Rayburn, Lowell Y.M., Semil P. Choksi, and Michael Bender. 2002. Correction of an *amontillado* (*amon*) cDNA artifact and identification of single nucleotide polymorphisms in the *amon* gene. *Dros. Inf. Serv*. 85: 23-24.

Correction of an *amontillado* (*amon*) cDNA artifact and identification of single nucleotide polymorphisms in the *amon* gene.

Rayburn, Lowell Y.M., Semil P. Choksi, and Michael Bender. Department of Genetics, Life Sciences Building, University of Georgia, Athens, GA 30602.

Comparing genomic sequence (Adams *et al*., 2000) with an *amon* cDNA sequence (Siekhaus and Fuller, 1999) showed that the first 43 nucleotides of this cDNA do not align with upstream *amon* genomic sequences. They do, however, match sequences within more 3' areas of the cDNA (see Figure 1A), suggesting that the first 43 nucleotides may be an artifact of cDNA cloning. The 43 nucleotides can be separated into two components. Nucleotides 3-21 (bold type) match those at positions 3268-3286 (bold type) of the cDNA, and nucleotides 18-47 (underlined) reverse complement the sequence from 3225-3254 (underlined) in the 3' end of the cDNA. The Berkeley Drosophila Genome Project (BDGP) database contains 12 *amon* EST sequences, eleven of which extend further 5' than the cDNA reported by Siekhaus and Fuller (1999) and do match genomic sequences. The two *amon* ESTs with the greatest

A

```
   1 CGTGCATTTA CATACATAGT TGTATCCAAA TTAAAGCAGA TTTTGGCAGC GCACAAAGAC
  61 TCCAGACGAA GTACATAGAA ACAGAATTTC AGGCGGAAGG AAAGCACAAC AAGCTCCATA
     ------------------------------------------------------------
3181 TCCCATTATT GCGGAACATG TTTGTCGCTT AACTTTGGTT GCAAGCCAAA ATCTGCTTTA
3241 ATTTGGATAC AACTCGTACG TGTAAAATGC ATTTACATAC ATAGTTTTCC TTTCAGTTTC
3301 TTTCGGGGGC CAAAAAAACC ACAACATAGT AACATTAATT GTGTGAAACC AGATTCACGA
     ------------------------------------------------------------
```

B

```
ATGTATATTA CTTACTTTAC CATGATAGCG TACTTCTAGT TTTGCTAGCG TACTTCTAGT
TTTGTAAGCC TAACAGTAGT AGCCCCTTGG AACATTCCTG TGCACGCCAC TGGCGAAGCT
GGTAAGACAA CGGTATCTCC GCTAGGGGTG CGTCGAGTGA GTGGCGGGTG GCTTATCGCC
TGCTAGTCGG AGCTTTCGAG CGAGCTTCGG CGGCGAAGCT TTTGCCACCG CTTTGGCGCC
TGGCGGACAT TCATAATGGC GTCCGTCGGT ACCGCTGGCG AACAGTACTC AGCGCGCAGC
GCCCCAGTGA AGTTCGTGGC AGCGCACAAA GAC
```

Figure 1. Correction of an *amon* cDNA artifact. A. Sequence of an *amon* cDNA (Siekhaus and Fuller, 1999). Dashes represent sequences not reproduced here. Sequences 5' to the arrow do not match upstream *amon* genomic sequences. The sequence in bold near the 5' end of the cDNA matches the bold sequence near the 3' end in the forward direction. The underlined sequence near the 5' end of the cDNA matches the underlined sequence near the 3' end in reverse complement orientation. Numbering is that of Siekhaus and Fuller (1999). B. Genomic sequence upstream of *amon*. Sequence in bold represents the extent of the two longest *amon* ESTs. Underlined sequences represent matches to consensus Initiator and Downstream Promoter Elements.

5' extent (BDGP clone ID #RE06156 and #RE58333) lengthen *amon* cDNA sequences by 64 nucleotides (Figure 1B). Good matches (4/6 nucleotides, underlined) to a consensus *Drosophila* Initiator sequence at the 5' end of the cDNA and to the *Drosophila* Downstream Promoter Element functional range set (Kutach and Kadonaga, 2000) suggest that these two *amon* ESTs may mark the 5' end of the *amon* message.

Table 1: Polymorphisms in the coding regions of *amon*.

amon Exon	cDNA bp *	*amon* genomic	*amon* cDNA	*red e* genomic	*TM3* genomic	Codon Change
2	664	t	t	t	c	gat --> gac
2	721	t	t	c	t	agt -->agc
3	952	g	a	g	g	agg -->aga
8	1432	c	a	a	c	atc --> ata
8	1456	c	g	c	c	ggc --> ggg
8	1477	t	c	c	t	aat --> aac
8	1489	c	t	t	c	tac --> tat
9	1726	t	t	c	t	cct --> ccg
10	2038	g	a	g	g	ttg --> tta
10	2092	a	a	g	g	gaa --> gag
10	2109	a	a	c	c	aca --> acc
10	2110	a	c	a	a	aaa --> caa
11	2125	c	c	g	c	tcc --> tcg
11	2132	c	c	t	t	ctg --> ttg
11	2176	c	c	t	t	acc --> act
11	2188	t	t	c	c	ttt --> ttc
11	2234	t	t	c	c	ttg --> ctg
12	2320	c	c	g	g	ccc --> ccg
12	2347	g	g	a	a	ccg --> cca

While sequencing EMS induced mutations in *amon*, we identified 19 single nucleotide polymorphisms in *amon* coding sequences. Table 1 shows polymorphisms between *amon* genomic sequence (Adams *et al.*, 2000), an *amon* cDNA (Siekhaus and Fuller, 1999), the *red e* parental chromosome used in our screen (Rayburn *et al.*, in preparation), and two third chromosome balancers (*TM3, Sb* and *TM3, Sb* y^+) obtained from the Bloomington stock center. Because we sequenced genomic DNA amplified from *red e/TM3* heterozygotes, polymorphisms between these chromosomes served as useful positive controls. One of these polymorphisms (a vs. c at position 2110) is predicted to result in a Lysine to Glutamine change at position 562, suggesting that Amon is functional with either Lysine or Glutamine at this position. The remaining 18 polymorphisms are not predicted to result in an amino acid change.

Acknowledgments: This work was supported by an NIH grant (GM53681) to M.B. and an NIH training grant (GM07103) to L.Y.M.R.

References: Adams, M.D., S.E. Celniker, R.A. Holt, C.A. Evans, J.D. Gocayne, *et al*. 2000, Science. 287: 2185-2195; Kutach, A.K., and J.T. Kadonaga 2000, Mol Cell Biol. 20: 4754-4764; Siekhaus, D.E., and R.S. Fuller 1999, J Neurosci. 19: 6942-6954.