTVSGTATQTDDAAPTQDGPQTQPSSENTESKSPKRLHVSNIPIFRFRDPDLRQMFQFGKILDVEIIFNERGSKGFGFV
TFENSADADRAREKLHGTVEGRKIEVNNATARVMTNKKTVNPNYTNQWKLNPVVGAVY
SPDFYAGTVLLCQANQEGSSMYSGPSSLVYTSAMPGFYPAPATAAAAYRGAHLRGRGR
TVYNTFRAAAPPPPIPAYGGVVYQDGFYGADIYGGYAAARYAQPTPATAAAAYSDSYGR
VYAADPYHHTLAPAPTYGVGAMNAFAPLTDKTRSHADDVGLVLSSLQASIYRGGYNR
FAPY

Alpha (variant 1; NM_183188.2; NP_899011.2; updated Nov'11)
MLASQGVLLHSYGVPMIVPAAPYFPGMLQGNQEAAAAPDTMAQPYASAQFAPPQNGIPAEYTAPHPHPAP
EYTGQTTVPDHTLNLYPPTQTHSEQSADTSAQTVSGTATQTDDAAPTQDGPQTQPSSENTESKSPKRLHV
SNIPFRFRDPDLRQMFQFGKILDVEIIFNERGSKGFGFVTFENSADADRAREKLHGTVEGRKIEVNN
TARVMTNKKTVNPNYTNQWKLNPVVGAVYSPDFYAGTVLLCQANQEGSSMYSGPSSLVYTSAMPGFYPAA
TAAAYRGAHLRGRGRTVYNTFRAAAPPPPIPAYGGVVYQEPVYGNKLLQGGYAAARYAQPTPATAAAAY
DSYGRVYAADPYHHTLAPAPTYGVGAMNAFAPLTDKTRSHADDVGLVLSSLQASIYRGGYNRFAPY

Gamma (variant 2; NM_021477.5; NP_067452.2; updated Nov'11)

MNCEREQLRGNQEAAAAPDTMAQPYASAQFAPPQNGIPAEYTAPHPHPAPEYTGQTTVPDHTLNLYPPTQ
THSEQSADTSAQTVSGTATQTDDAAPT DGQPQTQPSENTESKSPKRLHVSNI PFRFRDPDLRQMFGQFG
KILDVEIIFNERGSKGFGFVTFENSADADRAREKLHGTVEGRKIEVNNATARVMTNKKTVNPTYNGWKL
NPVVGAVYSPDFYAGTVLLCQANQEGSSMYSGPSSLVYTSAMPGFPPYPAATAAAAYRGAHLRGRGRTVYN
TFRAAAPPPPIPAYGGVVYQDGFYGADIYGGYAAARYAQPTPATAAAYSDSYGRVYAADPYHHTLAPAPT
YGVGAMNAFAPLTDKTRSHADDVGLVLSSLQASIYRGGYNRFAPY

CLUSTAL W (1.83) multiple sequence alignment

```
AY659954      MNCERE-----QLRGNQEAAAAPDTMAQPYASAQF
AY659955      MNCERE-----QLRGNQEAAAAPDTMAQPYASAQF
alpha         MLASQGVLLHSYGVP MIVPAAPYFPGLMQGNQEAAAAPDTMAQPYASAQF
gamma         MNCERE-----QLRGNQEAAAAPDTMAQPYASAQF
               *  . . :                               . : *****

AY659954      APPQNGIPAEYTAPHPHPAPEYTGQTTVPDHTLNLYPPTQTHSEQSADTS
AY659955      APPQNGIPAEYTAPHPHPAPEYTGQTTVPDHTLNLYPPTQTHSEQSADTS
alpha         APPQNGIPAEYTAPHPHPAPEYTGQTTVPDHTLNLYPPTQTHSEQSADTS
gamma         APPQNGIPAEYTAPHPHPAPEYTGQTTVPDHTLNLYPPTQTHSEQSADTS
               *****

AY659954      AQTVSGTATQTDDAAPT DGQPQTQPSENTESKSPKRLHVSNI PFRFRDP
AY659955      AQTVSGTATQTDDAAPT DGQPQTQPSENTESKSPKRLHVSNI PFRFRDP
alpha         AQTVSGTATQTDDAAPT DGQPQTQPSENTESKSPKRLHVSNI PFRFRDP
gamma         AQTVSGTATQTDDAAPT DGQPQTQPSENTESKSPKRLHVSNI PFRFRDP
               *****

AY659954      DLRQMFGQFGKILDVEIIFNERGSKGFGFVTFENSADADRAREKLHGTVV
AY659955      DLRQMFGQFGKILDVEIIFNERGSKGFGFVTFENSADADRAREKLHGTVV
alpha         DLRQMFGQFGKILDVEIIFNERGSKGFGFVTFENSADADRAREKLHGTVV
gamma         DLRQMFGQFGKILDVEIIFNERGSKGFGFVTFENSADADRAREKLHGTVV
               *****

AY659954      EGRKIEVNNATARVMTNKKTVNPTYNGWKLNPVVGAVYSPDFYAGTVLLC
AY659955      EGRKIEVNNATARVMTNKKTVNPTYNGWKLNPVVGAVYSPDFYAGTVLLC
alpha         EGRKIEVNNATARVMTNKKTVNPTYNGWKLNPVVGAVYSPDFYAGTVLLC
gamma         EGRKIEVNNATARVMTNKKTVNPTYNGWKLNPVVGAVYSPDFYAGTVLLC
               *****

AY659954      QANQEGSSMYSGPSSLVYTSAMPGFPPYPAATAAAAYRGAHLRGRGRTVYN
AY659955      QANQEGSSMYSGPSSLVYTSAMPGFPPYPAATAAAAYRGAHLRGRGRTVYN
alpha         QANQEGSSMYSGPSSLVYTSAMPGFPPYPAATAAAAYRGAHLRGRGRTVYN
gamma         QANQEGSSMYSGPSSLVYTSAMPGFPPYPAATAAAAYRGAHLRGRGRTVYN
               *****

AY659954      TFRAAAPPPPIPAYGGVVYQDGFYGADI-YGGYAAARYAQPTPATAAAYS
AY659955      TFRAAAPPPPIPAYGGVVYQDGFYGADI-YGGYAAARYAQPTPATAAAYS
alpha         TFRAAAPPPPIPAYGGVVYQEFVYGNKLLQGGYAAARYAQPTPATAAAYS
gamma         TFRAAAPPPPIPAYGGVVYQDGFYGADI-YGGYAAARYAQPTPATAAAYS
               *****: . ** . : *****

AY659954      DSYGRVYAAD-----PYHHTLAPPTYGVGAMNAFA
AY659955      DRNQFVFVATDEISCN TSAVTDEFMLPTPTTHLLQPPPTALV-----
alpha         DSYGRVYAAD-----PYHHTLAPPTYGVGAMNAFA
gamma         DSYGRVYAAD-----PYHHTLAPPTYGVGAMNAFA
               *   * : *                               .   * * * . * *   *

AY659954      PLTDAKTRSHADDVGLVLSSLQASIYRGGYNRFAPY
AY659955      -----P
alpha         PLTDAKTRSHADDVGLVLSSLQASIYRGGYNRFAPY
gamma         PLTDAKTRSHADDVGLVLSSLQASIYRGGYNRFAPY
```

CLUSTAL W (1.83) multiple sequence alignment (between AY659954 and gamma)

```

AY659954      MNCEREQLRGNQEAAAAPDTMAQPYASAQFAPPQNGIPAEYTAPHPHPAP
gamma         MNCEREQLRGNQEAAAAPDTMAQPYASAQFAPPQNGIPAEYTAPHPHPAP
              *****

AY659954      EYTGQTTVPDHTLNLYPPTQTHSEQSADTSAQTVSGTATQTDDAAPT DGQ
gamma         EYTGQTTVPDHTLNLYPPTQTHSEQSADTSAQTVSGTATQTDDAAPT DGQ
              *****

AY659954      PQTQPSSENTESK SQPKRLHVSNI PFRFRDPDLRQMFGQFGKILDVEI IFN
gamma         PQTQPSSENTESK SQPKRLHVSNI PFRFRDPDLRQMFGQFGKILDVEI IFN
              *****

AY659954      ERGSKGFGFVTFENSADADRAREKLGHTVVEGRKIEVNNATARVMTNKK T
gamma         ERGSKGFGFVTFENSADADRAREKLGHTVVEGRKIEVNNATARVMTNKK T
              *****

AY659954      VNPYTNGWKLNPVVGAVYSPDFYAGTVLLCQANQEGSSMYSGPSSLVYTS
gamma         VNPYTNGWKLNPVVGAVYSPDFYAGTVLLCQANQEGSSMYSGPSSLVYTS
              *****

AY659954      AMPGFPYPAATAAAAYRGAHLRGRGRTVYNTFRAAAPPPPIPAYGGVVYQ
gamma         AMPGFPYPAATAAAAYRGAHLRGRGRTVYNTFRAAAPPPPIPAYGGVVYQ
              *****

AY659954      DGFYGADIYGGYAAARYAQPTPATAAAYSDSYGRVYAADPYHHTLAPAPT
gamma         DGFYGADIYGGYAAARYAQPTPATAAAYSDSYGRVYAADPYHHTLAPAPT
              *****

AY659954      YGVGAMNAFAPLTD AKTRSHADDVGLVLSSLQASIYRGGYNRFAPY
gamma         YGVGAMNAFAPLTD AKTRSHADDVGLVLSSLQASIYRGGYNRFAPY
              *****

```

Comparison between alpha and gamma
>lcl|36203 unnamed protein product
Length=396

Score = 714 bits (1843), Expect = 0.0, Method: Compositional matrix adjust.
Identities = 380/390 (97%), Positives = 384/390 (98%), Gaps = 1/390 (0%)

Query	28	MQGNQEAAAAPDTMAQPYASAQFAPPQNGIPAEYTAPHPHPAPEYTGQTTVPDHTLNLYP	87
		++GNQEAAAAPDTMAQPYASAQFAPPQNGIPAEYTAPHPHPAPEYTGQTTVPDHTLNLYP	
Sbjct	8	LRGNQEAAAAPDTMAQPYASAQFAPPQNGIPAEYTAPHPHPAPEYTGQTTVPDHTLNLYP	67
Query	88	PTQTHSEQSADTSAQTVSGTATQTDDAAPT DGQPQTQPSSENTESK SQPKRLHVSNI PFRF	147
		PTQTHSEQSADTSAQTVSGTATQTDDAAPT DGQPQTQPSSENTESK SQPKRLHVSNI PFRF	
Sbjct	68	PTQTHSEQSADTSAQTVSGTATQTDDAAPT DGQPQTQPSSENTESK SQPKRLHVSNI PFRF	127
Query	148	RDPDLRQMFGQFGKILDVEI IFNERGSKGFGFVTFENSADADRAREKLGHTVVEGRKIEV	207
		RDPDLRQMFGQFGKILDVEI IFNERGSKGFGFVTFENSADADRAREKLGHTVVEGRKIEV	
Sbjct	128	RDPDLRQMFGQFGKILDVEI IFNERGSKGFGFVTFENSADADRAREKLGHTVVEGRKIEV	187
Query	208	NNATARVMTNKKTVNPYTNGWKLNPVVGAVYSPDFYAGTVLLCQANQEGSSMYSGPSSLV	267
		NNATARVMTNKKTVNPYTNGWKLNPVVGAVYSPDFYAGTVLLCQANQEGSSMYSGPSSLV	
Sbjct	188	NNATARVMTNKKTVNPYTNGWKLNPVVGAVYSPDFYAGTVLLCQANQEGSSMYSGPSSLV	247

Query	268	YTSAMPGFPYPAATAAAAYRGAHLRGRGRTVYNTFRAAAPPPPIPAYGGVVYQEPVYGNK	327
		YTSAMPGFPYPAATAAAAYRGAHLRGRGRTVYNTFRAAAPPPPIPAYGGVVYQ+ YG	
Sbjct	248	YTSAMPGFPYPAATAAAAYRGAHLRGRGRTVYNTFRAAAPPPPIPAYGGVVYQDGFYGAD	307
Query	328	LLQGGYAAARYAQPTPATAAAAYSDSYGRVYAADPYHHTLAPAPTYGVGAMNAFAPLTDAK	387
		+ GGYAAARYAQPTPATAAAAYSDSYGRVYAADPYHHTLAPAPTYGVGAMNAFAPLTDAK	
Sbjct	308	IY-GGYAAARYAQPTPATAAAAYSDSYGRVYAADPYHHTLAPAPTYGVGAMNAFAPLTDAK	366
Query	388	TRSHADDVGLVLSSLQASIYRGGYNRFAPY	417
		TRSHADDVGLVLSSLQASIYRGGYNRFAPY	
Sbjct	367	TRSHADDVGLVLSSLQASIYRGGYNRFAPY	396

A2BP1 isoform sequence alignments

Human FOX-1 (RBFox1 chr 16) isoforms

RNA binding protein fox-1 homolog 1 isoform 5 [Homo sapiens]

NCBI Reference Sequence: NP_001135805.1

RNA binding protein fox-1 homolog 1 isoform 5 [Homo sapiens]

Transcript Variant: This variant (5) differs in the 5' UTR and has multiple coding region differences compared to variant 1. The resulting isoform (5) has a shorter and distinct N-terminus and other internal differences compared to isoform 1.

MNCEREQLRGNQEAAAAPDTMAQPYASAQFAPPQNGIPAEYTAPHPHPAPEYTGQTTVPEHTLNLYPPAQ
THSEQSPADTSAQTVSGTATQTDDAAPT DGQPQTQPSSENTENKSQPKRLHVSNI PFRFRDPDLRQMFGQF
GKILDVEIIFNERGSKGFGFVTFENSADADRAREKLHGT VVEGRKIEVNNATARVM **TNKKTVNPTYNGWK**
LNPVVGAVYSPEFYAVPGFPYPAATAAAAYRGAHLRGRGRTVYNTFRAAAPPPPIPAYGGVVYQDGFYGA
DIYGGYAA YRYAQPTPATAAAYSDSYGRVYAADPYHHALAPPTYGVGAMNAFAPLTD AKTRSHADDVGL
VLSSLQASIYRGGYNRFAPY

Comparison of protein sequences between isoform 5 and published

>lcl|63771 unnamed protein product

Length=377

Score = 642 bits (1656), Expect = 0.0, Method: Compositional matrix adjust.
Identities = 349/377 (93%), Positives = 350/377 (93%), Gaps = 27/377 (7%)

Query	21	MAQPYASAQFAPPQNGIPAEYTAPHPHPAPEYTGQTTVPEHTLNLYPPAQTHSEQSPADT	80
		MAQPYASAQFAPPQNGIPAEYTAPHPHPAPEYTGQTTVPEHTLNLYPPAQTHSEQSPADT	
Sbjct	1	MAQPYASAQFAPPQNGIPAEYTAPHPHPAPEYTGQTTVPEHTLNLYPPAQTHSEQSPADT	60
Query	81	SAQTVSGTATQTDDAAPT DGQPQTQPSSENTENKSQPKRLHVSNI PFRFRDPDLRQMFGQF	140
		SAQTVSGTATQTDDAAPT DGQPQTQPSSENTENKSQPKRLHVSNI PFRFRDPDLRQMFGQF	
Sbjct	61	SAQTVSGTATQTDDAAPT DGQPQTQPSSENTENKSQPKRLHVSNI PFRFRDPDLRQMFGQF	120
Query	141	GKILDVEIIFNERGSKGFGFVTFENSADADRAREKLHGT VVEGRKIEVNNATARVMTNKK	200
		GKILDVEIIFNERGSKGFGFVTFENSADADRAREKLHGT VVEGRKIEVNNATARVMTNKK	
Sbjct	121	GKILDVEIIFNERGSKGFGFVTFENSADADRAREKLHGT VVEGRKIEVNNATARVMTNKK	180
Query	201	TVNPTYNGWKLNPNVVGAVYSPEFY-----AVPGFPYPA	233
		TVNPTYNGWKLNPNVVGAVYSPEFY A+PGFPYPA	
Sbjct	181	TVNPTYNGWKLNPNVVGAVYSPEFYAGTVLLCQANQEGSSMYSAPSSLVYTSAMPGFPYPA	240
Query	234	ATAAAAYRGAHLRGRGRTVYNTFRAAAPPPPIPAYGGVVYQDGFY GADIYGGYAA YRYAQ	293
		ATAAAAYRGAHLRGRGRTVYNTFRAAAPPPPIPAYGGVVYQDGFY GADIYGGYAA YRYAQ	
Sbjct	241	ATAAAAYRGAHLRGRGRTVYNTFRAAAPPPPIPAYGGVVYQDGFY GADIYGGYAA YRYAQ	300
Query	294	PTPATAAAYS DSYGRVYAADPYHHALAPPTYGVGAMNAFAPLTD AKTRSHADDVGLVLS	353
		PTPATAAAYS DSYGRVYAADPYHHALAPPTYGVGAMNAFAPLTD AKTRSHADDVGLVLS	
Sbjct	301	PTPATAAAYS DSYGRVYAADPYHHALAPPTYGVGAMNAFAPLTD AKTRSHADDVGLVLS	360
Query	354	SLQASIYRGGYNRFAPY	370
		SLQASIYRGGYNRFAPY	
Sbjct	361	SLQASIYRGGYNRFAPY	377

RNA binding protein fox-1 homolog 1 isoform 4 [Homo sapiens]

NCBI Reference Sequence: NP_001135806.1

RNA binding protein fox-1 homolog 1 isoform 4 [Homo sapiens]

Transcript Variant: This variant (6) differs in the 5' UTR and lacks an in-frame section of the 3' coding region, compared to variant 1. This results in a shorter isoform (4) with an alternate N-terminus compared to isoform 1. Variants 4 and 6 both encode isoform 4.

MNCEREQLRGNQEAAAAPDTMAQPYASAQFAPPQNGIPA EYTAPHPHPAPEYTGQTTVPEHTLNLYPPAQ
 THSEQSPADTSAQTVSGTATQTDDAAPT DGQPQTQPSSENTENKSQPKRLHVSNI PFRFRDPDLRQMFGQF
 GKILDVEIIFNERGSKGFGFVTFENSADADRAREK LHGTVVEGRKIEVNNATARVM **TNKKTVNPTYNGWK**
 LNPVVGAVYSPEFYAGTVLLCQANQEGSSMYSAPSSLVYTSAMPGFPYPAATAAAAYRGAHLRGRGR TVY
 NTFRAAAPPPPIPAYGGVVYQDGFY GADIYGGYAA YRYAQPTPATAAAYSDSYGRVYAADPYH HALAPAP
 TYGVGAMNAFAPLTD AKTRSHADDVGLV LSSLQASIYRGGYNRFAPY

Comparison of protein sequences between isoform 4 (variant 6) and published

>lcl|5495 unnamed protein product
 Length=377

Score = 713 bits (1840), Expect = 0.0, Method: Compositional matrix adjust.
 Identities = 377/377 (100%), Positives = 377/377 (100%), Gaps = 0/377 (0%)

Query	21	MAQPYASAQFAPPQNGIPA EYTAPHPHPAPEYTGQTTVPEHTLNLYPPAQTHSEQSPADT	80
		MAQPYASAQFAPPQNGIPA EYTAPHPHPAPEYTGQTTVPEHTLNLYPPAQTHSEQSPADT	
Sbjct	1	MAQPYASAQFAPPQNGIPA EYTAPHPHPAPEYTGQTTVPEHTLNLYPPAQTHSEQSPADT	60
Query	81	SAQTVSGTATQTDDAAPT DGQPQTQPSSENTENKSQPKRLHVSNI PFRFRDPDLRQMFGQF	140
		SAQTVSGTATQTDDAAPT DGQPQTQPSSENTENKSQPKRLHVSNI PFRFRDPDLRQMFGQF	
Sbjct	61	SAQTVSGTATQTDDAAPT DGQPQTQPSSENTENKSQPKRLHVSNI PFRFRDPDLRQMFGQF	120
Query	141	GKILDVEIIFNERGSKGFGFVTFENSADADRAREK LHGTVVEGRKIEVNNATARVMTNKK	200
		GKILDVEIIFNERGSKGFGFVTFENSADADRAREK LHGTVVEGRKIEVNNATARVMTNKK	
Sbjct	121	GKILDVEIIFNERGSKGFGFVTFENSADADRAREK LHGTVVEGRKIEVNNATARVMTNKK	180
Query	201	TVNPTYNGWKLNPNVVGAVYSPEFYAGTVLLCQANQEGSSMYSAPSSLVYTSAMPGFPYPA	260
		TVNPTYNGWKLNPNVVGAVYSPEFYAGTVLLCQANQEGSSMYSAPSSLVYTSAMPGFPYPA	
Sbjct	181	TVNPTYNGWKLNPNVVGAVYSPEFYAGTVLLCQANQEGSSMYSAPSSLVYTSAMPGFPYPA	240
Query	261	ATAAAAYRGAHLRGRGR TVYNTFRAAAPPPPIPAYGGVVYQDGFY GADIYGGYAA YRYAQ	320
		ATAAAAYRGAHLRGRGR TVYNTFRAAAPPPPIPAYGGVVYQDGFY GADIYGGYAA YRYAQ	
Sbjct	241	ATAAAAYRGAHLRGRGR TVYNTFRAAAPPPPIPAYGGVVYQDGFY GADIYGGYAA YRYAQ	300
Query	321	PTPATAAAYSDSYGRVYAADPYH HALAPAPTYGVGAMNAFAPLTD AKTRSHADDVGLVLS	380
		PTPATAAAYSDSYGRVYAADPYH HALAPAPTYGVGAMNAFAPLTD AKTRSHADDVGLVLS	
Sbjct	301	PTPATAAAYSDSYGRVYAADPYH HALAPAPTYGVGAMNAFAPLTD AKTRSHADDVGLVLS	360
Query	381	SLQASIYRGGYNRFAPY	397
		SLQASIYRGGYNRFAPY	
Sbjct	361	SLQASIYRGGYNRFAPY	377

RNA binding protein fox-1 homolog 1 isoform 4 [Homo sapiens]

NCBI Reference Sequence: NP_061193.2

RNA binding protein fox-1 homolog 1 isoform 4 [Homo sapiens]

Transcript Variant: This variant (4) differs in the 5' UTR and lacks an in-frame section of the 3' coding region, compared to variant 1. This results in a shorter isoform (4) with an alternate N-terminus compared to isoform 1. Variants 4 and 6 both encode isoform 4.

MNCEREQLRGNQEAAAAPDTMAQPYASAQFAPPQNGIPA EYTAPHPHPAPEYTGQTTVPEHTLNLYPPAQ
 THSEQSPADTSAQTVSGTATQTDDAAPT DGQPQTQPSSENTENKSQPKRLHVSNI PFRFRDPDLRQMFGQF
 GKILDVEIIFNERGSKGFGFVTFENSADADRAREK LHGTVVEGRKIEVNNATARVM **TNKKTVNPTYNGWK**
 LNPVVGAVYSPEFYAGTVLLCQANQEGSSMYSAPSSLVYTSAMPGFPYPAATAAAAYRGAHLRGRGR TVY
 NTFRAAAPPPPIPAYGGVVYQDGFY GADIYGGYAA YRYAQPTPATAAAYSDSYGRVYAADPYH HALAPAP
 TYGVGAMNAFAPLTD AKTRSHADDVGLV LSSLQASIYRGGYNRFAPY

Comparison between isoform 4 variant 6 and variant 4

>lcl|47453 unnamed protein product
 Length=397

Score = 811 bits (2095), Expect = 0.0, Method: Compositional matrix adjust.
Identities = 397/397 (100%), Positives = 397/397 (100%), Gaps = 0/397 (0%)

Query	1	MNCEREQLRGNQEAAAAPDTMAQPYASAQFAPPQNGIPAEYTAPHPHPAPEYTGQTTVPE	60
		MNCEREQLRGNQEAAAAPDTMAQPYASAQFAPPQNGIPAEYTAPHPHPAPEYTGQTTVPE	
Sbjct	1	MNCEREQLRGNQEAAAAPDTMAQPYASAQFAPPQNGIPAEYTAPHPHPAPEYTGQTTVPE	60
Query	61	HTLNLYPPAQTHSEQSPADTSAQTVSGTATQTDDAAPTGGQPQTQPSSENTENKSQPKRLH	120
		HTLNLYPPAQTHSEQSPADTSAQTVSGTATQTDDAAPTGGQPQTQPSSENTENKSQPKRLH	
Sbjct	61	HTLNLYPPAQTHSEQSPADTSAQTVSGTATQTDDAAPTGGQPQTQPSSENTENKSQPKRLH	120
Query	121	VSNIPFRFRDPDLRQMFGQFGKILDVEIIFNERGSKGFGFVTFENSADADRAREKLHGTV	180
		VSNIPFRFRDPDLRQMFGQFGKILDVEIIFNERGSKGFGFVTFENSADADRAREKLHGTV	
Sbjct	121	VSNIPFRFRDPDLRQMFGQFGKILDVEIIFNERGSKGFGFVTFENSADADRAREKLHGTV	180
Query	181	VEGRKIEVNNATARVMTNKKTVNPTYNGWKLNPPVVGAVYSPEFYAGTVLLCQANQEGSSM	240
		VEGRKIEVNNATARVMTNKKTVNPTYNGWKLNPPVVGAVYSPEFYAGTVLLCQANQEGSSM	
Sbjct	181	VEGRKIEVNNATARVMTNKKTVNPTYNGWKLNPPVVGAVYSPEFYAGTVLLCQANQEGSSM	240
Query	241	YSAPSSLVYTSAMPGFPYPAATAAAAYRGAHLRGRGRTVYNTFRAAAPPPIPAYGGVVY	300
		YSAPSSLVYTSAMPGFPYPAATAAAAYRGAHLRGRGRTVYNTFRAAAPPPIPAYGGVVY	
Sbjct	241	YSAPSSLVYTSAMPGFPYPAATAAAAYRGAHLRGRGRTVYNTFRAAAPPPIPAYGGVVY	300
Query	301	QDGYGADIYGGYAAARYAQPTPATAAAYSDSYGRVYAADPYHHALAPAPTYGVGAMNAF	360
		QDGYGADIYGGYAAARYAQPTPATAAAYSDSYGRVYAADPYHHALAPAPTYGVGAMNAF	
Sbjct	301	QDGYGADIYGGYAAARYAQPTPATAAAYSDSYGRVYAADPYHHALAPAPTYGVGAMNAF	360
Query	361	APLTDAKTRSHADDVGLVLSSLQASIYRGGYNRFAPY	397
		APLTDAKTRSHADDVGLVLSSLQASIYRGGYNRFAPY	
Sbjct	361	APLTDAKTRSHADDVGLVLSSLQASIYRGGYNRFAPY	397

Comparison of protein sequences between isoform 4 (variant 4) and published

>lcl|58971 unnamed protein product
Length=377

Score = 713 bits (1840), Expect = 0.0, Method: Compositional matrix adjust.
Identities = 377/377 (100%), Positives = 377/377 (100%), Gaps = 0/377 (0%)

Query	21	MAQPYASAQFAPPQNGIPAEYTAPHPHPAPEYTGQTTVPEHTLNLYPPAQTHSEQSPADT	80
		MAQPYASAQFAPPQNGIPAEYTAPHPHPAPEYTGQTTVPEHTLNLYPPAQTHSEQSPADT	
Sbjct	1	MAQPYASAQFAPPQNGIPAEYTAPHPHPAPEYTGQTTVPEHTLNLYPPAQTHSEQSPADT	60
Query	81	SAQTVSGTATQTDDAAPTGGQPQTQPSSENTENKSQPKRLHVSNIPIFRFRDPDLRQMFGQF	140
		SAQTVSGTATQTDDAAPTGGQPQTQPSSENTENKSQPKRLHVSNIPIFRFRDPDLRQMFGQF	
Sbjct	61	SAQTVSGTATQTDDAAPTGGQPQTQPSSENTENKSQPKRLHVSNIPIFRFRDPDLRQMFGQF	120
Query	141	GKILDVEIIFNERGSKGFGFVTFENSADADRAREKLHGTVVEGRKIEVNNATARVMTNKK	200
		GKILDVEIIFNERGSKGFGFVTFENSADADRAREKLHGTVVEGRKIEVNNATARVMTNKK	
Sbjct	121	GKILDVEIIFNERGSKGFGFVTFENSADADRAREKLHGTVVEGRKIEVNNATARVMTNKK	180
Query	201	TVNPTYNGWKLNPPVVGAVYSPEFYAGTVLLCQANQEGSSMYSAPSSLVYTSAMPGFPYPA	260
		TVNPTYNGWKLNPPVVGAVYSPEFYAGTVLLCQANQEGSSMYSAPSSLVYTSAMPGFPYPA	
Sbjct	181	TVNPTYNGWKLNPPVVGAVYSPEFYAGTVLLCQANQEGSSMYSAPSSLVYTSAMPGFPYPA	240
Query	261	ATAAAAYRGAHLRGRGRTVYNTFRAAAPPPIPAYGGVVYQDGYGADIYGGYAAARYAQ	320
		ATAAAAYRGAHLRGRGRTVYNTFRAAAPPPIPAYGGVVYQDGYGADIYGGYAAARYAQ	
Sbjct	241	ATAAAAYRGAHLRGRGRTVYNTFRAAAPPPIPAYGGVVYQDGYGADIYGGYAAARYAQ	300
Query	321	PTPATAAAYSDSYGRVYAADPYHHALAPAPTYGVGAMNAFAPLTDAKTRSHADDVGLVLS	380
		PTPATAAAYSDSYGRVYAADPYHHALAPAPTYGVGAMNAFAPLTDAKTRSHADDVGLVLS	
Sbjct	301	PTPATAAAYSDSYGRVYAADPYHHALAPAPTYGVGAMNAFAPLTDAKTRSHADDVGLVLS	360

```

Query   381  SLQASIYRGGYNRFAPY   397
        SLQASIYRGGYNRFAPY
Sbjct   361  SLQASIYRGGYNRFAPY   377

```

RNA binding protein fox-1 homolog 1 isoform 1 [Homo sapiens]

NCBI Reference Sequence: NP_665898.1

RNA binding protein fox-1 homolog 1 isoform 1 [Homo sapiens]

Transcript Variant: This variant (1), also known as gamma, encodes the longest isoform (1).

```

MLASQGVLLHPYGVPMIVPAAPYLPGLIQGNQEAAPDPTMAQPYASAQFAPPQNGIPAEYTAPHPHPAP
EYTGQTTVPEHTLNLYPPAQTHSEQSPADTSAQTVSGTATQTDDAAPTQGPQTQPSSENTENKSQPKRLH
VSNIPFRFRDPDLRQMFQGFQKILDVEIIFNERGSKGFGFVTFENSADADRAREKLGHTVVEGRKIEVNN
ATARVMTNKKTVNPYTNGWKLNPVVGAVYSPEFYAGTVLLCQANQEGSSMYSAPSSLVYTSAMPGFPYPA
ATAAAAYRGAHLRGRGRTVYNTFRAAAPPPPIPAYGGVVYQEPVYGNKLLQGGYAAAYRYAQPTPATAAAY
SDSYGRVYAADPYHHALAPPTYGVGAMNAFAPLTDKTRSHADDVGLVLSSLQASIYRGGYNRFAPY

```

Comparison of protein sequences between isoform 1 and published

```

>lcl|14673 unnamed protein product
Length=377

```

```

Score = 693 bits (1789), Expect = 0.0, Method: Compositional matrix adjust.
Identities = 370/378 (98%), Positives = 372/378 (98%), Gaps = 1/378 (0%)

```

```

Query   41  MAQPYASAQFAPPQNGIPAEYTAPHPHPAPEYTGQTTVPEHTLNLYPPAQTHSEQSPADT   100
        MAQPYASAQFAPPQNGIPAEYTAPHPHPAPEYTGQTTVPEHTLNLYPPAQTHSEQSPADT
Sbjct    1  MAQPYASAQFAPPQNGIPAEYTAPHPHPAPEYTGQTTVPEHTLNLYPPAQTHSEQSPADT    60

Query   101 SAQTVSGTATQTDDAAPTQGPQTQPSSENTENKSQPKRLHVSNIPIFRFRDPDLRQMFQGF   160
        SAQTVSGTATQTDDAAPTQGPQTQPSSENTENKSQPKRLHVSNIPIFRFRDPDLRQMFQGF
Sbjct    61 SAQTVSGTATQTDDAAPTQGPQTQPSSENTENKSQPKRLHVSNIPIFRFRDPDLRQMFQGF   120

Query   161 GKILDVEIIFNERGSKGFGFVTFENSADADRAREKLGHTVVEGRKIEVNNATARVMTNKK   220
        GKILDVEIIFNERGSKGFGFVTFENSADADRAREKLGHTVVEGRKIEVNNATARVMTNKK
Sbjct   121 GKILDVEIIFNERGSKGFGFVTFENSADADRAREKLGHTVVEGRKIEVNNATARVMTNKK   180

Query   221 TVNPYTNGWKLNPVVGAVYSPEFYAGTVLLCQANQEGSSMYSAPSSLVYTSAMPGFPYPA   280
        TVNPYTNGWKLNPVVGAVYSPEFYAGTVLLCQANQEGSSMYSAPSSLVYTSAMPGFPYPA
Sbjct   181 TVNPYTNGWKLNPVVGAVYSPEFYAGTVLLCQANQEGSSMYSAPSSLVYTSAMPGFPYPA   240

Query   281 ATAAAYRGAHLRGRGRTVYNTFRAAAPPPPIPAYGGVVYQEPVYGNKLLQGGYAAAYRYA   340
        ATAAAYRGAHLRGRGRTVYNTFRAAAPPPPIPAYGGVVYQ+ YG + GGYAAAYRYA
Sbjct   241 ATAAAYRGAHLRGRGRTVYNTFRAAAPPPPIPAYGGVVYQDGFYGADIY-GGYAAAYRYA   299

Query   341 QPTPATAAAYSDSYGRVYAADPYHHALAPPTYGVGAMNAFAPLTDKTRSHADDVGLVL   400
        QPTPATAAAYSDSYGRVYAADPYHHALAPPTYGVGAMNAFAPLTDKTRSHADDVGLVL
Sbjct   300 QPTPATAAAYSDSYGRVYAADPYHHALAPPTYGVGAMNAFAPLTDKTRSHADDVGLVL   359

Query   401 SSLQASIYRGGYNRFAPY   418
        SSLQASIYRGGYNRFAPY
Sbjct   360 SSLQASIYRGGYNRFAPY   377

```

RNA binding protein fox-1 homolog 1 isoform 2 [Homo sapiens]

NCBI Reference Sequence: NP_665899.1

Transcript Variant: This variant (2), also known as alpha, lacks an in-frame segment of the 3' coding region, compared to variant 1. This results in a shorter isoform (2) that lacks an internal segment, compared to isoform 1.

RNA binding protein fox-1 homolog 1 isoform 2 [Homo sapiens]

```

MLASQGVLLHPYGVPMIVPAAPYLPGLIQGNQEAAPDPTMAQPYASAQFAPPQNGIPAEYTAPHPHPAP
EYTGQTTVPEHTLNLYPPAQTHSEQSPADTSAQTVSGTATQTDDAAPTQGPQTQPSSENTENKSQPKRLH

```

VSNIPIFRFRDPDLRQMFQFGKILDVEIIFNERGSKGFGFVTFENSADADRAREKLGHTVVEGRKIEVNN
 ATARVM**TNKKTVNPTYNGWK**LNPVVGAVYSPEFYAGTVLLCQANQEGSSMYSAPSSLVYTSAMPGFPYPA
 ATAAAYRGAHLRGRGRTVYNTFRAAAPPPPIPAYGGVVYQEPVYGNKLLQGGYAAARYAQPTPATAAAY
 SDSYGRVYAADPYHHALAPPTYGVGAMASIYRGGYNRFAPY

Comparison of protein sequences between isoform 2 and published

>lcl|23333 unnamed protein product
 Length=377

Score = 627 bits (1617), Expect = 0.0, Method: Compositional matrix adjust.
 Identities = 344/378 (91%), Positives = 346/378 (92%), Gaps = 27/378 (7%)

Query	41	MAQPYASAQFAPPQNGIPA EYTAPHPHPAPEYTGQTTVPEHTLNLYPPAQTHSEQSPADT	100
		MAQPYASAQFAPPQNGIPA EYTAPHPHPAPEYTGQTTVPEHTLNLYPPAQTHSEQSPADT	
Sbjct	1	MAQPYASAQFAPPQNGIPA EYTAPHPHPAPEYTGQTTVPEHTLNLYPPAQTHSEQSPADT	60
Query	101	SAQTVSGTATQTDDAAPT DGQPQTQPSSENTENKSQPKRLHVSNIPIFRFRDPDLRQMFQGF	160
		SAQTVSGTATQTDDAAPT DGQPQTQPSSENTENKSQPKRLHVSNIPIFRFRDPDLRQMFQGF	
Sbjct	61	SAQTVSGTATQTDDAAPT DGQPQTQPSSENTENKSQPKRLHVSNIPIFRFRDPDLRQMFQGF	120
Query	161	GKILDVEIIFNERGSKGFGFVTFENSADADRAREKLGHTVVEGRKIEVNNATARVMTNKK	220
		GKILDVEIIFNERGSKGFGFVTFENSADADRAREKLGHTVVEGRKIEVNNATARVMTNKK	
Sbjct	121	GKILDVEIIFNERGSKGFGFVTFENSADADRAREKLGHTVVEGRKIEVNNATARVMTNKK	180
Query	221	TVNPTYNGWKLNPNVVGAVYSPEFYAGTVLLCQANQEGSSMYSAPSSLVYTSAMPGFPYPA	280
		TVNPTYNGWKLNPNVVGAVYSPEFYAGTVLLCQANQEGSSMYSAPSSLVYTSAMPGFPYPA	
Sbjct	181	TVNPTYNGWKLNPNVVGAVYSPEFYAGTVLLCQANQEGSSMYSAPSSLVYTSAMPGFPYPA	240
Query	281	ATAAAAYRGAHLRGRGRTVYNTFRAAAPPPPIPAYGGVVYQEPVYGNKLLQGGYAAARYA	340
		ATAAAAYRGAHLRGRGRTVYNTFRAAAPPPPIPAYGGVVYQ+ YG + GGYAAARYA	
Sbjct	241	ATAAAAYRGAHLRGRGRTVYNTFRAAAPPPPIPAYGGVVYQDGFYGADIY-GGYAAARYA	299
Query	341	QPTPATAAAYSDSYGRVYAADPYHHALAPPTYGVGAM-----	378
		QPTPATAAAYSDSYGRVYAADPYHHALAPPTYGVGAM	
Sbjct	300	QPTPATAAAYSDSYGRVYAADPYHHALAPPTYGVGAMNAFAPLTDKTRSHADDVGLVL	359
Query	379	----ASIYRGGYNRFAPY	392
		ASIYRGGYNRFAPY	
Sbjct	360	SSLQASIYRGGYNRFAPY	377

RNA binding protein fox-1 homolog 1 isoform 3 [Homo sapiens]

NCBI Reference Sequence: NP_665900.1

Transcript Variant: This variant (3), also known as beta, has an alternate exon in the 3' coding region, compared to variant 1. This results in a shorter isoform (3) with an alternate C-terminus, compared to isoform 1.

RNA binding protein fox-1 homolog 1 isoform 3 [Homo sapiens]

MLASQGVLLHPYGVPMIVPAAPYLPLGLIQGNQEAAAAPDTMAQPYASAQFAPPQNGIPA EYTAPHPHPAP
 EYTGQTTVPEHTLNLYPPAQTHSEQSPADTSAQTVSGTATQTDDAAPT DGQPQTQPSSENTENKSQPKRLH
 VSNIPFRFRDPDLRQMFQFGKILDVEIIFNERGSKGFGFVTFENSADADRAREKLGHTVVEGRKIEVNN
 ATARVM**TNKKTVNPTYNGWK**LNPVVGAVYSPEFYAGTVLLCQANQEGSSMYSAPSSLVYTSAMPGFPYPA
 ATAAAYRGAHLRGRGRTVYNTFRAAAPPPPIPAYGGVVYQEPVYGNKLLQGGYAAARYAQPTPATAAAY
 SDRNQVFVFAADEISCNTSAVTDEFMLPTPTTTHLLQPPPTALVP

Comparison of protein sequences between isoform 3 and published

>lcl|51101 unnamed protein product
 Length=377

Score = 563 bits (1452), Expect = 2e-165, Method: Compositional matrix adjust.
 Identities = 306/320 (96%), Positives = 309/320 (97%), Gaps = 1/320 (0%)

```

Query   41   MAQPYASAQFAPPQNGIPAEYTAPHPHPAPEYTGQTTVPEHTLNLYPPAQTHSEQSPADT   100
Sbjct   1   MAQPYASAQFAPPQNGIPAEYTAPHPHPAPEYTGQTTVPEHTLNLYPPAQTHSEQSPADT   60

Query   101  SAQTVSGTATQTDDAAPTGGQPTQPSSENTENKSQPKRLHVSNIPIFRFRDPDLRQMFQQF   160
Sbjct   61  SAQTVSGTATQTDDAAPTGGQPTQPSSENTENKSQPKRLHVSNIPIFRFRDPDLRQMFQQF   120

Query   161  GKILDVEIIFNERGSKGFGFVTFENSADADRAREKLHGTVVEGRKIEVNNATARVMTNKK   220
Sbjct   121  GKILDVEIIFNERGSKGFGFVTFENSADADRAREKLHGTVVEGRKIEVNNATARVMTNKK   180

Query   221  TVNPYTNGWKLNPVVGAVYSPEFYAGTVLLCQANQEGSSMYSAPSSLVYTSAMPGFPYPA   280
Sbjct   181  TVNPYTNGWKLNPVVGAVYSPEFYAGTVLLCQANQEGSSMYSAPSSLVYTSAMPGFPYPA   240

Query   281  ATAAAAYRGAHLRGRGRTVYNTFRAAAPPPIIPAYGGVVYQEPVYGNKLLQGGYAAARYA   340
Sbjct   241  ATAAAAYRGAHLRGRGRTVYNTFRAAAPPPIIPAYGGVVYQ+ YG + GGYAARYA   299

Query   341  QPTPATAAAYS DRNQFVFVA   360
Sbjct   300  QPTPATAAAYS DSYGRVYAA   319

```

Sequence of human A2BP1 as in Shibata et al., 2000 Human Molecular Genetics

MAQPYASAQF APPQNGIPAE YTAPHPHPAP EYTGQTTVPE HTLNLYPPAQ THSEQSPADT SAQTVSGTAT
 QTDDAAPTGG QPQTQPSSENT ENKSQPKRLH VSNIPFRFRD PDLRQMFQQF GKILDVEIIF NERGSKGFGF
 VTFENSADAD RAREKLHGTV VEGRKIEVNN ATARVMTNKK TVNPYTNGWK LNPVVGAVYS PEFYAGTVLL
 CQANQEGSSM YSAPSSLVYT SAMPGFPYPA ATAAAAYRGA HLRGRGRTVY NTFRAAAPP PIPAYGGVVY
 QDGFYGADIY GGYAARYAQ PTPATAAAYS DSYGRVYAAD PYHHALAPAP TYGVGAMNAF APLTDAKTRS
 HADDVGLVLS SLQASIYRGG YNRFAPY

MAQPYASAQFAPPQNGIPAEYTAPHPHPAPEYTGQTTVPEHTLNLYPPAQTHSEQSPADTSAQTVSGTATQTDDAAP
 TDGQPTQPSSENTENKSQPKRLHVSNIPIFRFRDPDLRQMFQQFGKILDVEIIFNERGSKGFGFVTFENSADADRARE
 KLHGTVVEGRKIEVNNATARVMTNKKTVNPYTNGWK LNPVVGAVYSPEFYAGTVLLCQANQEGSSMYSAPSSLVYTS
 AMPGFPYPAATAAAAAYRGAHLRGRGRTVYNTFRAAAPPPIIPAYGGVVYQDGFYGADIYGGYAAARYAQPTPATAAA
 YSDSYGRVYAADPYHHALAPAPTYGVGAMNAFAPLTDKTRSHADDVGLVLSLQASIYRGGYNRFAPY

Blast search of human A2BP1 (from Shibata's paper in pubmed) pulled out many similarities and the following is one

```

>  ref|NP\_061193.2|  RNA binding protein fox-1 homolog 1 isoform 4 [Homo sapiens]

ref|NP\_001135806.1|  RNA binding protein fox-1 homolog 1 isoform 4 [Homo sapiens]

sp|Q9NWB1.2|RFOX1\_HUMAN  RecName: Full=RNA binding protein fox-1 homolog 1;
AltName: Full=Ataxin-2-binding
protein 1; AltName: Full=Fox-1 homolog
A; AltName: Full=Hexaribonucleotide-binding protein 1
Length=397

```

[GENE ID: 54715_RBFOX1](#) | RNA binding protein, fox-1 homolog (C. elegans) 1
 [Homo sapiens] (Over 10 PubMed links)

Score = 768 bits (1982), Expect = 0.0, Method: Compositional matrix adjust.
 Identities = 377/377 (100%), Positives = 377/377 (100%), Gaps = 0/377 (0%)

```

Query   1   MAQPYASAQFAPPQNGIPAEYTAPHPHPAPEYTGQTTVPEHTLNLYPPAQTHSEQSPADT   60

```

Sbjct	21	MAQPYASAQFAPPQNGIPAEYTAPHHPHAPPEYTGQTTVPEHTLNLYPPAQTHSEQSPADT MAQPYASAQFAPPQNGIPAEYTAPHHPHAPPEYTGQTTVPEHTLNLYPPAQTHSEQSPADT	80
Query	61	SAQTVSGTATQTDDAAPTGDGQPQTQPSSENTENKSQPKRLHVSNI PFRFRDPDLRQMFGQF SAQTVSGTATQTDDAAPTGDGQPQTQPSSENTENKSQPKRLHVSNI PFRFRDPDLRQMFGQF	120
Sbjct	81	SAQTVSGTATQTDDAAPTGDGQPQTQPSSENTENKSQPKRLHVSNI PFRFRDPDLRQMFGQF	140
Query	121	GKILDVEIIFNERGSKGFGFVTFENSADADRAREKLGHTVVEGRKIEVNNATARVMTNKK GKILDVEIIFNERGSKGFGFVTFENSADADRAREKLGHTVVEGRKIEVNNATARVMTNKK	180
Sbjct	141	GKILDVEIIFNERGSKGFGFVTFENSADADRAREKLGHTVVEGRKIEVNNATARVMTNKK	200
Query	181	TVNPYTNGWKLNPVVGAVYSPEFYAGTVLLCQANQEGSSMYSAPSSLVYTSAMPGFPYPA TVNPYTNGWKLNPVVGAVYSPEFYAGTVLLCQANQEGSSMYSAPSSLVYTSAMPGFPYPA	240
Sbjct	201	TVNPYTNGWKLNPVVGAVYSPEFYAGTVLLCQANQEGSSMYSAPSSLVYTSAMPGFPYPA	260
Query	241	ATAAAAYRGAHLRGRGRTVYNTFRAAAPPPPIPAYGGVVYQDGFYGADIYGGYAAAYRYAQ ATAAAAYRGAHLRGRGRTVYNTFRAAAPPPPIPAYGGVVYQDGFYGADIYGGYAAAYRYAQ	300
Sbjct	261	ATAAAAYRGAHLRGRGRTVYNTFRAAAPPPPIPAYGGVVYQDGFYGADIYGGYAAAYRYAQ	320
Query	301	PTPATAAAYSDSYGRVYAADPYHHALAPAPTYGVGAMNAFAPLTDKTRSHADDVGLVLS PTPATAAAYSDSYGRVYAADPYHHALAPAPTYGVGAMNAFAPLTDKTRSHADDVGLVLS	360
Sbjct	321	PTPATAAAYSDSYGRVYAADPYHHALAPAPTYGVGAMNAFAPLTDKTRSHADDVGLVLS	380
Query	361	SLQASIYRGGYNRFAPY 377 SLQASIYRGGYNRFAPY	
Sbjct	381	SLQASIYRGGYNRFAPY 397	

Human FOX-2 (RBFox2 chr 22) isoforms

RNA binding protein fox-1 homolog 2 isoform 1 [Homo sapiens]

NCBI Reference Sequence: NP_001026865.1
Transcript Variant: This variant (1) represents the longest transcript and encodes isoform 1.

MEKKKMVTQGNQEPTTTTPDAMVQPFTTIPFPPPPQNGIPTEYGVPHQTQDYAGQTGEHNLTLYGSTQAHGE
QSSNSPSTQNGSLTTEGGAQTDGQQSQTSSENSESKSTPKRLHVSNIPIFRFRDPDLRQMFQFGKILDV
EIIIFNERGSKGFGFVTFENSADADRAREKLHGTVVVEGRKIEVNNATARVMTNKKMVT**PYANGWK**LSPVVG
AVYGPELYAASSFQADVSLGNDAAVPLSGRGGINTYIPLISLPLVPGFPYPTAATTAAAFRGAHLRGRGR
TVYGAVRAVPPTAIPAYPGVVYQDGFYAGADLYGGYAAARYAQPATATAATAAAAAAAAAAAYSDGYGRVYTAD
PYHALAPAASYGVGAVASLYRGGYSRFAPY

Highlighted in grey in the above sequence (SLPL) is the sequence stretch that is extra in isoform 1 compared to isoform 3

RNA binding protein fox-1 homolog 2 isoform 2 [Homo sapiens]

NCBI Reference Sequence: NP_055124.1
Transcript Variant: This variant (2) uses two alternate in-frame splice sites in the coding region, and lacks an alternate exon in the 3' coding region that results in a frameshift, compared to variant 1. The resulting protein (isoform 2) is shorter and has a distinct C-terminus compared to isoform 1.

MEKKKMVTQGNQEPTTTTPDAMVQPFTTIPFPPPPQNGIPTEYGVPHQTQDYAGQTGEHNLTLYGSTQAHGE
QSSNSPSTQNGSLTQTEGGAQTDGQQSQTSSENSESKSTPKRLHVSNIPIFRFRDPDLRQMFQFGKILDV
VEIIFNERGSKGFGFVTFENSADADRAREKLHGTVVVEGRKIEVNNATARVMTNKKMVT**PYANGWK**LSPVV
GAVYGPELYAASSFQADVSLGNDAAVPLSGRGGINTYIPLIIPGFPYPTAATTAAAFRGAHLRGRGRRTVY
GAVRAVPPTAIPAYPGVDMQPTDMHSLLLQPQPPLLQLPLQPLTVTMAGCTQPTPTMPLPLPLAMELALW
RVYTEVATADLPTEVT

Isoform 1 and 2 comparison

>lcl|21661 unnamed protein product
Length=367

Sort alignments for this subject sequence by:

		E value	Score	Percent identity	
			Query start position	Subject start position	
Score = 556 bits (1433), Expect = 3e-163, Method: Compositional matrix adjust.					
Identities = 296/304 (97%), Positives = 297/304 (98%), Gaps = 5/304 (2%)					
Query	1	MEKKKMVTQGNQEPTTTTPDAMVQPFTTIPFPPPPQNGIPTEYGVPHQTQDYAGQTGEHNLT	60		
		MEKKKMVTQGNQEPTTTTPDAMVQPFTTIPFPPPPQNGIPTEYGVPHQTQDYAGQTGEHNLT			
Sbjct	1	MEKKKMVTQGNQEPTTTTPDAMVQPFTTIPFPPPPQNGIPTEYGVPHQTQDYAGQTGEHNLT	60		
Query	61	LYGSTQAHGEQSSNSPSTQNGSLT-TEGGAQTDGQQSQTSSENSESKSTPKRLHVSNIPI	119		
		LYGSTQAHGEQSSNSPSTQNGSLT TEGGAQTDGQQSQTSSENSESKSTPKRLHVSNIPI			
Sbjct	61	LYGSTQAHGEQSSNSPSTQNGSLTQTEGGAQTDGQQSQTSSENSESKSTPKRLHVSNIPI	120		
Query	120	FRFRDPDLRQMFQFGKILDVEIIFNERGSKGFGFVTFENSADADRAREKLHGTVVVEGRK	179		
		FRFRDPDLRQMFQFGKILDVEIIFNERGSKGFGFVTFENSADADRAREKLHGTVVVEGRK			
Sbjct	121	FRFRDPDLRQMFQFGKILDVEIIFNERGSKGFGFVTFENSADADRAREKLHGTVVVEGRK	180		
Query	180	IEVNNATARVMTNKKMVT PYANGWK LSPVVGAVYGPELYAASSFQADVSLGNDAAVPLSG	239		
		IEVNNATARVMTNKKMVT PYANGWK LSPVVGAVYGPELYAASSFQADVSLGNDAAVPLSG			
Sbjct	181	IEVNNATARVMTNKKMVT PYANGWK LSPVVGAVYGPELYAASSFQADVSLGNDAAVPLSG	240		
Query	240	RGGINITYIPLISLPLVPGFPYPTAATTAAAFRGAHLRGRGRRTVYGAVRAVPPTAIPAYPG	299		
		RGGINITYIPLI +PGFPYPTAATTAAAFRGAHLRGRGRRTVYGAVRAVPPTAIPAYPG			
Sbjct	241	RGGINITYIPLI----IPGFPYPTAATTAAAFRGAHLRGRGRRTVYGAVRAVPPTAIPAYPG	296		

Query 300 VVYQ 303
V Q
Sbjct 297 VDMQ 300

Score = 13.9 bits (24), Expect = 8.2, Method: Compositional matrix adjust.
Identities = 4/6 (67%), Positives = 5/6 (83%), Gaps = 0/6 (0%)

Query 307 YGADLY 312
YG +LY
Sbjct 214 YGPELY 219

RNA binding protein fox-1 homolog 2 isoform 3 [Homo sapiens]

NCBI Reference Sequence: NP_001076045.1

Transcript Variant: This variant (3) uses an alternate in-frame splice site in the 3' coding region, compared to variant 1. The resulting protein (isoform 3) is 4-aa shorter than isoform 1.

MEKKKMVTQGNQEPTTTPDAMVQPFPTTIPFPPPPQNGIPTEYGVPHQTQDYAGQTGEHNLTLYGSTQAHGE
QSSNSPSTQNGSLTTEGGAQTDGQQSQTQSSENSESKSTPKRLHVSNIFFRFRDPDLRQMFQFGKILDV
EIIIFNERGSKGFGFVTFENSADADRAREKLHGTVVEGRKIEVNNATARVMTNKKMVT**PYANGWK**LSPVVG
AVYGPELYAASSFQADVSLGNDAAVPLSGRGGINTYIPLIIPGFPTYTAATTAFAFRGAHLRGRGRTVYG
AVRAVPPTAIPAYPGVVYQDGFYGDLYGGYAAARYAQPATATAATAAAAAAAYSDDGYGRVYTADPYHA
LAPAASYGVGAVASLYRGGYSRFAPY

Isoform 1 and 3 comparison

>lcl|63017 unnamed protein product
Length=376

Score = 684 bits (1765), Expect = 0.0, Method: Compositional matrix adjust.
Identities = 375/380 (99%), Positives = 376/380 (99%), Gaps = 4/380 (1%)

Query	1	MEKKKMVTQGNQEPTTTPDAMVQPFPTTIPFPPPPQNGIPTEYGVPHQTQDYAGQTGEHNLT	60
		MEKKKMVTQGNQEPTTTPDAMVQPFPTTIPFPPPPQNGIPTEYGVPHQTQDYAGQTGEHNLT	
Sbjct	1	MEKKKMVTQGNQEPTTTPDAMVQPFPTTIPFPPPPQNGIPTEYGVPHQTQDYAGQTGEHNLT	60
Query	61	LYGSTQAHGEQSSNSPSTQNGSLTTEGGAQTDGQQSQTQSSENSESKSTPKRLHVSNIFF	120
		LYGSTQAHGEQSSNSPSTQNGSLTTEGGAQTDGQQSQTQSSENSESKSTPKRLHVSNIFF	
Sbjct	61	LYGSTQAHGEQSSNSPSTQNGSLTTEGGAQTDGQQSQTQSSENSESKSTPKRLHVSNIFF	120
Query	121	RFRDPDLRQMFQFGKILDVEIIFNERGSKGFGFVTFENSADADRAREKLHGTVVEGRKI	180
		RFRDPDLRQMFQFGKILDVEIIFNERGSKGFGFVTFENSADADRAREKLHGTVVEGRKI	
Sbjct	121	RFRDPDLRQMFQFGKILDVEIIFNERGSKGFGFVTFENSADADRAREKLHGTVVEGRKI	180
Query	181	EVNNATARVMTNKKMVT PYANGWK LSPVVGAVYGPELYAASSFQADVSLGNDAAVPLSGR	240
		EVNNATARVMTNKKMVT PYANGWK LSPVVGAVYGPELYAASSFQADVSLGNDAAVPLSGR	
Sbjct	181	EVNNATARVMTNKKMVT PYANGWK LSPVVGAVYGPELYAASSFQADVSLGNDAAVPLSGR	240
Query	241	GGINTYIPLISLPLVPGFPTYTAATTAFAFRGAHLRGRGRTVYGAVRAVPPTAIPAYPGV	300
		GGINTYIPLI +PGFPTYTAATTAFAFRGAHLRGRGRTVYGAVRAVPPTAIPAYPGV	
Sbjct	241	GGINTYIPLI----IPGFPTYTAATTAFAFRGAHLRGRGRTVYGAVRAVPPTAIPAYPGV	296
Query	301	VYQDGFYGDLYGGYAAARYAQPATATAATAAAAAAAYSDDGYGRVYTADPYHALAPAAS	360
		VYQDGFYGDLYGGYAAARYAQPATATAATAAAAAAAYSDDGYGRVYTADPYHALAPAAS	
Sbjct	297	VYQDGFYGDLYGGYAAARYAQPATATAATAAAAAAAYSDDGYGRVYTADPYHALAPAAS	356
Query	361	YGVGAVASLYRGGYSRFAPY	380
		YGVGAVASLYRGGYSRFAPY	
Sbjct	357	YGVGAVASLYRGGYSRFAPY	376

RNA binding protein fox-1 homolog 2 isoform 4 [Homo sapiens]

NCBI Reference Sequence: NP_001076046.1

Transcript Variant: This variant (4) lacks an alternate exon in the 3' coding region that results in a frameshift, compared to variant 1. The resulting protein (isoform 4) is shorter and has a distinct C-terminus compared to isoform 1.

```
MEKKKMVTQGNQEPTTTPDAMVQPFTTIPFPPPPQNGIPT EYGVPH TQDYAGQTGEHNLTLYGSTQAHGE
QSSNSPSTQNGSLTTEGGAQTDGQSQTSSENSESKSTPKRLHVSNI PFRFRDPDLRQMFGQFGKILDV
EIIIFNERGSKGFGFVTFENSADADRAREKLHGT VVEGRKIEVNNATARVMTNKKMVT PYANGWKLSPVVG
AVYGPELYAASSFQADVSLGNDAAVPLSGRGGINTYIPLISLPLVPGFPYPTAATTAAAFRGAHLRGRGR
TVYGAVRAVPPTAIPAYPGVDMQPTDMHSLLLQPQPPLLQPLQPLTVTVMAGCTQPTPTMPLPLPLAMEL
ALWRVYTEVATADLPPTVET
```

RNA binding protein fox-1 homolog 2 isoform 5 [Homo sapiens]

NCBI Reference Sequence: NP_001076047.1

Transcript Variant: This variant (5) contains a distinct 5' UTR, 5' CDS, and uses an alternate in-frame splice site in the coding region, compared to variant 1. The resulting protein (isoform 5) is longer and has a distinct N-terminus compared to isoform 1.

```
MAEGAQPQPPQLGPGAAARGMKRESELELPVPGAGGDGADPGLSKRPRT EEEAADGGGGMQNEPLTPGY
HGFPARDSQGNQEPTTTPDAMVQPFTTIPFPPPPQNGIPT EYGVPH TQDYAGQTGEHNLTLYGSTQAHGE
QSSNSPSTQNGSLTQTEGGAQTDGQSQTSSENSESKSTPKRLHVSNI PFRFRDPDLRQMFGQFGKILD
VEIIFNERGSKGFGFVTFENSADADRAREKLHGT VVEGRKIEVNNATARVMTNKKMVT PYANGWKLSPVV
GAVYGPELYAASSFQADVSLGNDAAVPLSGRGGINTYIPLISLPLVPGFPYPTAATTAAAFRGAHLRGRG
RTVYGAVRAVPPTAIPAYPGVVYQDGFY GADLYGGY AAYRYAQPATATAATAAAAAAAYS DGYGRVYTA
DPYHALAPAASYGVGAVASLYRGGYSRFAPY
```

RNA binding protein fox-1 homolog 2 isoform 6 [Homo sapiens]

NCBI Reference Sequence: NP_001076048.1

Transcript Variant: This variant (6) contains a distinct 5' UTR and 5' CDS, compared to variant 1. The resulting protein (isoform 6) is longer and has a distinct N-terminus compared to isoform 1

```
MAEGAQPQPPQLGPGAAARGMKRESELELPVPGAGGDGADPGLSKRPRT EEEAADGGGGMQNEPLTPGY
HGFPARDSQGNQEPTTTPDAMVQPFTTIPFPPPPQNGIPT EYGVPH TQDYAGQTGEHNLTLYGSTQAHGE
QSSNSPSTQNGSLTTEGGAQTDGQSQTSSENSESKSTPKRLHVSNI PFRFRDPDLRQMFGQFGKILDV
EIIIFNERGSKGFGFVTFENSADADRAREKLHGT VVEGRKIEVNNATARVMTNKKMVT PYANGWKLSPVVG
AVYGPELYAASSFQADVSLGNDAAVPLSGRGGINTYIPLISLPLVPGFPYPTAATTAAAFRGAHLRGRGR
TVYGAVRAVPPTAIPAYPGVVYQDGFY GADLYGGY AAYRYAQPATATAATAAAAAAAYS DGYGRVYTAD
PYHALAPAASYGVGAVASLYRGGYSRFAPY
```

Human FOX-3 (RBFox3 chr 17) isoforms

RNA binding protein fox-1 homolog 3 [Homo sapiens]

NCBI Reference Sequence: NP_001076044.1

MAQPYPPAQYPPPPQNGIPAEYAPPPPHPTQDYSGQTPVPTEHGMTLYTPAQTHPEQPGSEASTQPIAGT
QTVPQTDEAAQTDSQPLHPSDPTEKQQPKRLHVSNIFFRFRDPDLRQMFGQFGKILDVEIIFNERGSKGF
GFVTFETSSDADRAREKLNGTIVEGRKIEVNNATARVMTNKKTGNPYTNGWKLNPVVGAVYGPEFYAVTG
FPYPTTGTAVAYRGAHLRGRGRAVYNTFRAAPPPPIPTYGAVVYQDGFYGAIEYGGYAAARYAQAAAA
AAYSDSYGRVYAAADPYHHTIGPAATYSIGTM

Monkey A2BP1 sequence

PREDICTED: RNA binding protein fox-1 homolog 1 isoform 3 [Pan troglodytes].

ACCESSION XP_001135062
VERSION XP_001135062.2 GI:332845232
DBSOURCE REFSEQ: accession [XM_001135062.2](#)
KEYWORDS .
SOURCE Pan troglodytes (chimpanzee)

MNCEREQLRGNQEAAAAPDTMAQPYASAQFAPPQNGIPAEYTAPHPHPAPEYTGQTTVPEHTLNLYPPAQ
THSEQSPADTSAQTVSGTATQTDDAAPT DGQPQTQPSSENTENKSQPKRLHVSNI PFRFRDPDLRQMFGQF
GKILDVEIIFNERGSKGFGFVTFENSADADRAREKLHGT VVEGRKIEVNNATARVM **TNKKTVNPTYNGWK**
LNPVVGAVYSPEFYAGTVLLCQANQEGSSMYSAPSSLVYTSASKPTVVALFV FVVYQDGFY GADIYGGYA
AYRYAQPTPATAAAYS DSYGRVYAADPYHHALAPPTYGVGAMNAFAPLTD AKTRSHADDVGLVLSSLQA
SIYRGGYNRFAPY

Multiple Sequence Alignment

Generated by MUSCLE [\[see reference\]](#) version 3.6 (using option: -maxiters 2).

Homo sapiens	1	MLASQGVLLHPYGVPMIVPAAPYLPGLIQGNQEAAAAPDTMAQPYASAQF	50
Chimpanzee	1	MLASQGVLLHPYGVPMIVPAAPYLPGLIQGNQEAAAAPDTMAQPYASAQF	50
dog	1	MLPSQGVLLHPYGVPMIVPAAPYFPGLIQGNQEAAAAPDTMAQPYASAQF	50
mouse	1	MLASQGVLLHSYGVPMIVPAAPYFPGLMQGNQEAAAAPDTMAQPYASAQF	50
rat	1	MLASQGVLLHSYGVPMVVPAAFFPGLMQGNQEAAAAPDTMAQPYASAQF	50
red jungle fowl	1	MSCEREQL-----RGNQEATAPPDTMAQPYASAQF	30
zebrafish	1	-MEEKGSKM-----VEQGNQDAPAPPETMAQPFPSAQF	32
NP_665899.1	51	APPQNGIPAEYTAPHPHPAPEYTGQTTVPEHTLNLYPPAQTHSEQSPADT	100
XP_510799.2	51	APPQNGIPAEYTAPHPHPAPEYTGQTTVPEHTLNLYPPAQTHSEQSPADT	100
XP_863750.1	51	APPQNGIPAEYTAPHPHPAPEYTGQTTVPEHTLNLYPPAQTHSEQSAADT	100
NP_899011.1	51	APPQNGIPAEYTAPHPHPAPEYTGQTTVPDHTLNLYPPTQTHSEQS-ADT	99
XP_220155.3	51	APPQNGIPAEYTAPHPHPAPEYTGQTTVPDHTLNLYPPTQTHSEQS-ADT	99
XP_414942.2	31	APPQNGIPAEYTAPHPHPAPDYTGQTTVPEHTLNMYPPAQTHSEQSAADT	80
NP_001005596.1	33	APPQNGIPAEYTPSHPHPTPDYSGQTPVAEHTLNLYTPAQTHSEPSGPDN	82
NP_665899.1	101	SAQTVSGTATQTDDAAPT DGQPQTQPSSENTENKSQPKRLHVSNI PFRFRD	150
XP_510799.2	101	SAQTVSGTATQTDDAAPT DGQPQTQPSSENTENKSQPKRLHVSNI PFRFRD	150
XP_863750.1	101	SAQTVSGTATQTDDAAPT DGQPQTQPSSENTENKSQPKRLHVSNI PFRFRD	150
NP_899011.1	100	SAQTVSGTATQTDDAAPT DGQPQTQPSSENTESKSQPKRLHVSNI PFRFRD	149
XP_220155.3	100	SAQTVSGTATQTDDAAPT DGQPQTQPSSENTENKSQPKRLHVSNI PFRFRD	149
XP_414942.2	81	NAQTVSGTATQTEDAAQT DGQQQTQSSENTENKSQPKRLHVSNI PFRFRD	130
NP_001005596.1	83	SIQAVSGTATQTDDSAQTSQQQTQSSEITEIKTQPKRLHVSNI PFRFRD	132
NP_665899.1	151	PDLRQMFGQFGKILDVEIIFNERGSKGFGFVTFENSADADRAREKLHGTV	200
XP_510799.2	151	PDLRQMFGQFGKILDVEIIFNERGSKGFGFVTFENSADADRAREKLHGTV	200
XP_863750.1	151	PDLRQMFGQFGKILDVEIIFNERGSKGFGFVTFENSADADRAREKLHGTV	200
NP_899011.1	150	PDLRQMFGQFGKILDVEIIFNERGSKGFGFVTFENSADADRAREKLHGTV	199
XP_220155.3	150	PDLRQMFGQFGKILDVEIIFNERGSKGFGFVTFENSADADRAREKLHGTV	199
XP_414942.2	131	PDLRQMFGQFGKILDVEIIFNERGSKGFGFVTFENSADADRAREKLHGTV	180
NP_001005596.1	133	PDLRQMFGQFGKILDVEIIFNERGSKGFGFVTFESSADADRAREKLHGTV	182
NP_665899.1	201	VEGRKIEVNNATARVM TNKKTVNPTYNGWK LNPVVGAVYSPEFYAGTVLL	250
XP_510799.2	201	VEGRKIEVNNATARVM TNKKTVNPTYNGWK LNPVVGAVYSPEFYAGTVLL	250
XP_863750.1	201	VEGRKIEVNNATARVM TNKKTVNPTYNGWK LNPVVGAVYSPEFYAGTVLL	250

NP 899011.1	200	VEGRKIEVNNATARVM	TNKKTVN	PNPYTNGWK	LNPVVGAVYSPDFYAGTVLL	249
XP 220155.3	200	VEGRKIEVNNATARVM	TNKKAVN	PNPYTNGWK	LNPVVGAVYSPDFYAGTVLL	249
XP 414942.2	181	VEGRKIEVNNATARVM	TNKKTVN	PNPYTNGWK	LNPVVGAVYSPEFYAGRVLL	230
NP 001005596.1	183	VEGRKIEVNNATARVM	TNKKTVN	PNPYANGWK	LNPVVGAVYSPEFYA-----	227
NP 665899.1	251	CQANQEGSSMYSAPSSLVYTSAMP	GFP---	YPAAT-AAAA	YRG AHLRGRG	296
XP 510799.2	251	CQANQEGSSMYSAPSSLVYTSAM	SGIKSLCYPAYS-AISAHCKLHLL	GRG		299
XP 863750.1	251	CQANQEGSSMYSAPSSLVYTSAMP	GFP---	YPAAT-AAAA	YRG AHLRGRG	296
NP 899011.1	250	CQANQEGSSMYS	SGPSSLVYTSAMP	GFP---	YPAAS-AAAA	YRG AHLRGRG
XP 220155.3	250	CQANQEGSSMYS	SGPSSLVYTSAMP	GFP---	YPAAT-AAAA	YRG AHLRGRG
XP 414942.2	231	CQANQEGSPVYSAPSSLVYT	STMPGFP---	YPAAT-AAAA	YRG AHLRGRG	276
NP 001005596.1	228	-----	VPGFP---	YPAATAAAAA	YRG AHLRGRG	252
NP 665899.1	297	RTVYNTFRAAAPPPPI	PAYGGVVYQ	EPVYGNKLLQGGY	AA YRQAQTPAT	346
XP 510799.2	300	NSVVGSLGPSTPPPPHPNS	GGVVYQ	EPVYGNKLLQGGY	AA YRQAQTPAT	349
XP 863750.1	297	RTVYNTFRAAAPPPPI	PAYGGVVYQ	EPVYGNKLLQGGY	AA YRQAQTPAT	346
NP 899011.1	296	RTGYNTLRRAAAPPPPI	PAYGGVVYQ	EPVYGNKLLQGGY	AA YRQAQTPAT	345
XP 220155.3	296	RTVYNTFRAAAPPPPI	PAYGGVVYQ	DGFYGADIY-GGY	AA YRQAQTPAT	344
XP 414942.2	277	RTVYNTFRAAAPPPPI	PAYGGVVYQ	DGFYGADIY-GGY	AA YRQAQPATAT	325
NP 001005596.1	253	RTVYNTFRAAAPPPHI	PAYGGVVYQ	DGFYGADIY-GGY	TAYRYTQPATAT	301
NP 665899.1	347	AAAYS	SDSYGRVYAADPYH	HALAPAPTYGV	GAM-----	378
XP 510799.2	350	AAAYS	SDSYGRVYAADPYH	HALAPAPTYGV	GAMNAFAPLTD	AKTRSHADDV
XP 863750.1	347	AAAYS	SDSYGRVYAADPYH	HHTLAPAPTYSV	GAM-----	378
NP 899011.1	346	AAAYS	SDSYGRVYAADPYH	HHTLAPAPTYGV	GAMNAFAPLTD	AKTRSHADDV
XP 220155.3	345	AAAYS	SDSYGRVYAADPYH	HHTLAPAPTYGV	GAMNAFAPLTD	AKTRSHADDV
XP 414942.2	326	AAAYS	SDSYGRVYAADPYH	HHTLAPAA	TYGVGAVNAFAPLTD	AKTRSHADDV
NP 001005596.1	302	AAAYS	SDSYGRVYAADPY	NHALAPAA	TYSVGAMNAFAPLTD	AKTRSHADDV
NP 665899.1	379	-----	ASIYRGGYN	RFAPY		392
XP 510799.2	400	GLVLSSLQ	ASIYRGGYN	RFAPY		421
XP 863750.1	379	-----	ASIYRGGYN	RFAPY		392
NP 899011.1	396	GLVLSSLQ	ASIYRGGYN	RFAPY		417
XP 220155.3	395	GLVLSSLQ	ASIYRGGYN	RFAPY		416
XP 414942.2	376	GLVLSSLQ	ASIYRGGY	SRFAPY		397
NP 001005596.1	352	GLVLSSLQ	ASIYRGGY	SRFAPY		373

Protein Acc.	Gene	Organism
NP 665899.1	RBFOX1	H.sapiens (man)
XP 510799.2	RBFOX1	P.troglodytes (Chimpanzee)
XP 863750.1	RBFOX1	C.lupus (dog)
NP 899011.1	Rbfox1	M.musculus (house mouse)
XP 220155.3	A2bp1	R.norvegicus (Norway rat)
XP 414942.2	RBFOX1	G.gallus (red jungle fowl)

Alignment of different isoforms of human RBFOX-1

CLUSTAL W (1.83) multiple sequence alignment

```
hFOX1-iso1      MLASQGVLLHPYGVPMIVPAAPYLPGLIQGNQEAAAAPDTMAQPYASAQF
hFOX1-iso2      MLASQGVLLHPYGVPMIVPAAPYLPGLIQGNQEAAAAPDTMAQPYASAQF
hFOX1-iso3      MLASQGVLLHPYGVPMIVPAAPYLPGLIQGNQEAAAAPDTMAQPYASAQF
hFOX1-iso4      MNCERE-----QLRGNQEAAAAPDTMAQPYASAQF
hFOX1-iso5      MNCERE-----QLRGNQEAAAAPDTMAQPYASAQF
*   .:.                               :.:*****

hFOX1-iso1      APPQNGIPAEYTAPHPHPAPEYTGQTTVPEHTLNLYPPAQTHSEQSPADT
hFOX1-iso2      APPQNGIPAEYTAPHPHPAPEYTGQTTVPEHTLNLYPPAQTHSEQSPADT
hFOX1-iso3      APPQNGIPAEYTAPHPHPAPEYTGQTTVPEHTLNLYPPAQTHSEQSPADT
hFOX1-iso4      APPQNGIPAEYTAPHPHPAPEYTGQTTVPEHTLNLYPPAQTHSEQSPADT
hFOX1-iso5      APPQNGIPAEYTAPHPHPAPEYTGQTTVPEHTLNLYPPAQTHSEQSPADT
*****

hFOX1-iso1      SAQTVSGTATQTDDAAPT DGQPQTQPSSENTENKSQPKRLHVSNI PFRFRD
hFOX1-iso2      SAQTVSGTATQTDDAAPT DGQPQTQPSSENTENKSQPKRLHVSNI PFRFRD
hFOX1-iso3      SAQTVSGTATQTDDAAPT DGQPQTQPSSENTENKSQPKRLHVSNI PFRFRD
hFOX1-iso4      SAQTVSGTATQTDDAAPT DGQPQTQPSSENTENKSQPKRLHVSNI PFRFRD
hFOX1-iso5      SAQTVSGTATQTDDAAPT DGQPQTQPSSENTENKSQPKRLHVSNI PFRFRD
*****

hFOX1-iso1      PDLRQMFGQFGKILDVEIIFNERGSKGFGFVTFENSADADRAREKLHGTV
hFOX1-iso2      PDLRQMFGQFGKILDVEIIFNERGSKGFGFVTFENSADADRAREKLHGTV
hFOX1-iso3      PDLRQMFGQFGKILDVEIIFNERGSKGFGFVTFENSADADRAREKLHGTV
hFOX1-iso4      PDLRQMFGQFGKILDVEIIFNERGSKGFGFVTFENSADADRAREKLHGTV
hFOX1-iso5      PDLRQMFGQFGKILDVEIIFNERGSKGFGFVTFENSADADRAREKLHGTV
*****

hFOX1-iso1      VEGRKIEVNNATARVMTNKKT VNPYTNGWKLNPVVGAVYSPEFYAGTVLL
hFOX1-iso2      VEGRKIEVNNATARVMTNKKT VNPYTNGWKLNPVVGAVYSPEFYAGTVLL
hFOX1-iso3      VEGRKIEVNNATARVMTNKKT VNPYTNGWKLNPVVGAVYSPEFYAGTVLL
hFOX1-iso4      VEGRKIEVNNATARVMTNKKT VNPYTNGWKLNPVVGAVYSPEFYAGTVLL
hFOX1-iso5      VEGRKIEVNNATARVMTNKKT VNPYTNGWKLNPVVGAVYSPEFYA-----
*****

hFOX1-iso1      CQANQEGSSMYSAPSSLVYTSAMPGFPYPAATAAAAYRGAHLRGRGRTVY
hFOX1-iso2      CQANQEGSSMYSAPSSLVYTSAMPGFPYPAATAAAAYRGAHLRGRGRTVY
hFOX1-iso3      CQANQEGSSMYSAPSSLVYTSAMPGFPYPAATAAAAYRGAHLRGRGRTVY
hFOX1-iso4      CQANQEGSSMYSAPSSLVYTSAMPGFPYPAATAAAAYRGAHLRGRGRTVY
hFOX1-iso5      -----VPGFPYPAATAAAAYRGAHLRGRGRTVY
                               :.:*****

hFOX1-iso1      NTFRAAAPPPPIPAYGGVVYQEPVYGNKLLQGGYAAAYRYAQPTPATAAAY
hFOX1-iso2      NTFRAAAPPPPIPAYGGVVYQEPVYGNKLLQGGYAAAYRYAQPTPATAAAY
hFOX1-iso3      NTFRAAAPPPPIPAYGGVVYQEPVYGNKLLQGGYAAAYRYAQPTPATAAAY
hFOX1-iso4      NTFRAAAPPPPIPAYGGVVYQDGFYGADI-YGGYAAAYRYAQPTPATAAAY
hFOX1-iso5      NTFRAAAPPPPIPAYGGVVYQDGFYGADI-YGGYAAAYRYAQPTPATAAAY
*****: .** .: *****

hFOX1-iso1      SDSYGRVYAA-D-----PYHHALAPPTYGVGAMNAF
```

```

hFOX1-iso2      SDSYGRVYAA-D-----PYHHALAPAPTYGVGAM---
hFOX1-iso3      SDRNQFVFVADEISCNTSAVTDEFMLPTPTTTHLLQPPPTALV-----
hFOX1-iso4      SDSYGRVYAA-D-----PYHHALAPAPTYGVGAMNAF
hFOX1-iso5      SDSYGRVYAA-D-----PYHHALAPAPTYGVGAMNAF
                  **      *:. *      .      * * * *.** *

hFOX1-iso1      APLTDAKTRSHADDVGLVLSSLQASIYRGGYNRFAPY
hFOX1-iso2      -----ASIYRGGYNRFAPY
hFOX1-iso3      -----P
hFOX1-iso4      APLTDAKTRSHADDVGLVLSSLQASIYRGGYNRFAPY
hFOX1-iso5      APLTDAKTRSHADDVGLVLSSLQASIYRGGYNRFAPY

```

Alignment of different isoforms of human RBFOX-1,2,3

CLUSTAL W (1.83) multiple sequence alignment

```

hFOX1-iso1      MLASQGVLLHPYGVPMIVPAA-----
hFOX1-iso2      MLASQGVLLHPYGVPMIVPAA-----
hFOX1-iso3      MLASQGVLLHPYGVPMIVPAA-----
hFOX1-iso4      MNCERE-----
hFOX1-iso5      MNCERE-----
hFOX2-iso1      MEKKK-----
hFOX2-iso2      MEKKK-----
hFOX2-iso3      MEKKK-----
hFOX2-iso4      MEKKK-----
hFOX2-iso5      MAEGAQP---QPPQLGPGAAARGMKRESELELPVPGAGGDGADPGLSK
hFOX2-iso6      MAEGAQP---QPPQLGPGAAARGMKRESELELPVPGAGGDGADPGLSK
hFOX3           -----

hFOX1-iso1      -----P-----YLPGLIQGNQEAAAAPDTMAQPYA
hFOX1-iso2      -----P-----YLPGLIQGNQEAAAAPDTMAQPYA
hFOX1-iso3      -----P-----YLPGLIQGNQEAAAAPDTMAQPYA
hFOX1-iso4      -----QLRGNQEAAAAPDTMAQPYA
hFOX1-iso5      -----QLRGNQEAAAAPDTMAQPYA
hFOX2-iso1      -----MVTQGNQEPTTTPDAMVQPFT
hFOX2-iso2      -----MVTQGNQEPTTTPDAMVQPFT
hFOX2-iso3      -----MVTQGNQEPTTTPDAMVQPFT
hFOX2-iso4      -----MVTQGNQEPTTTPDAMVQPFT
hFOX2-iso5      RPRTEEAAADGGGGMQNEPLTPGYHGFPARDSQGNQEPTTTPDAMVQPFT
hFOX2-iso6      RPRTEEAAADGGGGMQNEPLTPGYHGFPARDSQGNQEPTTTPDAMVQPFT
hFOX3           -----MAQPYP
                        *.**:.

hFOX1-iso1      SAQFA-PPQNGIPAEYTAPHPHPAPEYTGQTTV-PEHTLNLYPPAQTHSE
hFOX1-iso2      SAQFA-PPQNGIPAEYTAPHPHPAPEYTGQTTV-PEHTLNLYPPAQTHSE
hFOX1-iso3      SAQFA-PPQNGIPAEYTAPHPHPAPEYTGQTTV-PEHTLNLYPPAQTHSE
hFOX1-iso4      SAQFA-PPQNGIPAEYTAPHPHPAPEYTGQTTV-PEHTLNLYPPAQTHSE
hFOX1-iso5      SAQFA-PPQNGIPAEYTAPHPHPAPEYTGQTTV-PEHTLNLYPPAQTHSE
hFOX2-iso1      TIPFPPPPQNGIPT EYGVPH---QDYAGQTG---EHNLTLYGSTQAHGE
hFOX2-iso2      TIPFPPPPQNGIPT EYGVPH---QDYAGQTG---EHNLTLYGSTQAHGE
hFOX2-iso3      TIPFPPPPQNGIPT EYGVPH---QDYAGQTG---EHNLTLYGSTQAHGE
hFOX2-iso4      TIPFPPPPQNGIPT EYGVPH---QDYAGQTG---EHNLTLYGSTQAHGE
hFOX2-iso5      TIPFPPPPQNGIPT EYGVPH---QDYAGQTG---EHNLTLYGSTQAHGE
hFOX2-iso6      TIPFPPPPQNGIPT EYGVPH---QDYAGQTG---EHNLTLYGSTQAHGE

```


hFOX3	PAQYPPPPQNGIPA EYAPPPHPTQDYSGQTPVPTEHGMTLYTPAQTHPE .:. *****: ** . :*:*** ** :. ** .: *: *
hFOX1-iso1	QSPADTSAQTVSGTAT--QTDDAAPT DGQ-PQTQPSSENTENKSQPKRLHV
hFOX1-iso2	QSPADTSAQTVSGTAT--QTDDAAPT DGQ-PQTQPSSENTENKSQPKRLHV
hFOX1-iso3	QSPADTSAQTVSGTAT--QTDDAAPT DGQ-PQTQPSSENTENKSQPKRLHV
hFOX1-iso4	QSPADTSAQTVSGTAT--QTDDAAPT DGQ-PQTQPSSENTENKSQPKRLHV
hFOX1-iso5	QSPADTSAQTVSGTAT--QTDDAAPT DGQ-PQTQPSSENTENKSQPKRLHV
hFOX2-iso1	QSSNSPSTQ--NGSLT---TEGGAQTDGQQSQTQSSENSESKSTPKRLHV
hFOX2-iso2	QSSNSPSTQ--NGSLT---QTEGGAQTDGQQSQTQSSENSESKSTPKRLHV
hFOX2-iso3	QSSNSPSTQ--NGSLT---TEGGAQTDGQQSQTQSSENSESKSTPKRLHV
hFOX2-iso4	QSSNSPSTQ--NGSLT---TEGGAQTDGQQSQTQSSENSESKSTPKRLHV
hFOX2-iso5	QSSNSPSTQ--NGSLT---QTEGGAQTDGQQSQTQSSENSESKSTPKRLHV
hFOX2-iso6	QSSNSPSTQ--NGSLT---TEGGAQTDGQQSQTQSSENSESKSTPKRLHV
hFOX3	QPGSEASTQPIAGTQTVPQTDEAAQTDSQ-PL-HPSDPTE-KQQPKRLHV *. ..*: * *: * *: . * **.* . :. *: : * *. *****
hFOX1-iso1	SNIPFRFRD PDLRQMFGQFGKILDVEIIFNERGSKGFGFVTFENSADADR
hFOX1-iso2	SNIPFRFRD PDLRQMFGQFGKILDVEIIFNERGSKGFGFVTFENSADADR
hFOX1-iso3	SNIPFRFRD PDLRQMFGQFGKILDVEIIFNERGSKGFGFVTFENSADADR
hFOX1-iso4	SNIPFRFRD PDLRQMFGQFGKILDVEIIFNERGSKGFGFVTFENSADADR
hFOX1-iso5	SNIPFRFRD PDLRQMFGQFGKILDVEIIFNERGSKGFGFVTFENSADADR
hFOX2-iso1	SNIPFRFRD PDLRQMFGQFGKILDVEIIFNERGSKGFGFVTFENSADADR
hFOX2-iso2	SNIPFRFRD PDLRQMFGQFGKILDVEIIFNERGSKGFGFVTFENSADADR
hFOX2-iso3	SNIPFRFRD PDLRQMFGQFGKILDVEIIFNERGSKGFGFVTFENSADADR
hFOX2-iso4	SNIPFRFRD PDLRQMFGQFGKILDVEIIFNERGSKGFGFVTFENSADADR
hFOX2-iso5	SNIPFRFRD PDLRQMFGQFGKILDVEIIFNERGSKGFGFVTFENSADADR
hFOX2-iso6	SNIPFRFRD PDLRQMFGQFGKILDVEIIFNERGSKGFGFVTFENSADADR
hFOX3	SNIPFRFRD PDLRQMFGQFGKILDVEIIFNERGSKGFGFVTFETSSDADR *****: *: ***** . *: *****
hFOX1-iso1	AREKLHGT VVEGRKIEVNNATARVM TNKKTVN PYTNGWKLNPVVGAVYSP
hFOX1-iso2	AREKLHGT VVEGRKIEVNNATARVM TNKKTVN PYTNGWKLNPVVGAVYSP
hFOX1-iso3	AREKLHGT VVEGRKIEVNNATARVM TNKKTVN PYTNGWKLNPVVGAVYSP
hFOX1-iso4	AREKLHGT VVEGRKIEVNNATARVM TNKKTVN PYTNGWKLNPVVGAVYSP
hFOX1-iso5	AREKLHGT VVEGRKIEVNNATARVM TNKKTVN PYTNGWKLNPVVGAVYSP
hFOX2-iso1	AREKLHGT VVEGRKIEVNNATARVM TNKKMVT PYANGWKLSPVVGAVYGP
hFOX2-iso2	AREKLHGT VVEGRKIEVNNATARVM TNKKMVT PYANGWKLSPVVGAVYGP
hFOX2-iso3	AREKLHGT VVEGRKIEVNNATARVM TNKKMVT PYANGWKLSPVVGAVYGP
hFOX2-iso4	AREKLHGT VVEGRKIEVNNATARVM TNKKMVT PYANGWKLSPVVGAVYGP
hFOX2-iso5	AREKLHGT VVEGRKIEVNNATARVM TNKKMVT PYANGWKLSPVVGAVYGP
hFOX2-iso6	AREKLHGT VVEGRKIEVNNATARVM TNKKMVT PYANGWKLSPVVGAVYGP
hFOX3	AREKLNGTIVEGRKIEVNNATARVM TNKKTGN PYTNGWKLNPVVGAVYGP *****: *: ***** . *: ***** . ***** . *
hFOX1-iso1	EFYAGTVLLCQANQEGSSMYSAP-----SSLVYTSAMPGFYP-A
hFOX1-iso2	EFYAGTVLLCQANQEGSSMYSAP-----SSLVYTSAMPGFYP-A
hFOX1-iso3	EFYAGTVLLCQANQEGSSMYSAP-----SSLVYTSAMPGFYP-A
hFOX1-iso4	EFYAGTVLLCQANQEGSSMYSAP-----SSLVYTSAMPGFYP-A
hFOX1-iso5	EFYA-----VPGFPYP-A
hFOX2-iso1	ELYAASSFQADVSL--GNDAAVPLSGRGGINTYIPLISLPLVPGFPYPTA
hFOX2-iso2	ELYAASSFQADVSL--GNDAAVPLSGRGGINTYIPL----IIPGFYPYPTA
hFOX2-iso3	ELYAASSFQADVSL--GNDAAVPLSGRGGINTYIPL----IIPGFYPYPTA
hFOX2-iso4	ELYAASSFQADVSL--GNDAAVPLSGRGGINTYIPLISLPLVPGFPYPTA
hFOX2-iso5	ELYAASSFQADVSL--GNDAAVPLSGRGGINTYIPLISLPLVPGFPYPTA
hFOX2-iso6	ELYAASSFQADVSL--GNDAAVPLSGRGGINTYIPLISLPLVPGFPYPTA
hFOX3	EFYA-----VTGFPYP-T

*: **

: .***** :

hFOX1-iso1	ATAAAAYRGAHLRGRGRTVYNTFRAAAPPPPIPAYGGVVYQEPVYGNKLL
hFOX1-iso2	ATAAAAYRGAHLRGRGRTVYNTFRAAAPPPPIPAYGGVVYQEPVYGNKLL
hFOX1-iso3	ATAAAAYRGAHLRGRGRTVYNTFRAAAPPPPIPAYGGVVYQEPVYGNKLL
hFOX1-iso4	ATAAAAYRGAHLRGRGRTVYNTFRAAAPPPPIPAYGGVVYQDGFYGADI-
hFOX1-iso5	ATAAAAYRGAHLRGRGRTVYNTFRAAAPPPPIPAYGGVVYQDGFYGADI-
hFOX2-iso1	ATTAAAFRGAHLRGRGRTVYGAVRA-VPPTAIPAYPGVVYQDGFYGADL-
hFOX2-iso2	ATTAAAFRGAHLRGRGRTVYGAVRA-VPPTAIPAYPGVDMQPTDMHSLLL
hFOX2-iso3	ATTAAAFRGAHLRGRGRTVYGAVRA-VPPTAIPAYPGVVYQDGFYGADL-
hFOX2-iso4	ATTAAAFRGAHLRGRGRTVYGAVRA-VPPTAIPAYPGVDMQPTDMHSLLL
hFOX2-iso5	ATTAAAFRGAHLRGRGRTVYGAVRA-VPPTAIPAYPGVVYQDGFYGADL-
hFOX2-iso6	ATTAAAFRGAHLRGRGRTVYGAVRA-VPPTAIPAYPGVVYQDGFYGADL-
hFOX3	TGTAVAYRGAHLRGRGRAVYNTFRAAAPPPPIPTYGAVVYQDGFYGAEI-
	: :*. :*****:*. :. ** **..*:* . * *

hFOX1-iso1	QGGYAAAYRYAQPTPATA-----AAYSDSYGRVYAA-DPY
hFOX1-iso2	QGGYAAAYRYAQPTPATA-----AAYSDSYGRVYAA-DPY
hFOX1-iso3	QGGYAAAYRYAQPTPATA-----AAYS DRNQFVFVADEI
hFOX1-iso4	YGGYAAAYRYAQPTPATA-----AAYSDSYGRVYAA-DPY
hFOX1-iso5	YGGYAAAYRYAQPTPATA-----AAYSDSYGRVYAA-DPY
hFOX2-iso1	YGGYAAAYRYAQPATATAATAA-----AAAAAAYS DGYGRVYTA-DPY
hFOX2-iso2	QPQPPLLQPLQPLTVTVMAGCTQPTPTMPLPLPLAMELALWRVYTE-VAT
hFOX2-iso3	YGGYAAAYRYAQPATATAATAA-----AAAAAAYS DGYGRVYTA-DPY
hFOX2-iso4	QPQPPLLQPLQPLTVTVMAGCTQPTPTMPLPLPLAMELALWRVYTE-VAT
hFOX2-iso5	YGGYAAAYRYAQPATATAATAA-----AAAAAAYS DGYGRVYTA-DPY
hFOX2-iso6	YGGYAAAYRYAQPATATAATAA-----AAAAAAYS DGYGRVYTA-DPY
hFOX3	YGGYAAAYRYAQPA A-A-----AAYSDSYGRVYAAADPY
	. : ** .. . * . *:.

hFOX1-iso1	-----HHALAPAPTYGVGAMNAFAPLTD AKTRSHADD
hFOX1-iso2	-----HHALAPAPTYGVGAM-----
hFOX1-iso3	SCNTSAVTDEFMLPTPTTTHLLQPPPTALV-----
hFOX1-iso4	-----HHALAPAPTYGVGAMNAFAPLTD AKTRSHADD
hFOX1-iso5	-----HHALAPAPTYGVGAMNAFAPLTD AKTRSHADD
hFOX2-iso1	-----HALAPAASYGVGAV-----
hFOX2-iso2	-----ADLP PTE-----
hFOX2-iso3	-----HALAPAASYGVGAV-----
hFOX2-iso4	-----ADLP PTE-----
hFOX2-iso5	-----HALAPAASYGVGAV-----
hFOX2-iso6	-----HALAPAASYGVGAV-----
hFOX3	-----HHTIGPAATYSIGT-----
	: *.

hFOX1-iso1	VGLVLSSLQASIYRGGYNRFAPY
hFOX1-iso2	-----ASIYRGGYNRFAPY
hFOX1-iso3	-----P
hFOX1-iso4	VGLVLSSLQASIYRGGYNRFAPY
hFOX1-iso5	VGLVLSSLQASIYRGGYNRFAPY
hFOX2-iso1	-----ASLYRGGYSRFAPY
hFOX2-iso2	-----VT
hFOX2-iso3	-----ASLYRGGYSRFAPY
hFOX2-iso4	-----VT
hFOX2-iso5	-----ASLYRGGYSRFAPY
hFOX2-iso6	-----ASLYRGGYSRFAPY
hFOX3	-----M

CLUSTAL W (1.83) multiple sequence alignment (comparison human isoforms, 1-5,
published A2BP1 sequence)

```
iso-1      MLASQGVLLHPYGVPMIVPAAPYLPGLIQGNQEAAAAPDTMAQPYASAQF
iso-2      MLASQGVLLHPYGVPMIVPAAPYLPGLIQGNQEAAAAPDTMAQPYASAQF
iso-3      MLASQGVLLHPYGVPMIVPAAPYLPGLIQGNQEAAAAPDTMAQPYASAQF
iso-4      MNCERE-----QLRGNQEAAAAPDTMAQPYASAQF
iso-5      MNCERE-----QLRGNQEAAAAPDTMAQPYASAQF
pub         -----MAQPYASAQF
                        *****
```

```
iso-1      APPQNGIPAEYTAPHPHPAPEYTGQTTVPEHTLNLYPPAQTHSEQSPADT
iso-2      APPQNGIPAEYTAPHPHPAPEYTGQTTVPEHTLNLYPPAQTHSEQSPADT
iso-3      APPQNGIPAEYTAPHPHPAPEYTGQTTVPEHTLNLYPPAQTHSEQSPADT
iso-4      APPQNGIPAEYTAPHPHPAPEYTGQTTVPEHTLNLYPPAQTHSEQSPADT
iso-5      APPQNGIPAEYTAPHPHPAPEYTGQTTVPEHTLNLYPPAQTHSEQSPADT
pub         APPQNGIPAEYTAPHPHPAPEYTGQTTVPEHTLNLYPPAQTHSEQSPADT
                        *****
```

```
iso-1      SAQTVSGTATQTDDAAPTGDGQPQTQPSSENTENKSQPKRLHVSNI PFRFRD
iso-2      SAQTVSGTATQTDDAAPTGDGQPQTQPSSENTENKSQPKRLHVSNI PFRFRD
iso-3      SAQTVSGTATQTDDAAPTGDGQPQTQPSSENTENKSQPKRLHVSNI PFRFRD
iso-4      SAQTVSGTATQTDDAAPTGDGQPQTQPSSENTENKSQPKRLHVSNI PFRFRD
iso-5      SAQTVSGTATQTDDAAPTGDGQPQTQPSSENTENKSQPKRLHVSNI PFRFRD
pub         SAQTVSGTATQTDDAAPTGDGQPQTQPSSENTENKSQPKRLHVSNI PFRFRD
                        *****
```

```
iso-1      PDLRQMFGQFGKILDVEIIFNERGSKGFGFVTFENSADADRAREKLHGTV
iso-2      PDLRQMFGQFGKILDVEIIFNERGSKGFGFVTFENSADADRAREKLHGTV
iso-3      PDLRQMFGQFGKILDVEIIFNERGSKGFGFVTFENSADADRAREKLHGTV
iso-4      PDLRQMFGQFGKILDVEIIFNERGSKGFGFVTFENSADADRAREKLHGTV
iso-5      PDLRQMFGQFGKILDVEIIFNERGSKGFGFVTFENSADADRAREKLHGTV
pub         PDLRQMFGQFGKILDVEIIFNERGSKGFGFVTFENSADADRAREKLHGTV
                        *****
```

```
iso-1      VEGRKIEVNNATARVMTNKKTVNPTYNGWKLNPVVGAVYSPEFYAGTVLL
iso-2      VEGRKIEVNNATARVMTNKKTVNPTYNGWKLNPVVGAVYSPEFYAGTVLL
iso-3      VEGRKIEVNNATARVMTNKKTVNPTYNGWKLNPVVGAVYSPEFYAGTVLL
iso-4      VEGRKIEVNNATARVMTNKKTVNPTYNGWKLNPVVGAVYSPEFYAGTVLL
iso-5      VEGRKIEVNNATARVMTNKKTVNPTYNGWKLNPVVGAVYSPEFYA-----
pub         VEGRKIEVNNATARVMTNKKTVNPTYNGWKLNPVVGAVYSPEFYAGTVLL
                        *****
```

```
iso-1      CQANQEGSSMYSAPSSLVYTSAMPGFFPYPAATAAAAYRGAHLRGRGRTVY
iso-2      CQANQEGSSMYSAPSSLVYTSAMPGFFPYPAATAAAAYRGAHLRGRGRTVY
iso-3      CQANQEGSSMYSAPSSLVYTSAMPGFFPYPAATAAAAYRGAHLRGRGRTVY
iso-4      CQANQEGSSMYSAPSSLVYTSAMPGFFPYPAATAAAAYRGAHLRGRGRTVY
iso-5      -----VPGFFPYPAATAAAAYRGAHLRGRGRTVY
pub         CQANQEGSSMYSAPSSLVYTSAMPGFFPYPAATAAAAYRGAHLRGRGRTVY
                        : *****
```

```
iso-1      NTFRAAAPPPPIPAYGGVVYQEPVYGNKLLQGGYAAARYAQPTPATAAAY
iso-2      NTFRAAAPPPPIPAYGGVVYQEPVYGNKLLQGGYAAARYAQPTPATAAAY
iso-3      NTFRAAAPPPPIPAYGGVVYQEPVYGNKLLQGGYAAARYAQPTPATAAAY
iso-4      NTFRAAAPPPPIPAYGGVVYQDGFYGADI-YGGYAAARYAQPTPATAAAY
iso-5      NTFRAAAPPPPIPAYGGVVYQDGFYGADI-YGGYAAARYAQPTPATAAAY
pub         NTFRAAAPPPPIPAYGGVVYQDGFYGADI-YGGYAAARYAQPTPATAAAY
                        *****: .** .: *****
```

```

iso-1      SDSYGRVYAA-D-----PYHHALAPAPTYGVGAMNAF
iso-2      SDSYGRVYAA-D-----PYHHALAPAPTYGVGAM---
iso-3      SDRNQFVFVAADEISCNTSAVTDEFMLPTPTTTHLLQPPPTALV-----
iso-4      SDSYGRVYAA-D-----PYHHALAPAPTYGVGAMNAF
iso-5      SDSYGRVYAA-D-----PYHHALAPAPTYGVGAMNAF
pub        SDSYGRVYAA-D-----PYHHALAPAPTYGVGAMNAF
          **      *:. * *      .      * * * . * * *

```

```

iso-1      APLTDAKTRSHADDVGLVLSSLQASIYRGGYNRFAPY
iso-2      -----ASIYRGGYNRFAPY
iso-3      -----P
iso-4      APLTDAKTRSHADDVGLVLSSLQASIYRGGYNRFAPY
iso-5      APLTDAKTRSHADDVGLVLSSLQASIYRGGYNRFAPY
pub        APLTDAKTRSHADDVGLVLSSLQASIYRGGYNRFAPY

```

Homo sapiens RNA binding protein, fox-1 homolog (C. elegans) 3 (RBFox3), mRNA
NCBI Reference Sequence: NM_001082575.1
Ref NM_001082575.1 Homo sapiens RNA binding protein, fox-1 homolog (C. elegans) 3 (RBFox3), mRNA

```

CAAAGCAGACCTCACCTCCCACTACCAGCTTGAAGTCACAGCAGCCAGAGGAAATTCTGCCACCATTTTC
CCAGGTCTGCAGCCCTCCAGCTGGGAACCTGCTCCTGGAGCCATCCCTCTGCAAACAGAGAGCCAGAG
TGCCTCGGGGAAAATTGGCTGAATAAAAGAGCGATCAGGACGCCACGGCTCCGCCTGAAGCGATGGCCCCA
GCCCTACCCCCCGCCCGCAGTACCCCCCTCCGCCACAGAACGGCATCCCTGCCGAGTACGCCCCGCCCCCA
CCGCACCCACGCAGGACTACTCCGCCAGACCCCGGTCCCCACAGAGCATGGCATGACCCTGTACACAC
CAGCACAGACCCACCCGAGCAGCCAGGCTCCGAGGCCAGCACACAGCCCATCGCCGGGACCCAGACAGT
GCCGCAGACAGACGAGGCGGCACAGACGCCAGCCGCTCCACCCCTCCGACCCTACAGAGAAGCAG
CAGCCCAAGCGGCTACACGTCTCCAAACATCCCCCTCCGGTTTCAGGGACCCGACTTGCGGCAAATGTTTCG
GGCAATTTCGAAAAATTTTAGACGTGGAGATCATTTTAAACGAGCGGGGCTCCAAGGGTTTGGGTTTGT
AACTTTTGAAGTACGTGCTGACCGAGCCCGGAGAGCTGAATGGGACGATCGTAGAGGGACGG
AAAATTGAGGTCAATAATGCCACGGCCCGAGTGATGACCAAGAAGACGGGGAAACCCCTACCAACG
GCTGGAAGCTAAATCCAGTGGTTCGGCGCAGTCTACGGGCCTGAATTCTATGCAGTGACGGGGTTCCTTA
CCCCACACCGGCACAGCCGTTGCCATCCGGGGCGCACATCTTCGGGGCCGGGGCCGGGCGTGTATAAT
ACATTCGGGGCTGCGCCACCCCAACCCCATCCGACTTACGGAGCGGTCGTGTATCAGGATGGATTTT
ATGGTGCTGAGATTTATGGAGGCTACGCAGCCTACAGATACGCTCAGCCCGCTGCAGCGGCGGCAGCCTA
CAGCGACAGTTACGGCAGAGTCTACGCAGCTGCCGACCCGTACCATCACACCATCGGGCCCGCGGCGACCC
TACAGCATTTGGAACCATGTGAACCTTCCACCGTTTCTTCTCGGACCATGAAGGGCAAAAACAAAAAA
CAAAAAAATCACAAAACAAAAAAACAAAAAAGATGTTAAGATCCAAGCAACAAAAAACCACCAACC
AAACCAAGAGGCATCCAACCAAGTCCAAGTCCCGCGTCTGGCCACACGCCCCGACCGAGGGAGCACGCC
GGCAGGGGCGCCGAGGAGCGGCCCCAGGACAGGACGGCCCCACCGCGTCTTGGCTGGCAGCACAGTGGGA
ACACGCCCTCCGTCTCAGGCAGTGGGGGAGTTGGAGGGGAAGGGGCCCTCCCTTGTGGGACCCGTGGGG
GCTCTGTTTTCCATCCAGTCTTCTTTCCAGCCCCCAACTCCCAAGACAGACAGTGTGGAGCCAGCGG
CGGCGGAGCAGGCCCCGGGCTGAGCAGGCAGGCGCTGCTAGCAAGACTTGATCTTTGTGGCCAGCTGTGC
CAGGGGGCCCGGGGCTGAGGGGTGCGGGCAGCTTTCATCCAGGGGCTCCACTGGGCCCCGTACCCCT
CCTGTTCGCGTCCCCTGCGTCCCACCTCCCTCCTGCCCGGCAGTCCCGCCCGTGCCTCCAGCCTGGCGAGG
AAGCCGTCCAACAGTAGCCCCGGGGCCAGTCCCAACAGAAAGGGCTGACGTGGCTCCAGGACTCAGGGG
CGCTCCATGGGAGGACGAAGGAAGCCAGCCAGCCAGGAGCCACTCCTCACACCTCCAAGTGTGGCCAAG
TGGGCCCTGAGGCCAAGGACTTACTTGCTCTTCCCTGGCCATCTCTCCCTTTCTGGAGGAGGCCCGGGCC
TGTGTACACCAAGGCTGACCTCGTGTCTGCTGCTGGGACCCAGCCCTCCCTGCCGCTCCCCTGTGAGCCC
AGTCCACCGTGGGCGCCAGGGCCAGGACGGGCCAGCGCCCGGCTGCATCGCGAGGTTGGGAGTCACAG
TGGCTGTGGGCCTGGACGGGCACAGCCAGAGCAGGGGCCCATGGGAAGGGCAAGGGATGGGGAAGCCTGG
GCCGGCCCCCTTCCCTGCTCCCAAGGCAGGTGTCCAGGTGGCGGGAGCAGCACCAGGACAGCCAGGCTTA
CCCGTGGGAGGAGCAGGAGCAGAGCAGGTGGCAGGGAGGAACCCCTGGCGAGGCAGGAGCACTGAAGT
AGGGAAGCAGCAAAAAATACAGGCTCCCAACGTGGCTCCACTGTCTCATGAAGTGTCAAAAATTTAAAAA
TACACCTCACTTTCTATTACAGCATCAGCTATTGAAATGGAATTCTCCTTTTCTATTCCCGTTGTACATAG
CCCCACGCCCTGCCTCCGGCTTTGTCTCTGTACAGAGCCCCCTGTCCCCTCTGCTGTTCCGGACCCCTT
TCTTGCAGCAGCTCAACCCCCGACTCACTCAGATCCCCAGGACTGCAGCCGAGCCCCGGGCTTCCTTTC
TTACCATTTCTGTATGCTTCCAAGGTGTGACCATTCAACTAACAGTATTATTAAGATTATTAATAAAGAT
TTCTTTCTTCAACACAGGAAAAAAAAAAAAAAAAAAAA

```

Homo sapiens RNA binding protein, fox-1 homolog (C. elegans) 1 (RBFOX1), transcript variant 1, mRNA

NCBI Reference Sequence: NM_145891.2

Ref NM_145891.2 Homo sapiens RNA binding protein, fox-1 homolog (C. elegans) 1 (RBFOX1), transcript variant 1, mRNA

ACCTTCCATTTTGGTAGCTTCAACTTTGCAGCTGCTGTGTCTCTTCGTCGTCTGGGAAGAAAACCTTTCAC
ATTAAGAGTTGCTGATGCGGATTTTCTTTCTTCTCTCCCGCGTTGATGAGTGCTTGGCTCCTGACAG
AAGGGATTTGGCTCCCAGCTTTGTAGTTTCGGAAGAAGTTGGGTCTATAGATTTCCCCCTAACTCTCCATT
GATGTGTTGAGCTTCAGAGGGAATAATAACTCTACGTAAAGCATGCTGGCGTCTCAAGGAGTTCTCTGC
ATCCTTATAGGCGCTATGATTGTACCGCAGCTCCTTACCTTCTGGACTGATTCAGGGTAATCAGGA
AGCAGCCGCTGCCCCTGACACAATGGCTCAGCCTTACGCTTCGGCCAGTTTGCTCCCCCGCAGAACGGT
ATCCCCGCGGAATACAGGCCCTCATCCCCACCCGCGCCAGAGTACACAGGCCAGACCACGGTTCCCCG
AGCACACATTAAACCTGTACCCTCCCGCCCAGACGCACTCCGAGCAGAGCCCGCGGACACGAGCGCTCA
GACCTCTCTGGCACCGCCACACAGACAGATGACGCAGCACCGACGGATGGCCAGCCCCAGACACAACCT
TCTGAAAACACGGAACAAGTCTCAGCCCAAGCGCTGCATGTCTCCAATATCCCCTTCAGGTTCCGGG
ATCCGGACCTCAGACAAATGTTTGGTCAATTTGGTAAAAATCTTAGATGTTGAAATTATTTTAAATGAGCG
AGGCTCAAAGGATTTGGTTTCGTAACCTTCGAAAAATAGTGCCGATGCGGACAGGGCGAGGGAGAAATTA
CACGGCACCGTGGTAGAGGGCCGTAATAATCGAGGTAAATAATGCCACAGCACGTGTAATGACAAATAAAA
AGACCGTCAACCTTTATACAAATGGCTGGAATTTGAATCCAGTTGTGGGTGCAGTCTACAGTCCCGAATT
CTATGCAGGCACGGTCTGTGTGTCAGGCCAACAGGAGGGATCTTCCATGTACAGTGCCCCCAGTTCA
CTTGATATATACTTCTGCAATGCCAGCTTCCCGTATCCAGCAGCCACCGCCGCGCCGCTACCGAGGGG
CGCACCTGCGAGGCCGCGGTGCGACCGTGTACAACCTTCAGGGCCGCGCGCCCCCGCCCCGATCCC
GGCCTACGGCGGAGTAGTGTATCAAGAGCCTGTGTATGGCAATAAATTGCTGCAGGGTGGTTATGCTGCA
TACCGTACGCCAGCCTACCCCTGCCACTGCCGTGCTTACAGTGACAGTTACGGACGAGTTTATGCTG
CCGACCCCTACCACCACGCACTTGCTCCAGCCCCACCTACGGCGTTGGTGCCATGAATGCTTTTGCACC
TTTGACTGATGCCAAGACTAGGAGCCATGCTGATGATGTGGGTCTCGTTCTTTCTTATTGCAGGCTAGT
ATATACCGAGGGGGATACAACCGTTTGTCTCCATACTAAATGACAAAACCATAAAAACCTTCCAATGTGG
GGAGAAAGGAAGCTTTCCGAGGCCTGAGTATTGCAATACATGCAGTAGTACATCATTTTAGCAACTCTAA
AAAAAAAAAAAAATACAAATAAAAAAGGAAAAAAAAAATTACATTTTTTATCTTATACCTCAGATATTTTGTTC
TGTGTATTTTAAATATTGTGGGTCTTTAATTTCTGAAGGTTCCGTAGTTTGGTTGCTGGCTGTAGGAGTTT
TTGTGGTTAGCTTAGACAGATGCTAGATAATGAATAAAAACTGGTTTAGGGCCATATCCAGAGTGCTATA
TTATGTAATGAATTATATATGCTGAATATTAAGCTACTGGGGTTATCAGCTATTTGGGAAGAGTGTAAG
TGACTACAGTAGTCATTTTTTCTGCACCTGCATTATTTTATTTTGCGAAGGGGAGGTTGGGAGGGGCT
TAGGGGATTGGAACGGGGTTTGGCTGAAAGAAAAAAAAAAAAATGTAAGTATGAATCTAAACGACCCA
CTGCACCAACAATCATTTATCAATGGTTCTAAGTTACTCATTGCCAGTTCAGCCAAAGGTCATGTTGTT
AAGGGGGTCTCTAGTAGCACTTGTGCATCTGAGTTGAATGAAGCTGTGCAACCCACCTTTTAAACCA
TTCCACCCGCGAGTATTCAGGTTCTTAACCAAGTCGCTATTTAGGAAAAAAAAAACCCACTAGTTAGGCCATC
AACAAGCATTCTTTTTATATTTCTTCCAGTATAATAAATTATTGATATCACTGCTGACTTTTTATATTATG
GGAGGAAAAAAAAATAACATTAATAAAAAAGGTGATAAAAAAGCACTGTTTCTATTTTTTTCTTTTTTTCCAA
AAAAAGAAAGTAATAAAAACTTAAATCTTTGTACCAGTTAAAAAAATGTATAAAATTTACATCTGTGC
AGTGGAGTTGTTAAGTTCTAGAAACAGTCTATGAAGCTTTAGTTTTCAGCTAGTAGAACAACCTGTAGAG
ACAGACGTATAATTTTATGGAATTACATGATAATCATATTCGGATTTATAGAAGCATTTTACAAGTATT
GCAATCATTGAGTAGAGATAATCATGGTATTTTTCATCAGCTTGGTACTTTTTGAAACGTGACTGCGTTGT
GTGAACAATCTGCAATTTTTTCAGTCCGTGAGATCCTGCCCTTCCACCTCTTTCCCATACCCCAAGGATT
ATCTCAAAATGATCTCTTTAGTTCTGTCTCCAAGGCCAGGACACTTGTGAGAAGGATGCAAAAAAGAA
AAAAGTACATCCACCTCTTAATCCCAAAAGAACAAAGTCTCCCCACCTTTCTCGGATGCAGGGCCAGAG
TGACACAGCCGAAAAATTGCAGTTTGTCTGTACTTCTGTTTGAACTTTCACGTTGTCTGTTTACAAGT
TAACCTAAGTTGGGGTATCCGTCACGGGTCTTCTGTTTGTATTTAAATAAAAAACAACAGCAGCAGGCT
GTCCCTGAGTAGTTTTGTCTGCCATAGGTTAAGTCCCTCATGTGTACAGTGACAGGCCCTGTGGCCCGCACTT
CAGTAAGTTATCAACTCTCACCGCTGTGAACCTGCCAATCCGCTGTAAACAACCTGCTTTAAAAACAAAAC
CAACAAAAACTTTAAAAAAAATGTGTGATCCAGCTTTCTCTTGCCATCCTATGTGCATGCCGTAAAGATC
AGTTGGATATTAACCATCAATAAAGTTTCACAAGATTTGAAAACAAAGTTTCTGTAGCTTCGATACCTA
AGACAGACGATAGTGATCTTGAAGAGAGAAAGGGGAGGGGGAGCTACTTATCAGCCAAAAGCATATA
AAGTGTCTTTTTACACAGACTTCTTTGGGTGGGTGAGGGGAGGGGCAAGAGGGTGTATATAGCATCACT
AAGACATTCTCATTTCCCCACCTGGAAAACAGTGTTATGGCAATGGGTGCCTGGTTGATGTTCTTAAAGG
AAACGAATTATTAACCACTATGACATCCTCCAGAGGGAAGAAAGAGTAGGAGCAGGGGGCTATGGAGAA
TAAATTTCTCCCAATTGCGCCTCAGATTTCAAAATCCAGAATTTGTATTGTGTTTCGAATCACAAACAT
AGAATTTCTACTGTGTGGTTAAAGTAAATTCATTTGCAGTTTTGATTTTCATCAATGAGCTGTACTTT
CCCCCATGACTGTATGTAGTTTTAATAAAATCATTTAGAGTGAGTGAGTGCCAACCGATGTTGCAGAATC
TTTTGTCTAGGCACTCAAGATGCCAATAAGTCATTTTAAATGTATGTCAGAGATGTAAACAAACATTT

TGGATTTTTTTTAAACAGTATTTATTTGGAATGTTTTCATTTATCTAAATAACTATTGCTATTATGAATT
ATGGAAAATTAATATTATGTGTGGCATATAGTGACTTCTTAACACACACATCACGCAATCTGCAAACCCA
GAAAATGTGTATATCTGTCTTTAGAAATTAGTGTTTATATCACTTACAGTGTTTGTGAATAAAGAAAAC
TGGTTTGTAATATCAAAAAATAAAAGCTTAGTCTGAAAAGG