Development 122, 1373-1383 (1996)
Printed in Great Britain © The Company of Biologists Limited 1996
DEV6225

lag-1, a gene required for *lin-12* and *glp-1* signaling in *Caenorhabditis elegans*, is homologous to human CBF1 and *Drosophila* Su(H)

Sioux Christensen[1,‡], Voula Kodoyianni[2,5,†], Marcus Bosenberg[2,†], Lisa Friedman[3] and Judith Kimble[1-5,*]

[1]Department of Genetics, [2]Laboratory of Molecular Biology, [3]Department of Biochemistry, [4]Department of Medical Genetics University of Wisconsin-Madison, and [5]Howard Hughes Medical Institute, Madison, WI 53706, USA

*Author for correspondence
‡Present address: Salk Institute for Biological Studies, La Jolla, CA, USA
†These authors contributed equally to this work

SUMMARY

The homologous receptors LIN-12 and GLP-1 mediate diverse cell-signaling events during development of the nematode *Caenorhabditis elegans*. These two receptors appear to be functionally interchangeable and have sequence similarity to *Drosophila* Notch. Here we focus on a molecular analysis of the *lag-1* gene (*lin-12* and *glp-1*), which plays a central role in LIN-12- and GLP-1-mediated signal transduction. We find that the predicted LAG-1 protein is homologous to two DNA-binding proteins: human C Promoter Binding Factor (CBF1) and *Drosophila* Suppressor of Hairless (Su(H)). Furthermore, we show that LAG-1 binds specifically to the DNA sequence RTGGGAA, previously identified as a CBF-1/Su(H)-binding site. Finally, we report that the 5′ flanking regions and first introns of the *lin-12*, *glp-1* and *lag-1* genes are enriched for potential LAG-1-binding sites. We propose that LAG-1 is a transcriptional regulator that serves as a primary link between the LIN-12 and GLP-1 receptors and downstream target genes in *C. elegans*. In addition, we propose that LAG-1 may be a key component of a positive feedback loop that amplifies activity of the LIN-12/GLP-1 pathway.

Key words: *lag-1*, *lin-12*, *glp-1*, *Caenorhabditis elegans*, *Drosophila*, C Promotor Binding Factor (CBF1), Suppressor of Hairless (Su(H))

INTRODUCTION

Two homologous receptors, LIN-12 and GLP-1, regulate a variety of cell interactions during *C. elegans* development. Best understood are the roles of LIN-12 in AC/VU lateral signaling (Greenwald et al., 1983) and GLP-1 in germline induction (Austin and Kimble, 1987) and embryonic induction (Priess et al., 1987; Hutter and Schnabel, 1994; Mango et al., 1994; Mello et al, 1994; Moskowitz et al., 1994). In addition to these well characterized interactions, LIN-12 and GLP-1 control numerous other cell fate decisions during both embryonic and post-embryonic development (Greenwald et al., 1983; Lambie and Kimble, 1991; Newman and Sternberg, 1995; I. Moskowitz and J. Rothman, personal communication). Double mutants lacking the activities of both *lin-12* and *glp-1* die as L1 larvae with an array of cell transformations that result in the loss of the excretory cell, loss of the rectum and a twisted nose. This collection of double mutant defects is called the Lag phenotype, for *lin-12* and *glp-1* (Lambie and Kimble, 1991). Recent studies indicate that the Lag phenotype results from disruption of specific inductive interactions in the 24- to 87-cell embryo (I. Moskowitz and J. Rothman, personal communication). Therefore, certain cell-fate decisions appear to be controlled specifically by LIN-12 (e.g. AC/VU) or GLP-1 (e.g. germline induction), while others make use of either LIN-12 or GLP-1 (Lag defects).

In an effort to identify additonal components of the LIN-12/GLP-1 pathway, a screen for mutants with a phenotype similar to that of the *lin-12 glp-1* double mutants was performed (Lambie and Kimble, 1991). From this screen, two new genes were identified: *lag-1* and *lag-2*. Strong loss-of-function *lag-1* or *lag-2* mutants exhibit a Lag phenotype similar to that of *lin-12 glp-1* double mutant. Because all twenty independent alleles isolated in this screen map to either *lag-1* or *lag-2,* it seems likely that *lag-1* and *lag-2* define central components of the LIN-12 and GLP-1 signal transduction pathways.

Sequence similarities between the LIN-12 and GLP-1 proteins in *C. elegans* and *Drosophila* Notch (Greenwald, 1985; Wharton et al., 1985; Yochem et al., 1988; Austin and Kimble, 1989; Yochem and Greenwald, 1989) suggest that this class of receptors may share a common signaling mechanism. Consistent with this idea, two putative ligands for LIN-12 and GLP-1, called LAG-2 and APX-1, are homologous to the Notch ligands, Delta and Serrate (Henderson et al., 1994; Tax et al., 1994; Mello et al., 1994). The observation that these ligand and receptor families are also present in vertebrates suggests that this pathway may be functionally conserved throughout metazoan phylogeny (reviewed by Artavanis-Tsakonas et al., 1995; Henrique et al., 1995).

In this paper, we focus on a molecular analysis of the *lag-1* gene (Lambie and Kimble, 1991). We find that the predicted LAG-1 protein is homologous to a class of DNA-binding

proteins that includes human CBF1/KBF2/RBP-Jκ (henceforth called CBF1) and *Drosophila* Su(H) (Matsunami et al., 1989; Schweisguth and Posakony, 1992). This homology provides important clues about how *lag-1* might function in LIN-12/GLP-1 signaling. For example, we show that LAG-1 binds specifically to the DNA sequence RTGGGAA, previously identified as a CBF1/Su(H)-binding site (Tun et al., 1994; Brou et al., 1994). Because *lag-1* appears to be required for the majority of cell interactions mediated by LIN-12 and GLP-1 in *C. elegans*, we propose that the LAG-1 DNA-binding protein may be the primary downstream effector for the LIN-12 and GLP-1 receptors. In addition, we suggest that positive feedback in the LIN-12/GLP-1 signaling pathway may rely, at least in part, on LAG-1 binding to sites located in the genomic sequences of the *lin-12*, *glp-1* and *lag-1* genes.

MATERIALS AND METHODS

Strains

C. elegans genetic nomenclature follows guidelines of Horvitz et al. (1979). Most mutations used in this study are described in Hodgkin et al. (1988). Alleles used include: *dpy-13(e184)*; *unc-44(e362, e1260)*; *lag-1(q385, q418, q476)* (Lambie and Kimble, 1991); *smg-3(r929)* (provided by B. Cali and P. Anderson); *unc-8(e15)*; *deb-1(st385)* (provided by R. Barstead and R. Waterston); *unc-24(e138)*. In addition, we used the chromosomal rearrangement, *DnT1*, as a balancer for *lag-1*. The strain *unc-44(e1260) lag-1(q385)/deb-1(st385)* was used for mutant rescue experiments.

Three-factor mapping and positioning of lag-1 on the physical map

lag-1 was mapped to the right of *unc-44* and to the left of *unc-8* by standard three factor mapping: 8/38 Unc non-Deb recombinant self-progeny of an *unc-44 deb-1/lag-1* hermaphrodite segregated Lag homozygotes, while 30/38 recombinants did not, placing *lag-1* to the right of *unc-44* and close to *unc-44* in the *unc-44* to *deb-1* interval. 1/24 Unc non-Dpy recombinant self-progeny of a *dpy-13 unc-8/lag-1* hermaphrodite segregated Lag animals, placing *lag-1* close to and to the left of *unc-8* in the *dpy-13* to *unc-8* interval. Therefore, *lag-1* maps on the right arm of linkage group IV between *unc-44* and *unc-8* (Fig. 1A). Since *unc-44* has been cloned (Otsuka et al., 1995), this position provides a left boundary for *lag-1* on the physical map: *unc-44* coding sequences are present on phage DD#LRF1, which overlaps the left end of YAC Y66E2 (Fig. 1B).

To define a right boundary for *lag-1* on the physical map, we mapped *lag-1* with respect to the *rP8* polymorphism; *rP8* is detected by cosmid F58F9, which spans the right end of Y66E2 (Cali, 1995). From hermaphrodites of genotype *unc-44 lag-1(+) rP8 smg-3 unc-24/lag-1(q385)*, non-Unc-44 Smg-3 Unc-24 progeny were selected and used to establish lines carrying the recombinant chromosome. The progeny of each recombinant were examined for presence of *lag-1(q385)* and DNA prepared from each recombinant line was analyzed by Southern blot for the polymorphism *rP8* using F58F9 as probe. Of eleven recombinant lines analyzed, two were homozygous for both *lag-1(+)* and *rP8*. The remaining nine recombinant lines were all heterozygous for *lag-1(q385)*, DNA; of these nine lines, four carried *rP8*, while five did not. Therefore, *lag-1* maps to the left of *rP8*. Together these data place *lag-1* in an interval of ~360 kb spanned by a single YAC, Y66E2.

Mutant rescue

Germline transformation was performed using standard techniques (Mello et al., 1991). Cosmids located between DD#LRF1 and F58F9 were injected (3 μg/ml) together with pRF4 (50-100 μg/ml) as a dominant marker for germline transformation (Mello et al., 1991) into *unc-44 lag-1(q385)/deb-1* heterozygotes; the broods of transgenic animals were scored for viable *unc-44* progeny. From a pool of four cosmids that rescued *lag-1* (Fig. 1B), individual cosmids were assayed for rescue as above. In addition, JK#L78 (described in DNA sequencing and analysis section) and subclones of both JK#L78 and the cosmid F43A12 were tested for rescuing activity as described above.

Isolation of cDNA clones

pJK525 was used to screen approximately 240,000 plaques from each of two *C. elegans* cDNA libraries (kindly provided by P. Okkema). One full-length cDNA, pJK526, was isolated from the mixed stage library, while one partial cDNA, pJK527, was recovered from the embryonic library.

5′ end determination

Total RNA was prepared from mixed stage wild-type *C. elegans* and poly(A)$^+$ RNA was selected using the Oligo-tex kit (Qiagen) according to manufacturer instructions. First strand cDNAs were produced using a themostable reverse transcriptase (Epicenter) at 70°C following manufacturers directions using a *lag-1*-specific primer, SKC37 (5′GCGACATCACTGTGGCTACT3′). Second strand synthesis was performed using the RACE-1 primer (5′GACTCGAGTCGACATCGA(T)$_{17}$3′) as described by Frohman et al. (1988). cDNAs were PCR amplified using the RACE-1 primer and a nested *lag-1*-specific primer, SKC27 reverse (5′CGACTGGAACATTGCTTCTG3′). The primary amplification product was reamplified using the RACE-2 primer (5′GACTCGAGTCGAACTCG3′) (Frohman et al., 1988) and the *lag-1*-specific primer SKC31 (5′ATCCCAGGTTGGCTGCGGTG3′). Amplification products were reamplified using a primer complementary to the transpliced leader sequence SL1 (5′CTCAAACTTGGGTAATTAAACC3′) and SKC31. The resulting PCR product was ligated into the pT7Blue vector (Novagen) and sequenced.

The 5′ end of the *lag-1* cDNA was amplified from an independent *C. elegans* cDNA library (kindly provided by R. Barstead) using the vector-derived primer MB9 (5′GGAGCGATTTGCAGGCATTTGCTC3′) and a *lag-1*-specific primer SKC32 (5′GGTGGTGGGCAGAAGAATCT3′). The primary PCR product was reamplified with a nested, vector-derived primer, MB10 (5′GGCATGCCGGTAGAGGTGTGGTCA3′) and a second *lag-1*-specific primer SKC27reverse (see above). The PCR product was cloned into pT7Blue and sequenced to determine the 5′ end.

Northern blot

Northern blots were performed as described (Sambrook et al., 1989) using approximately 4 μg poly (A)$^+$ mRNA per lane. The northern blot was probed with a full-length cDNA insert excised from pJK526 and labeled using the Prime-a-Gene kit (Promega) according to manufacturers directions.

PCR walking

To clone genomic sequences between the end of the genomic phage clone JK#L78 and the first exon of *lag-1*, total genomic DNA was amplified using an oligo complementary to sequences in the first exon of *lag-1*, SKC38 (5′GAAATGCCTCTCGCCTACTC3′) and an oligo complementary to sequences within the first intron of *lag-1*, SKC-BSreverse (5′CAGGGACGCATCATCTTGTGTG3′), located near the 5′ end of JK#L78. The resulting PCR fragment was cloned into pT7Blue and sequenced.

To isolate genomic sequence upstream of the first *lag-1* exon, total genomic DNA was digested overnight with *Cla*I and cloned into Bluescript (Stratagene). The ligation mixture was PCR amplified using the T7 primer (Promega) and the *lag-1*-specific primer SKC39 (5′GGAAACTGTGCGTGCACGTG3′). The resulting PCR products were gel purified, cloned into pT7Blue and sequenced.

DNA sequencing and analysis

All sequencing reactions were performed using Sequenase Version 2.0 (USB) according to manufacturers' directions. The cDNA inserts from phage clones pJK526 and pJK527 were subcloned into Bluescript vectors and both strands of each insert were sequenced. Genomic fragments containing the *lag-1* locus, as identified by Southern blot analysis (data not shown), were subcloned from cosmid F43A12 (pJK525) or phage JK#L78 (pJK550, pJK551, pJK552) (see Fig. 1C) and sequenced as described for cDNAs except that intron sequences were determined by sequencing a single DNA strand. There is perfect agreement between the proposed coding sequences of the cDNA and genomic clones. Approximately 99.5% of the genomic sequence has been unambiguously determined. Genomic and cDNA sequences were aligned using the Align program (DNASTAR) to establish the exon-intron structure of *lag-1*. LAG-1, Su(H) and CBF1 sequences were compared using the MegAlign program of DNASTAR.

Identification of lag-1 mutations

DNA was prepared from 5-10 *lag-1* homozygous mutant larvae and fragments corresponding to *lag-1* exons were generated by PCR, cloned into pT7 Blue and sequenced using standard procedures (Sambrook et al., 1989). For each allele, PCR products covering all coding sequences, the majority of the 3′ UTR and the splice junctions of flanking introns were sequenced. Two or more clones were sequenced for each region; exons containing nucleotide changes were confirmed by sequencing at least two additional independent clones. Oligonucleotide primers used to amplify each exon are listed below; numbers in parentheses indicate the position of the 5′ end of each primer in the genomic sequence, with position 1 corresponding to the first nucleotide of the cDNA clone pJK526.

Exon 1: SKC E0.1- 5′CGATTTCTCTTCGTTTCCACTG3′ (-21); SKC E0.2-5′CCGATTTAACGATTCGAAAGCG3′ (134).

Exons 2 & 3: SKC E1.1- 5′GCTCCCTTTTTGACACGTCCC3′ (7037); SKC E1.2- 5′CCTTTCACTGGGCAACATGAC3′ (7389).

Exons 4 & 5: SKC E2.1- 5′CGGAGACCTGTAATTTCC3′ (8250); SKC E2.2- 5′CAGTTGGAACACATTTGC3′ (8756).

Exons 6 & 7: SKC E3.1- 5′GAAATTCCAGCGGTCAAGTAG3′ (11683); SKC E4.2- 5′CTGCCTGCTCGAAAAGAACC3′ (12211).

Exon 8: SKC E5A.1- 5′CGAACATATGGTCGTTAACG3′ (13170); SKC E5A.2- TTCGCGGCACAATAATCCTG3′ (13695).

Exons 9 & 10: SKC E5.5- 5′GGAAAAGTTCGACATCCTGG3′ (13574); SKC E5B.2- 5′TGCTGGCAAATCTGAATTCC3′ (14210).

Exon 11: SKC E6.1- 5′GTGATGTTGAGCAATGAGAA3′ (14551); SKC E6.2- 5′CTGAATTCAGAACTCTATTCTC3′ (15164).

Exon 12: SKC E7.2- 5′GGTATTTCGAAGTCAAATTCC3′ (15591); SKC-12- 5′AGGGAGATACCACCGTGTGA3′ (15993).

DNA binding

The plasmid pCLA48-673 was constructed by subcloning *lag-1* sequence coding for amino acids 48-673 into the *Eco*RV site of pCITE-4a(+) (Novagen). 2 μl of rabbit reticulocyte lysate or reticulocyte lysate programmed with pCLA48-673 were incubated with 0.1 pMol ^{32}P-end-labeled MB22 oligonucleotide in the absence or presence of competitor oligonucleotides and subjected to electrophoretic mobility shift assay (EMSA). Shifted bands were quantitated using a PhosphorImager and ImageQuant software (Molecular Dynamics). Buffer conditions, EMSA and in vitro transcription/translation were as described (Zimber-Strobl et al., 1994). 5 pMol competitor oligonucleotide was present unless otherwise specified. The sequences of MB20, MB21, MB22 and MB24 are shown in Fig. 5B.

RESULTS

Cloning *lag-1*

The *lag-1* locus was initially positioned on the genetic and physical maps to an interval defined by *unc-44* and the polymorphism *rP8* on chromosome IV (see Materials and Methods) (Fig. 1A,B). This region is spanned by a single YAC, Y66E2, of approximately 360 kb (Fig. 1B). To locate *lag-1* within Y66E2, cosmids underlying the YAC were tested for rescue of *lag-1(q385)* homozygotes by germline transformation (see Materials and Methods). Whereas *lag-1(q385)* homozygotes die as first stage larvae (Lambie and Kimble, 1991), transgenic *lag-1(q385)* homozygotes carrying either of two overlapping cosmids, M02G2 and F43A12, survived to adulthood (Fig. 1B). Although the somas of rescued animals have a wild-type morphology, the germlines fail to proliferate, a phenotype similar to that of weak *lag-1* and *glp-1* null mutants. Since *C. elegans*, transgenes are often poorly expressed in the germline (Mello and Fire, 1995), the inability of transgenic *lag-1* to rescue the germline defect may indicate that *lag-1* expression is required in the germline for continued mitotic proliferation. Alternatively, both cosmids that rescue the somatic *lag-1* defects may lack regulatory elements required for germline rescue.

Restriction mapping and Southern blot analysis of M02G2 and F43A12 indicated that the largest genomic fragment common to both cosmids was a 7 kb *Hin*dIII fragment. However, this fragment was not sufficient to rescue the *lag-1* mutant phenotype. To identify *lag-1* within F43A12, the cloned 7 kb *Hin*dIII fragment, pJK525 (Fig. 1C), was used as a probe to isolate phage from a genomic library. One phage, JK#L78, which contained the 7 kb *Hin*dIII fragment and flanking sequences, rescued *lag-1* somatic defects in a transgenic assay. Subclones of JK#L78, (pJK550, pJK551 and pJK552, Fig. 1C), were generated and tested for rescuing activity; none of these subclones was able to rescue *lag-1(q385)* homozygotes. We therefore used pJK525 to isolate cDNAs. Two cDNAs were isolated and sequenced: pJK526 is 2550 nt long and is likely to be full length (see below, Fig. 2), while pJK527 is 1854 nt long and appears to be truncated at the 5′ end. The 3′ ends of pJK526 and pJK527 were identical and the clones are colinear along their common length. By northern blot analysis using pJK525 as a probe, a single band of approximately 2.5 kb was detected in poly(A)$^+$ mRNA prepared from mixed stage animals (Fig. 3), consistent with the idea that pJK526 corresponds to the full-length transcript.

Exon/intron structure of *lag-1*

To deduce the exon/intron structure of *lag-1*, the genomic *lag-1* region was sequenced (see Materials and Methods). We first sequenced ~ 14.2 kb of genomic sequence derived from genomic subclones pJK525, pJK550, pJK551and pJK552 (Fig. 1C). Comparison of the genomic and cDNA sequences revealed at least 12 exons (11 within pJK#L78 and one at the 5′ end found in the cDNA, but not in the phage) (Fig. 1C); introns ranged in size from 48 nt to 6117 nt. All exon/intron boundaries are flanked by conserved *C. elegans* splice junction sequences (Emmons, 1988). Because the phage clone JK#L78 did not contain the first exon or 5′ flanking sequences, a PCR walk was used to obtain the 5′ portion of *lag-1*. In this way, genomic DNA corresponding to the first exon plus approxi-

mately 1.2 kb of its 5′ flanking region was cloned and sequenced (Fig. 1C).

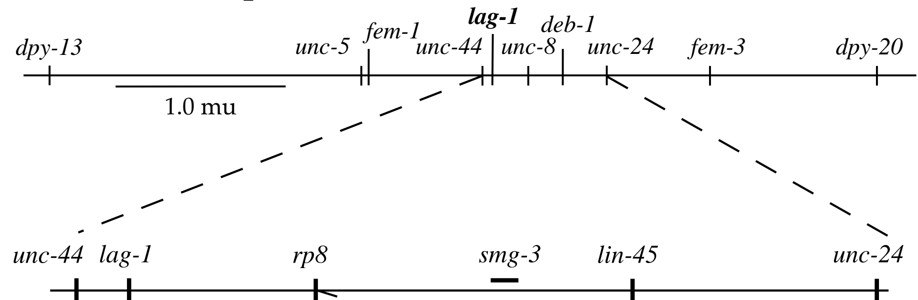
The genomic *lag-1* sequence contains a consensus *C. elegans* splice acceptor sequence (TTTCAG) immediately upstream of the first nucleotide of the longer cDNA clone, pJK526. This may indicate either that the *lag-1* message is transpliced in vivo, or that pJK526 represents a truncated transcript. To distinguish between these possibilities, the 5′ ends of two additional *lag-1* cDNA clones were examined. First, PCR was used to amplify the 5′-end of *lag-1* from a random primed *C. elegans* cDNA library; such libraries do not require transcription of the full-length cDNA and thus are more likely to contain complete 5′ cDNA sequences. Second, the product of a *lag-1*-specific RACE reaction was amplified with a primer complementary to the *C. elegans* transpliced leader sequence SL1. For both of the resulting clones, pJK533 (random primed library) and pJK534 (RACE reaction), the 5′ end corresponded to that identified in pJK526. Furthermore, pJK534 contained SL1 sequences immediately upstream of the first nucleotide predicted by pJK526, suggesting that the *lag-1* message can be transpliced in vivo.

Sequence analysis of the predicted LAG-1 protein

Conceptual translation of the full-length *lag-1* cDNA predicts a protein of 673 amino acids with an M_r ~74×10^3. Searches of protein databases revealed a striking homology between LAG-1 and the mammalian CBF1 and *Drosophila* Su(H) proteins (Fig. 4A) (Matsunami et al., 1989; Furukawa et al., 1991; Schweisguth and Posakony, 1992); we suggest that proteins of this class be called CSL proteins, for CBF1, Su(H) and LAG-1. Over a region of 412 amino acids, LAG-1 is 60% identical (71% similar) to human CBF1 and 62% identical (72% similar) to *Drosophila* Su(H) (Fig. 4B). Two potential nuclear localization sequences (Chelsky et al., 1989) are found within the predicted LAG-1 protein and a consensus MAP kinase site (P X S/T P) (Clark-Lewis et al., 1991) is conserved among the CSL proteins (Ser^{560} of LAG-1)(Fig. 2). Within the conserved region, four intron splice sites are identical between *lag-1* and the corresponding human and mouse genes (Fig. 2, filled arrowheads) (Amakawa et al., 1993). The N terminus and C terminus of the *C. elegans, Drosophila* and vertebrate proteins are variable in length and are not conserved (Fig. 4B).

Molecular identification of *lag-1* mutants

To confirm the molecular identification of *lag-1* and to begin exploring the functional domains of the LAG-1 protein, we sequenced the coding regions of three strong loss-of-function *lag-1* alleles. All three contain single nucleotide changes within the *lag-1* coding region, confirming identity of the locus as *lag-1* (Fig. 4B). Two mutations introduce premature stop codons: *lag-1(q418)* is predicted to truncate LAG-1 N-terminal to the conserved region, while *lag-1(q385)* removes only the C-terminal 48 residues. In *C. elegans*, mRNAs containing premature stop codons are often unstable; such mRNAs can be



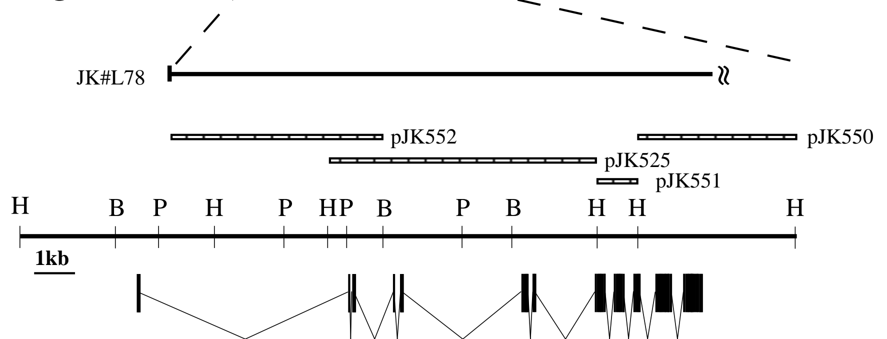


Fig. 1. Genetic and molecular identification of *lag-1*. (A) Genetic map of the *lag-1* region on the right arm of linkage group IV showing the position of *lag-1* with respect to markers used in three-factor mapping of the *lag-1* locus. The expanded region below shows the position of flanking genes and the polymorphism *rP8* used to correlate the genetic and physical maps in the *lag-1* region. (B) Physical map between *unc-44* and *rP8*. DD#LRF1 is a genomic phage clone containing part of *unc-44*; the cosmid clone F58F9, detects the RFLP polymorphism *rP8*. The results of mutant rescue experiments using various cosmid and phage are shown to the left of the relevant DNAs. (C) The *lag-1* gene. *Above,* solid line indicates the sequenced portion of JK#L78, including its breakpoint in the first *lag-1* intron. pJK525 is the 7 kb subclone used to screen genomic and cDNA lambda libraries (see text). The pJK525, pJK550, pJK551 and pJK552 subclones cover most of the genomic *lag-1* locus (see Results) H, *Hin*dIII; B, *Bsi*WI; P, *Pst*I. *Below*, exon-intron structure of *lag-1* based on comparison of genomic and cDNA sequence. Filled boxes represent exons; lines represent introns.

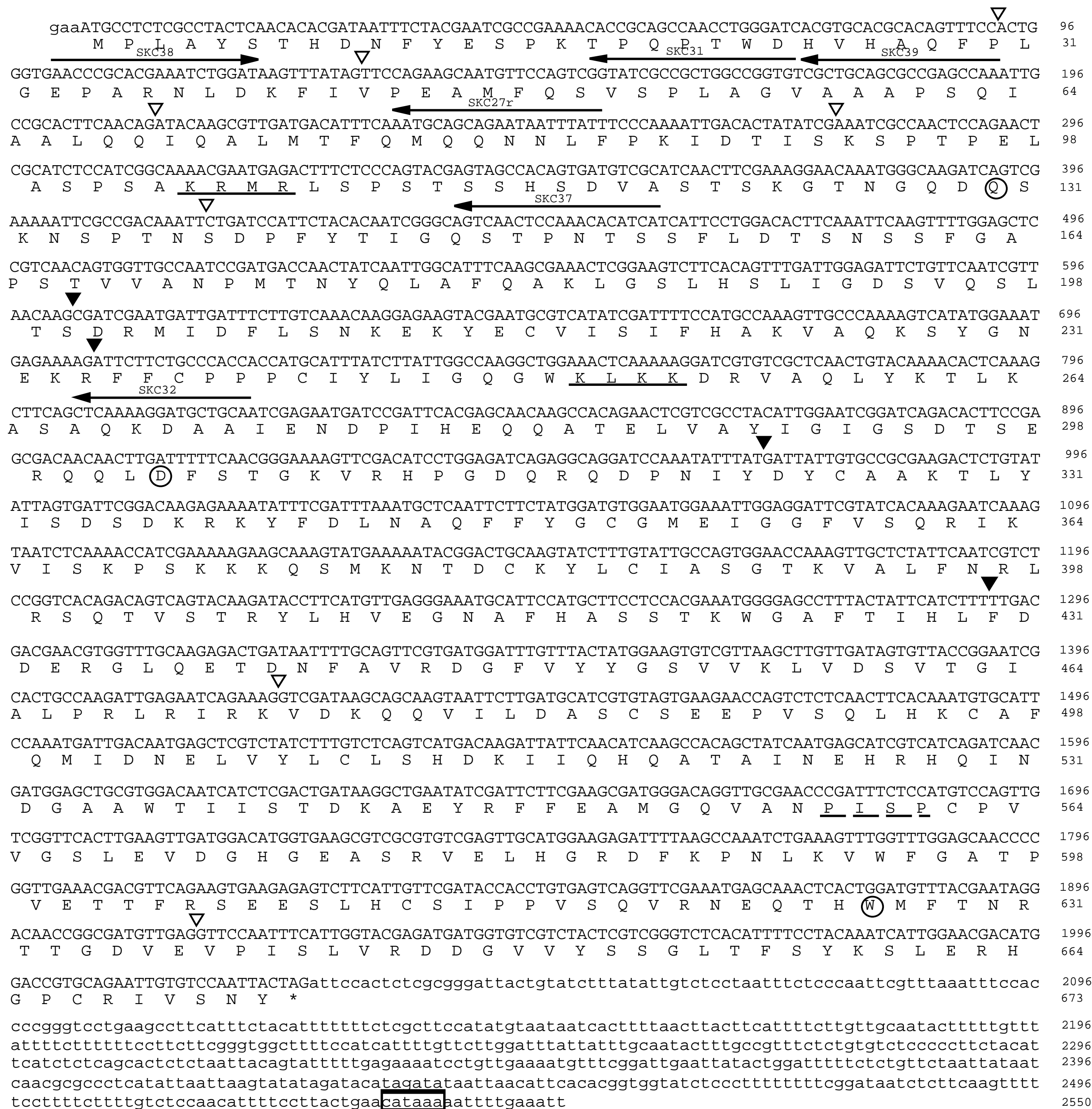


Fig. 2. The *lag-1* cDNA sequence and predicted protein. Nucleotide sequence of the *lag-1* cDNA and the predicted amino acid sequence of LAG-1 protein. cDNA numbering begins at the first nucleotide of pJK526, which is likely to be nearly full length, missing only its transpliced leader (see text). Amino acid numbering begins at the first methionine. Lower case letters represent predicted 5′ and 3′ untranslated regions; upper case letters represent predicted coding sequence. Oligonucleotide primers described in the text are shown below the corresponding sequence. Open arrowheads indicate positions of the eleven introns; filled arrowheads mark those splice sites conserved between *lag-1* and its mouse and human homologues (Amakawa et al., 1993). Potential nuclear localization sequences are indicated by solid underlines. The potential MAP kinase site is idicated by a dashed underline. Amino acids altered in three *lag-1* mutant alleles are circled (see Fig. 4A and text). The asterisk indicates the predicted termination codon. The proposed cleavage and polyadenylation signal is boxed. On both cDNA clones, a stretch of poly(A) begins immediately after the last nucleotide shown.

stabilized in a *smg* mutant background (Hodgkin et al., 1989). We therefore examined the phenotype of a *smg-1; lag-1(q385)* double mutant, but found no change from the *lag-1(q385)* single mutant phenotype. Therefore, the strong *lag-1(q385)* phenotype is likely to result from loss of LAG-1 protein activity rather than from message instability. In addition to these two nonsense

3.5 kb –
1.75 kb –
poly (A)+

Fig. 3. Northern blot. A band of ~2.5 kb is observed in poly(A)$^+$ enriched mRNA using the pJK526 insert as a probe. The positions of the 28S and 18S ribosomal RNA bands are indicated.

mutants, which establish the *lag-1* null phenotype by molecular criteria, we identified one missense mutation, *lag-1(q476)*, in which a semi-conserved aspartic acid at position 303 is converted to histidine. The mutant phenotype of *lag-1(q476)* is indistinguishable from that of the two nonsense mutants, suggesting that this amino acid replacement eliminates LAG-1 activity.

DNA binding by LAG-1

The homology of LAG-1 to CBF1 and Su(H) provides important clues to its function. CBF1 can activate transcription in combination with EBNA2 (Waltzer et al., 1994; Henkel et al., 1994; Grossman et al., 1994; Zimber-Strobl et al., 1994) and can repress transcription in certain cases (Hsieh and Hayward, 1995; Dou et al., 1994). Both CBF1 and Su(H) bind a defined DNA sequence, GTGGGAA, in vitro (Tun et al., 1994; Brou et al., 1994). To determine whether LAG-1 may have a similar activity, we tested the ability of LAG-1 to bind DNA containing the CBF1-binding site GTGGGAA as well as the related sequence ATGGGAA. Mobility shift assays were performed with in vitro translated LAG-1 and MB22, a radioactively labeled double-stranded oligonucleotide derived from the

A

```
198 L T S D R M I D F L S N K E K Y E C V I S I F H A K V A Q K S Y G N E K R F F C P P P C I Y L I G Q G W K L K K D R V A lag-1
105 L T R D A M E K Y M - - R E R N D M V I V I L H A K V A Q K S Y G N E K R F F C P P P C I Y L F G S G W R R R Y E E M L Su(H)
31  L T R E A M R N Y L - - K E R G D Q T V L I L H A K V A Q K S Y G N E K R F F C P P P C V Y L M G S G W K K K K E Q M E human

                                                                                           ●
258 Q L Y K T L K A S A Q K D A A I E N D P I H E Q Q A T E L V A Y I G I G S D T S E R Q Q L D F S T G K V R H P G D Q R Q lag-1
163 Q - - - - - Q G E G E - - - - - - - - - - - - - Q G A Q L C A F I G I G S S D Q D M Q Q L D L N - G K - - - - - - - - - Su(H)
89  R - - - - - D G C S E - - - - - - - - - - - - - Q E S Q P C A F I G I G N S D Q E M Q Q L N L E - G K - - - - - - - - - human

318 D P N I Y D Y C A A K T L Y I S D S D K R K Y F D L N A Q F F Y G C G M E I G G F V S Q R I K V I S K P S K K K Q S M K lag-1
195 - - - - - Q Y C A A K T L F I S D S D K R K H F M L S V K M F Y G N G H D I G V F N S K R I K V I S K P S K K K Q S L K Su(H)
121 - - - - - N Y C T A K T L Y I S D S D K R K H F M L S V K M F Y G N S D D I G V F L S K R I K V I S K P S K K K Q S L K human

378 N T D C K Y L C I A S G T K V A L F N R L R S Q T V S T R Y L H V E G N A F H A S S T K W G A F T I H L F D D E R G L Q lag-1
250 N A D - - - L C I A S G T N V A L F N R L R S Q T V S T R Y L H V E N G H F H A S S T Q W G A F T I H L L D D N E S - - Su(H)
176 N A D - - - L C I A S G T K V A L F N R L R S Q T V S T R Y L H V E G G N F H A S S Q Q W G A F F I H L L D D D E S - - human

438 E T D N F A V R D G F V Y Y G S V V K L V D S V T G I A L P R L R I R K V D K Q Q V I L D A S C S E E P V S Q L H K C A lag-1
305 E S E E F Q V R D G Y I H Y G A T V K L V C S V T G M A L P R L I I R K V D K Q M A L L E A - - - D D P V S Q L H K C A Su(H)
231 E G E E F T V R D G Y I H Y G Q T C K L V C S V T G M A L P R L I I R K V D K H T A L L D A - - - D D P V S Q L H K C A human

498 F Q M I D N E L V Y L C L S H D K I I Q H Q A T A I - - N E H R H Q I N D G A A W T I I S T D K A E Y R F F E A M G Q V lag-1
362 F Y M K D T D R M Y L C L S Q E K I I Q F Q A T P C P K E P N K E M I N D G A C W T I I S T D K A E Y Q F Y E G M G P V Su(H)
288 F Y L K D T E R M Y L C L S Q E R I I Q F Q A T P C P K E P N K E M I N D G A S W T I I S T D K A E Y T F Y E G M G P V human

556 A N P I S P C P V V G S L E V D G H G E A S R V E L H G R D F K P N L K V W F G A T P V E T T F R S E E S L H C S I P P lag-1
422 A S P V T P V P I V N S L N L N G G G D V A M L E L S G D N F T P H L Q V W F G D V E A E T M Y R C T E T L L C V V P E Su(H)
348 L A P V T P V P V V E S L Q L N G G G D V A M L E L T G Q N F T P N L R V W F G D V E A E T M Y R C G E S M L C V V P D human

                        ■
616 V S Q V R N E Q T H W M F T N R T T G D V E V P I S L V R D D G V V Y S S G L T F S Y K S L - E R H G P C lag-1
482 I S Q F R G E - - - W L W V R Q P T - - - Q V P I S L V R N D G I I Y A T G L T F T Y T P E P G P R P H C Su(H)
408 I S A F R E G - - - W R W V R Q P V - - - Q V P V T L V R N D G I I Y S T S L T F T Y T P E P G P R P H C human
```

B

N C CBF1
N C Su(H)
N C LAG-1
q418 Glu130 → TAG
q476 Asp 303 → His
q385 Trp 593 → TAG

Fig. 4. Sequence similarity between LAG-1, Drosophila Su(H) and human CBF1. (A) Comparison of amino acid sequences in the highly conserved region of LAG-1, Su(H) (Furukawa et al., 1991; Schweisguth and Posakony, 1992) and CBF1 (Amakawa et al., 1993). Amino acid identities are shaded. Dashes represent gaps introduced to maximize alignment; the LAG-1 protein has several insertions within the N-terminal half of the conserved region. To indicate mutations, *lag-1(q476)* is marked by a filled circle and *lag-1(q385)* by a filled square. (B) Conserved and non-conserved regions within LAG-1, Su(H) and CBF1. Shaded portions indicate the conserved core shown in part (A); open regions represent the non-conserved N-terminal and C-terminal regions. Molecular lesions associated with each of three *lag-1* mutations are depicted

	Sequence	% Reduction pMol Competitor 0.1	1.0	10.0
MB20	gatcGGC**ATGGGAA**CC CCG**TACCCTT**GGctag	40	85	98
MB22	gatcGAC**GTGGGAA**AG CTG**CACCCTT**TCctag	47	91	98
MB24	gatcGAC**G**AC**GGAA**AG CTG**C**TG**CCTT**TCctag	0	2	8
MB21	TCGTGGAGGAAGCTGGAATG AGCACCTCCTTCGACCTTAC	0	0	3

Fig. 5. Specific DNA binding by LAG-1. (A) Electrophoretic mobility shift assay (EMSA) with in vitro translated LAG-1. A shifted complex appears in the presence of LAG-1 (lane 2), which is reduced by a 50 fold excess of competitors containing the sequence GTGGGAA (MB22) (Lane 3) or ATGGGAA (MB20) (lane 4), but not by a competitor lacking a putative binding site (MB21) (lane 5). (B) Comparison of relative DNA binding by LAG-1 to naturally occurring sequence elements. The sequence of the double-stranded competitors is shown with the putative binding site in bold. Binding is expressed as percent reduction of LAG-1 bandshift with 0.1 pMol ^{32}P-end labeled MB22 oligonucleotide in the presence of varying amounts of competitor oligonucleotide. Double-stranded oligonucleotides were based on the consensus binding site defined for RBP-Jk/CBF1 (Tun et al., 1994) and were derived from: the LCS1 region of the *lin-12* promoter (Wilkinson et al., 1994) (MB20), the second *lin-12* intron (MB22) (Yochem et al, 1988), a mutated form of MB22 lacking an RTGGGAA site (MB24) and *lag-1* sequence that does not contain a putative binding site (MB21).

second intron of *lin-12*. The sequences of MB22 and competitor oligonucleotides are shown in Fig. 5B.

LAG-1 bound to MB22, which contains the consensus CBF1-binding site GTGGGAA (Fig. 5A, lane 2). Binding was competed by addition of excess MB22 (Fig. 5A,B) or excess MB20, which contains the related sequence ATGGGAA (Fig. 5A,B). The MB20 oligonucleotide is derived from the LCS1 region of the *lin-12* promoter, which is required for rescue of *lin-12* null mutants (Wilkinson et al., 1994). Binding was not competed by either of two control oligonucleotides: MB21 lacks an RTGGGAA (Fig. 5A,B) and MB24 is identical to MB22 except that two bases within the conserved binding site are altered (Fig. 5B). We conclude that all CSL proteins may bind similar sequences to control transcription of downstream genes.

Identification of LAG-1-binding sites in the genomic sequences of *lag-1, lin-12 and glp-1*

The distribution of potential LAG-1-binding sites in the genomic sequences of *lin-12*, *glp-1* and *lag-1* reveals an intriguing pattern (Fig. 6). If distributed randomly, one RTGGGAA site would be expected in 4 kb on average. Within 15 kb spanning *glp-1,* 9 sites were found; within 18 kb spanning *lin-12,* 20 sites were found; and within 16 kb spanning *lag-1,* 18 sites were observed (Fig. 6; Table 1). Furthermore, in *lin-12, glp-1* and *lag-1*, these sites are preferentially located in 5′ flanking regions and in large introns at the 5′ end of the genes. Particularly striking is the presence of 12 potential LAG-1-binding sites within a 2.4 kb region in the first intron of *lag-1*. 6 of these sites occur as tandem repeats with similar spacing, orientatation and conserved linker sequences (Table 1). This region of the first intron is present in the rescuing phage, pJK#L78, and may provide regulatory sequences required for *lag-1* activity. In contrast, upon examination of the genomic sequences of *lag-2, ced-4, tra-2,* and *xol-1,* sites were either not present (*lag-2* and *ced-4*) or found at a random frequency (*tra-2* and *xol-1*). In the latter case, potential LAG-1-binding sites were not preferentially located in predicted regulatory regions. A search of all available *C. elegans* genomic sequence using FindPatterns (Devereux et al., 1984) reveals potential LAG-1 sites distributed at the expected frequency of approximately 1 site every 4 kb. Furthermore, no unique sequence was found with a cluster of LAG-1-binding sites as dense (12 sites in 2.4 kb) as that observed in the first *lag-1* intron.

Fig. 6. Potential LAG-1-binding sites in genes required for LIN-12/GLP-1 signaling. Filled triangles denote the presence and orientation of potential LAG-1-binding sites (RTGGGAA) in *glp-1*, *lin-12*, *lag-1* and *lag-2*; boxes represent exons, lines represent non-coding regions. The asterisk denotes the position of the LAG-1-binding site in the LCS1 region of *lin-12*. It is possible that LAG-1 may bind to additional, non-consensus sites.

DISCUSSION

LAG-1 is a conserved central component in signaling by both LIN-12 and GLP-1

Only two genes, *lag-1* and *lag-2,* were identified in extensive mutageneses designed to find loci that participate in both *lin-12*- and *glp-1*-mediated signaling (Lambie and Kimble, 1991). Therefore, *lag-1* and *lag-2* are believed to encode core components of the LIN-12/GLP-1 pathway. This prediction has been confirmed for *lag-2*, which encodes a homolog of Delta and is believed to function as a ligand for GLP-1 and LIN-12 (Henderson et al., 1994; Tax et al., 1994). In this paper, we

Table 1. Potential LAG-1 binding sites in *lin-12*, *glp-1* and *lag-1*

Site	Location*	Direction	Sequence	Position
lin-12				
1	-5421	→	TTC**GTGGGAA**TTG	promoter
2	-5208	→	TTT**ATGGGAA**TTT	promoter
3	-4283	←	TAA**ATGGGAA**CCA	promoter
4	-3664	→	TAA**ATGGGAA**GTT	promoter
5	-2854	→	CCA**GTGGGAA**ATG	promoter
6	-2091	←	TTA**GTGGGAA**ATT	promoter
7	-2060	←	ATA**ATGGGAA**AAT	promoter
8	-1666	←	GGC**ATGGGAA**CCC	promoter
9	-822	←	AGA**GTGGGAA**CGG	promoter
10	733	←	ATG**GTGGGAA**CCC	intron 1
11	892	→	GTT**ATGGGAA**AGT	intron 1
12	1107	→	GCC**ATGGGAA**AAC	intron 1
13	1831	←	GCA**ATGGGAA**GTG	intron 2
14	1894	←	CCT**GTGGGAA**TTG	intron 2
15	3039	→	GAC**GTGGGAA**AGT	intron 2
16	3712	→	GTC**GTGGGAA**ACG	intron 3
17	4400	→	CTT**GTGGGAA**ACC	intron 3
18	5198	←	TCG**ATGGGAA**ACT	intron 4
19	6785	→	TCG**GTGGGAA**ATG	exon 5
20	7085	→	AAA**ATGGGAA**AAG	exon 5
glp-1				
1	-5517	←	GAA**ATGGGAA**ATG	promoter
2	-3940	←	GTT**ATGGGAA**GTT	promoter
3	-3203	→	ACA**GTGGGAA**ATG	promoter
4	-3196	→	GAA**ATGGGAA**ACG	promoter
5	-2329	→	GTC**ATGGGAA**CCT	promoter
6	-1581	←	TGC**GTGGGAA**ATT	promoter
7	242	→	GTG**GTGGGAA**ACA	intron 1
8	368	→	AGA**ATGGGAA**CAC	intron 1
9	2607	→	TTG**ATGGGAA**ACA	exon 4
lag-1				
1	674	←	AGG**ATGGGAA**AGT	intron 1
2	689	←	CGT**GTGGGAA**AAA	intron 1
3	781	→	GAA**ATGGGAA**GGC	intron 1
4	1632	←	CGT**ATGGGAA**ACT	intron 1
5	1765	←	AGC**GTGGGAA**AGT	intron 1
6	1780	←	GTG**GTGGGAA**AAA	intron 1
7	2079	←	CGA**GTGGGAA**AAA	intron 1
8	2096	←	ATC**GTGGGAA**AAA	intron 1
9	2297	→	TCT**GTGGGAA**TAC	intron 1
10	2405	→	CTG**GTGGGAA**GAT	intron 1
11	2917	←	TCG**GTGGGAA**AAT	intron 1
12	3078	←	ATC**ATGGGAA**CCT	intron 1
13	5802	→	GTG**GTGGGAA**TAG	intron 1
14	7840	←	GAC**GTGGGAA**AAT	intron 5
15	8092	→	GTA**GTGGGAA**CTT	intron 5
16	9001	←	CGG**GTGGGAA**AAC	intron 5
17	9119	→	TCG**GTGGGAA**TCT	intron 5
18	11371	→	AGC**GTGGGAA**CTA	intron 7
lag-1 tandem sites				
a	689, 674	←	CGT**GTGGGAA**AAACTA--GG**ATGGGAA**AGT	
b	1780,1765	←	GTG**GTGGGAA**AAACGA--GC**GTGGGAA**AGT	
c	2096, 2079	←	ATC**GTGGGAA**AAAGCGACGA**GTGGGAA**AAA	

*The numbering of the *lag-1* sequence is described in Materials and Methods. For *lin-12* and *glp-1*, the A in the initiation codon is denoted as base pair 1. LAG-1 binding sites were numbered according to the position of the R in the RTGGGAA sequence element.

report that this prediction also holds true for *lag-1,* which appears to encode the primary downstream effector for LIN-12/GLP-1-mediated interactions. Furthermore, LAG-1 bears a striking sequence similarity to vertebrate CBF1 and *Drosophila* Su(H), which suggests that at least one downstream effector of the LIN-12/Notch/GLP-1 signaling pathway is conserved.

LAG-1 controls most signaling interactions mediated by LIN-12 and GLP-1

The *lag-1* gene is required for numerous cell interactions mediated by LIN-12 and GLP-1. The larval lethality of *lag-1* null alleles demonstrates that LAG-1 is required for those embryonic cell-fate decisions relying on expression of either *lin-12* or *glp-1*; furthermore, the defect in germline proliferation of weak *lag-1* mutants indicates that LAG-1 is required for germline induction (Lambie and Kimble, 1991). However, no vulval phenotype has been observed in weak *lag-1* alleles (Lambie and Kimble, 1991), even though *lin-12* regulates lateral signaling in two types of cell interactions required for vulval development (AC/VU and VPCs) (Greenwald et al., 1983). The lack of vulval defects in *lag-1* mutants may reflect a lower threshold requirement for *lag-1* in these cell interactions, or it may indicate that *lag-1* does not participate in *lin-12-mediated* cell-fate decisions affecting vulval development.

A maternal requirement for *lag-1* is suggested by the embryonic lethality of progeny derived from animals homozygous for either of two weak alleles, *lag-1(q416)* (Christensen, 1995) and *lag-1(om13)* (Qiao et al., 1995). The effects of stronger *lag-1(lf)* mutations on the early embryo cannot be examined, because homozygotes either die as larvae or become sterile adults. The finding that *lag-1* antisense RNA injections in the adult hermaphrodite gonad results in embryonic lethality (V. Kodoyianni, unpublished) is consistent with the idea that *lag-1* is required maternally for embryogenesis.

Although the full extent of *lag-1* function remains to be determined, LAG-1 activity is clearly crucial for numerous cell interactions during *C. elegans* development and is therefore one of the key players in both LIN-12 and GLP-1 signaling. We suggest that the two LAG-1 homologs, vertebrate CBF1 and *Drosophila* Su(H), may also play a broad role in signaling by their counterpart receptors. Indeed, recent evidence suggests that Su(H) has a broader role than previously suspected in *Drosophila* development (Lecourtois and Schweisguth, 1995) and that CBF1 may participate in signaling by the vertebrate Notch pathway (Oka et al., 1995).

CSL proteins link receptor activity to target genes during signal transduction

Two major conclusions have emerged during the past year concerning the function and regulation of CSL proteins (for CBF1, Su(H) and LAG-1). First, they appear to be transcriptional regulators, either activators (Grossman et al., 1994; Henkel et al., 1994; Waltzer et al., 1994; Zimber-Strobl et al., 1994; Bailey and Posokony, 1995; Lecourtois and Schwiesguth, 1995) or repressors (Dou et al., 1994; Hsieh and Hayward, 1995). Second, they appear to link the Notch receptor with downstream target genes (Fortini and Artavanis-Tsakonas, 1994; Bailey and Posokony, 1995; Jarriault et al., 1995; Lecourtois and Schwiesguth, 1995). Su(H) protein binds the intracellular domain of Notch in yeast two-hybrid experiments and Notch activation appears to regulate the translocation of Su(H) from cytoplasm to nucleus (Tamura et al., 1995; Fortini and Artavanis-Tsakonas, 1994). In addition, Su(H) transcriptionally regulates the expression of target genes in the Enhancer of Split cluster (E(Sp)-C), in response to Notch activation (Bailey and Posakony, 1995; Lecourtois and Schwiesguth, 1995). Similarly, CBF1 mediates activation of expression of HES-1, the mammalian homolog of E(Spl), in the presense of activated

Notch1 (Jarriault et al., 1995). Therefore, although the molecular details of how activation of Notch leads to activation of Su(H) are not yet understood, CBF1 and Su(H) clearly link Notch and Notch-related homologs with transcriptional activation of E(Spl)-C related genes.

Several lines of evidence suggest that the function and regulation of LAG-1 may be similar to that of CBF1 and Su(H). First, LAG-1 binds to the consensus CBF1/Su(H)-binding site RTGGGAA (this paper). Second, LAG-1 appears to act downstream of the LIN-12/GLP-1 receptors: genetically, a weak *lag-1* mutation is epistatic to a *glp-1* gain-of-function mutation (L. W. Berry and T. Schedl, personal communication) and molecularly, *lag-1* mRNA is present in the mitotic region of the germline, the receiving tissue of a *glp-1*-mediated interaction (V. Kodoyianni, unpublished data). However, downstream target genes of LIN-12/GLP-1 signaling have not yet been identified in *C. elegans*.

LAG-1 may mediate positive feedback in the LIN-12/GLP-1 signaling pathways

Positive feedback has been suggested to regulate both in GLP-1 inductive signaling (Kodoyianni et al., 1992) and LIN-12 lateral signaling (Seydoux and Greenwald, 1989; Wilkinson et al., 1994). For induction processes, feedback was hypothesized because of an all-or-none phenotype in mutants carrying either of two unusual *glp-1* alleles: within a single worm, one germline arm may exhibit no proliferation, while the other may be indistinguishable from wild type. One interpretation of this all-or-none phenotype relies on positive feedback in the GLP-1 signaling pathway: partially active GLP-1 may reach a threshold sufficient to establish positive feedback, which then ensures continuous *glp-1* expression or activity (Kodoyianni et al., 1992) . For lateral signaling events, positive feedback was hypothesized from a bias in cell fate choice that was observed in *lin-12* mosaics but not wild-type animals (Seydoux and Greenwald, 1989). In *Drosophila*, a positive feedback mechanism has also been hypothesized to play a role in the *Notch* signaling pathway (Heitzler and Simpson, 1991).

Positive feedback appears to occur at the transcriptional level in the case of *lin-12* (Wilkinson et al., 1994). Specifically, a 67 bp region in the 5′ flanking region of *lin-12* is required to maintain *lin-12* expression in the VU precursor cell during the AC/VU decision. This 67 bp region, LCS1 (*lin-12* conserved sequence) was identified by conservation with *C. briggsae lin-12* and contains a LAG-1-binding site. Deletion of LCS1 from a *lin-12* transgene results in a reduction of its ability to regulate the AC/VU decision, but not other *lin-12*-mediated processes. Therefore, the LCS1 element appears to specifically control positive feedback for the AC/VU decision.

Examination of the *lin-12, glp-1 and lag-1* genomic sequences reveals a greater number of potential LAG-1-binding sites than would be predicted on a random basis (Fig. 6). Furthermore, the LAG-1-binding sites are clustered in the 5′ flanking regions and large first introns of these genes, including one in LCS1. Indeed, MB20, one of the oligonucleotides used to compete LAG-1 DNA-binding, is derived from LCS1. Based on the non-random distribution of potential LAG-1-binding sites in *glp-1*, *lin-12 and lag-1,* we suggest that these genes may be transcriptionally regulated by LAG-1 in a positive feedback loop (Fig. 7). However, the importance of LAG-1-binding sites for positive feedback in the LIN-12/GLP-1 pathway remains to be tested.

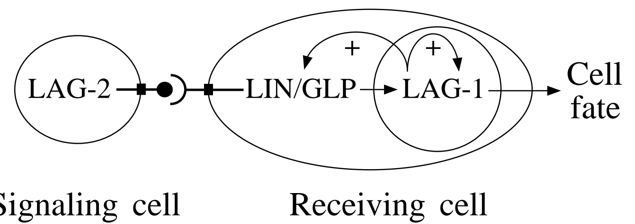


Fig. 7. Model for positive feedback regulation by the LAG-1 DNA-binding protein. Activation of GLP-1/LIN-12 may affect the ability of LAG-1 to transcriptionally up regulate its own synthesis, as well as that of the GLP-1/LIN-12 receptors.

Specificity in the LIN-12/GLP-1 signaling pathway

The central role of LAG-1 in signaling by both LIN-12 and GLP-1 suggests that LAG-1 itself does not confer specificity in the cellular responses to activation of these receptors. Similarly, specificity does not appear to depend on which receptor or ligand is used: LIN-12 and GLP-1 are functionally interchangeable (Mango et al., 1991; Lambie and Kimble, 1991; Fitzgerald et al., 1993) as are their predicted ligands LAG-2 and APX-1 (Gao and Kimble, 1995; Fitzgerald and Greenwald, 1995). Thus, specificity is likely to be generated by cell- or tissue-specific proteins that may interact with the core components of the pathway. For example, the CSL proteins may direct such proteins to specific promoters. Consistent with this idea, the LCS1 region is relatively large (67 nt) and appears to be specific for the AC/VU decision (Wilkinson et al., 1994). Therefore, as yet unidentified factors may work together with LAG-1 at the LCS1 regulatory site to regulate the *lin-12* promoter during lateral signaling between the AC and VU cells. As an extension of this idea, we suggest that the number of potential LAG-1-binding sites in the *glp-1, lin-12 and lag-1* genomic sequences may reflect the complexity of transcriptional regulation required to express these genes in specific cells and at specific times during development. Molecular analyses of individual sites, initiated with LCS1, will clarify the importance of these sites and their individual functions during cell signaling.

Extent of conservation in Notch-related pathways

LAG-1 is likely to provide the primary link between the LIN-12 and GLP-1 receptors and downstream genes in *C. elegans*, and a similar role may be played by Su(H) in *Drosophila* and by CBF1 in vertebrates (Fortini et al., 1994; Bailey and Posakony, 1995; Jarriault et al., 1995; Lecourtois and Schweisguth, 1995). Therefore, the Delta-like ligands, Notch-like receptors and CSL proteins are core components of the LIN-12/Notch/GLP-1 pathway that have been conserved throughout phylogeny. In *Drosophila* and vertebrates, regulation of the E(Spl) complex by the CSL proteins is also conserved (Bailey and Posakony, 1995; Jarriault et al., 1995; Lecourtois and Schweisguth, 1995). In nematodes, no connection has yet been made between LAG-1 and a nematode homolog of E(Spl), and therefore the extent to which the pathway is conserved downstream of *lag-1* is not yet clear.

We are grateful to A. Coulson and J. Sulston for providing cosmids; C. White for the genomic library and P. Okkema and B. Barstead for cDNA libraries; B. Cali for providing the RFLP mapping strain and

collaboration in recombinant analysis. We thank L. Berry and T. Schedl, and I. Moskowitz and J. Rothman for communicating unpublished results, Leanne Olds for her expertise with illustration, all the members of the Kimble laboratory for encouragement and discussions, and Lisa Kadyk for comments on the manuscript. This work was supported in part by grants from the NIH and the Council for Tobacco Research; M. B. is a Jane Coffin Childs Postdoctoral Fellow; J. K. is an investigator with the Howard Hughes Medical Institute.

REFERENCES

Amakawa, R., Jing, W., Ozawa, K., Matsunami, N., Hamaguchi, Y., Matsuda, F., Kawaichi, M. and Honjo, T. (1993). Human Jκ recombination signal binding protein gene (IGKJRB): comparison with its mouse homologue. *Genomics* **17**, 306-315.

Artavanis-Tsakonas, S., Matsuno, K. and Fortini, M. E. (1995). Notch signaling. *Science* **268**, 225-232.

Austin, J. and Kimble, J. (1987). *glp-1* is required in the germ line for regulation of the decision between mitosis and meiosis in *C. elegans. Cell* **51**, 589-599.

Austin, J. and Kimble, J. (1989). Transcript analysis of *glp-1* and *lin-12*, homologous genes required for cell interactions during development of *C. elegans. Cell* **58**, 565-571.

Bailey, A. M. and Posakony, J. W. (1995). Supppressor of Hairless directly activates transcripton of *Enhancer of split* Complex genes in response to Notch receptor activity. *Genes Dev.* **9**, 2609-2622.

Brou, C., Logeat, F., Lecourtois, M., Vanderkerckhove, J., Kourilsky, P., Schweisguth, F. and Israel, A. (1994). Inhibition of the DNA-binding activity of *Drosophila* Suppressor of Hairless and of its human homolog, KBF2/RBP-Jκ, by direct protein-protein interaction with *Drosophila* Hairless. *Genes Dev.* **8**, 2491-2503.

Cali, B. (1995). Analysis of nonsense-mediated mRNA decay in *Caenorhabditis elegans*. (Ph. D. Thesis)

Chelsky, D., Ralph, R. and Jonak, G. (1989). Sequence requirements for synthetic peptide-mediated translocation to the nucleus. *Mol. Cell Biol.* **9**, 2487-2492.

Christensen, S. (1995). The genetic and molecular characterization of the *C. elegans* signal-transduction gene *lag-1*. (Ph.D. Thesis)

Clark-Lewis, I., Sanghera, J. S. and Pelech, S. L. (1991). Definition of a consensus sequence for peptide substrate recognition by p44^{mpk}, the meiosis activated myelin basic protein kinase. *J. Biol. Chem* **266**, 15180-15184.

Devereux, J., Haeberli, P. and Smithies, O. (1984). A comprehensive set of sequence analysis programs for the VAX. *Nucleic Acids Research* **12**, 387-395.

Dou, S., Zeng, X., Cortes, P., Erdjument-Bromage, H., Tempst, P., Honjo, T. and Vales, L. D. (1994). The recombination signal sequence-binding protein RBP-2N functions as a transciptional repressor. *Mol. Cell Biol.* **14**, 3310-3319.

Emmons, S. W. (1988). The Genome. In The Nematode *Caenorhabditis elegans*, W. B. Wood, ed. (Cold Spring Harbor: Cold Spring Harbor Laboratory), pp. 47-79.

Fitzgerald, K., Wilkinson, H. and Greenwald, I. (1993). *glp-1* can substitute for *lin-12* in specifying cell fate decisions in *Caenorhabditis elegans*. *Development* **119**, 1019-1027.

Fitzgerald, K. and Greenwald, I. (1995). Interchangeablility of *Caenorhabditis elegans* DSL proteins and intrinsic signalling activity of their extracellular domains in vivo. *Development* **121**, 4275-4282.

Fortini, M. E. and Artavanis-Tsakonas, S. (1994). The Suppressor of Hairless protein participates in Notch receptor signaling. *Cell* **79**, 273-282.

Frohman, M. A., Dush, M. K. and Martin, G. R. (1988). Rapid production of full-length cDNAs from rare transcripts: amplification using a single gene-specific oligonucleotide primer. *Proc. Nat. Acad. Sci. USA* **85**, 8998-9002.

Furukawa, T., Kawaichi, M., Matsunami, N., Ryo, H., Nishida, Y. and Honjo, T. (1991). The *Drosophila* RBP-Jκ gene encodes the binding protein for the immunoglobulin Jκ recombination signal sequence. *J. Biol. Chem.* **266**, 23334-23340.

Gao, D. and Kimble, J. (1995). APX-1 can substitute for its homolog LAG-2 to direct cell interactions throughout *Caenorhabditis elegans* development. *Proc. Nat. Acad. Sci. USA* **92**, 9839-9842.

Greenwald, I. S., Sternberg, P. W. and Horvitz, H. R. (1983). The *lin-12* locus specifies cell fates in *C. elegans. Cell* **34**, 435-444.

Greenwald, I. (1985). *lin-12*, a nematode homeotic gene, is homologous to a set of mammalian proteins that includes epidermal growth factor. *Cell* **43**, 583-590.

Grossman, S. R., Johannsen, E., Tong, X., Yalamanchili, R. and Kieff, E. (1994). The Epstein-Barr virus nuclear antigen 2 transactivator is directed to response elements by the Jκ recombination signal binding protein. *Proc. Natl. Acad. Sci. USA* **91**, 7568-7572.

Heitzler, P. and Simpson, P. (1991). The choice of cell fate in the epidermis of *Drosophila*. *Cell* **64**, 1083-1092.

Henderson, S. T., Gao, D., Lambie, E. J. and Kimble, J. (1994). *lag-2* may encode a signaling ligand for the GLP-1 and LIN-12 receptors of *C. elegans*. *Development* **120**, 2913-2924.

Henkel, T., Ling, P. D., Hayward, S. D. and Peterson, M. G. (1994). Mediation of Epstein-Barr Virus EBNA2 transactivation by recombination signal-binding protein Jκ. *Science* **265**, 92-95.

Henrique, D., Adam, J., Myat, A., Chitnis, A., Lewis, J. and Ish-Horowicz, D. (1995). Expression of a *Delta* homologue in prospective neurons in the chick. *Nature* **375**, 787-790.

Hodgkin, J., Edgley, M., Riddle D. L. and Albertson, D. G. (1988). Genetic nomenclature, list of mapped genes, genetic map, physical maps. In *The Nematode Caenorhabditis elegans.* (ed. W. B. Wood). pp. 491-584. Cold Spring Harbor: Cold Spring Harbor Laboratory.

Hodgkin, J., Papp, A., Pulak, R., Ambros, V. and Anderson, P. (1989). A new kind of informational suppression in the nematode *Caenorhabditis elegans*. *Genetics* **123**, 301-313.

Horvitz, H. R., Brenner, S., Hodgkin, J. and Herman, R. K. (1979). A uniform genetic nomenclature for the nematode Caenorhabditis elegans. *Mol. Gen. Genet.* **175**, 129-33.

Hsieh, J. J.-D. and Hayward, S. D. (1995). Masking of the CBF1/RBP Jκ transcriptional repression domain by Epstein-Barr virus EBNA2. *Science* **268**, 560-563.

Hutter, H. and Schnabel, R. (1994). *glp-1* and inductions establishing embryonic axes in *C. elegans*. *Development* **120**, 2051-2064.

Jarriault, S., Brou, C., Logeat, F., Schroeter, E. H., Kopan, R. and Israël, A. (1995). Signaling downstream of activated mammalian Notch. *Nature* **377**, 355-358.

Kodoyianni, V., Maine, E. M. and Kimble, J. (1992). Molecular basis of loss-of-function mutations in the *glp-1* gene of *Caenorhabditis elegans*. *Mol. Biol. Cell* **3**, 1199-1213.

Lambie, E. J. and Kimble, J. (1991). Two homologous regulatory genes, *lin-12* and *glp-1*, have overlapping functions. *Development* **112**, 231-240.

Lecourtois, M. and Schweisguth, F. (1995). The neurogenic Suppressor of Hairless DNA-binding protein mediates the transcriptional activation of the *Enhancer of split* Complex genes triggered by Notch signaling. *Genes Dev.* **9**, 2598-2608.

Mango, S. E., Maine, E. M. and Kimble, J. (1991). Carboxy-terminal truncation activates *glp-1* protein to specify vulval fates in Caenorhabditis elegans. *Nature* **352**, 811-815.

Mango, S. E., Thorpe, C. J., Martin, P. R., Chamberlain, S. H. and Bowerman, B. (1994). Two maternal genes, *apx-1* and *pie-1*, are required to distinguish the fates of equivalent blastomeres in the early *Caenorhabditis* embryo. *Development* **120**, 2305-2315.

Matsunami, N., Hamaguchi, Y., Yamamoto, Y., Kuze, K., Kangawa, K., Matsuo, H., Kawaichi, M. and Honjo, T. (1989). A protein binding to the Jκ recombination sequence of immunoglobulin genes contains a sequence related to the integrase motif. *Nature* **342**, 934-937.

Mello, C. C., Kramer, J. M., Stinchcomb, D. and Ambros, V. (1991). Efficient gene transfer in *C. elegans*: extrachromosomal maintenance and integration of transforming sequences. *EMBO J.* **10**, 3959-3970.

Mello, C. C., Draper, B. W. and Priess, J. R. (1994). The maternal genes *apx-1* and *glp-1* and establishment of dorsal-ventral polarity in the early *C. elegans* embryo. *Cell* **77**, 95-106.

Mello, C. and Fire, A. (1995). DNA Transformation. In *Caenorhabditis elegans*; Modern Biological Analysis of an Organism. H. F. Epstein and D. C. Shakes, eds. Academic Press, pp. 452-482.

Moskowitz, I. P. G., Gendreau, S. B. and Rothman, J. H. (1994). Combinatorial specification of blastomere identity by *glp-1* dependent cellular interactions in the nematode *Caenorhabditis elegans*. *Development* **120**, 3325-3338.

Newman, A. P., White, J. G. and Sternberg, P. A. (1995). The Caenorhabditis elegans *lin-12* gene mediates induction of ventral uterine specialization by the anchor cell. *Development* **121**, 263-271.

Oka, C., Nakano, T., Wakeham, A., de la Pompa, J. L., Mori, C., Sakai, T. Okazaki, S., Kawaichi, M. Shiota, K., Mak, T. W. and Honjo T. (1995).

Disruption of the mouse RBP-Jκ gene results in early embryonic death. *Development* **121**, 3291-3301.
Otsuka, A. J., Franco, R., Yang, B., Shim, K.-H., Tang, L. Z., Zhang, Y. Y., Boontrakulpoontawee, P., Jeyaprakash, A., Hedgecock, E., Wheaton, V. I. and Sobery, A. (1995). An ankyrin-related gene (*unc-44*) is necessary for proper axonal guidance in *Caenorhabditis elegans*. *J. Cell Biol.* **129**, 1081-1092.
Priess, J. R., Schnabel, H. and Schnabel, R. (1987). The *glp-1* locus and cellular interactions in early *C. elegans* embryos. *Cell* **51**, 601-611.
Qiao, L., Lissemore, J. L., Shu, P., Smardon, A., Gelber, M. B. and Maine, E. M. (1995). Enhancers of *glp-1*, a gene required for cell-signaling in *Caenorhabditis elegans*, define a set of genes required for germline development. *Genetics* **141**, 551-569.
Sambrook, J., Fritsch, E. F. and Maniatis, T. (1989). *Molecular Cloning: A Laboratory Manual.* 2nd Edition New York: Cold Spring Harbor Laboratory Press.
Schweisguth, F. and Posakony, J. W. (1992). *Suppressor of Hairless*, the *Drosophila* homolog of the mouse recombination signal-binding protein gene, controls sensory organ cell fates. *Cell* **69**, 1199-1212.
Seydoux, G. and Greenwald, I. (1989). Cell autonomy of *lin-12* function in a cell fate decision in *C. elegans*. *Cell* **57**, 1237-1245.
Tamura, K., Taniguchi, Y., Minoguchi, S., Sakai, T., Tun, T., Furukawa, T. and Honjo, T. (1995). Physical interaction between a novel domain of the receptor Notch and the transcription factor RBP-Jκ/Su(H). *Curr. Biol.* **5**, 1416-1423.
Tax, F. E., Yeargers, J. J. and Thomas, J. H. (1994). Sequence of *C. elegans lag-2* reveals a cell-signalling domain shared with *Delta* and *Serrate* of *Drosophila*. *Nature* **368**, 150-154.
Tun, T., Hamaguchi, Y., Matsunami, N., Furukawa, T., Honjo, T. and Kawaichi, M. (1994). Recognition sequence of a highly conserved DNA binding protein RBP-Jκ. *Nucleic Acids Res.* **22**, 965-971.
Waltzer, L., Logeat, F., Brou, C., Israel, A., Sergeant, A. and Manet, E. (1994). The human Jk recombination signal sequence binding protein (RBP-Jk) targets the Epstein-Barr virus EBNA2 protein to its DNA responsive elements. *EMBO J* **13**, 5633-5638.
Wharton, K. A., Johansen, K. M., Xu, T. and Artavanis-Tsakonas, S. (1985). Nucleotide sequence from the neurogenic locus *Notch* implies a gene product which shares homology with proteins containing EFG-like repeats. *Cell* **43**, 567-581.
Wilkinson, H. A., Fitzgerald, K. and Greenwald, I. (1994). Reciprocal changes in expression of the receptor LIN-12 and its ligand LAG-2 prior to commitment in a *C. elegans* cell fate decision. *Cell* **79**, 1187-1198.
Yochem, J., Weston, K. and Greenwald, I. (1988). The *Caenorhabditis elegans lin-12* gene encodes a transmembrane protein with overall similarity to *Drosophila Notch*. *Nature* **335**, 547-550.
Yochem J. and Greenwald, I (1989). *glp-1* and *lin-12*, homologous genes implicated in distinct cell-cell interactions in *C. elegans*, encode similar transmembrane proteins. *Cell* **58**, 553-563.
Zimber-Strobl, U., Stroble, L. J., Meitinger, C., Hinrichs, R., Sakai, T., Furukawa, T., Honjo, T. and Bornkamm, G. W. (1994). Epstein-Barr virus nuclear antigen 2 exerts its transactivating function through interaction with recombination signal binding protein RBP-Jκ, the homologue of *Drosophila Suppressor of Hairless*. *EMBO J.* **13**, 4973-4982.

(Accepted 5 February 1996)

Note added in proof:

The nucleotide sequences reported in this paper have GenBank accession numbers U49794 and U49795.