

The *Caenorhabditis elegans* Genes *sqv-2* and *sqv-6*, Which Are Required for Vulval Morphogenesis, Encode Glycosaminoglycan Galactosyltransferase II and Xylosyltransferase*[§]

Received for publication, September 11, 2002,
and in revised form, February 4, 2003
Published, JBC Papers in Press, February 12, 2003,
DOI 10.1074/jbc.C200518200

Ho-Yon Hwang[‡], Sara K. Olson[§],
Jillian R. Brown[§], Jeffrey D. Esko[§],
and H. Robert Horvitz[‡]¶

From the [‡]Howard Hughes Medical Institute,
Department of Biology, Massachusetts Institute of
Technology, Cambridge, Massachusetts 02139 and the
[§]Department of Cellular and Molecular Medicine,
Glycobiology Research and Training Center, University
of California, San Diego, La Jolla, California 92093-0687

In mutants defective in any of eight *Caenorhabditis elegans sqv* (squashed vulva) genes, the vulval extracellular space fails to expand during vulval morphogenesis. Strong *sqv* mutations result in maternal-effect lethality, caused in part by the failure of the progeny of homozygous mutants to initiate cytokinesis and associated with the failure to form an extracellular space between the egg and the eggshell. Recent studies have implicated glycosaminoglycans in these processes. Here we report the cloning and characterization of *sqv-2* and *sqv-6*. *sqv-6* encodes a protein similar to human xylosyltransferases. Transfection of *sqv-6* restored xylosyltransferase activity to and rescued the glycosaminoglycan biosynthesis defect of a xylosyltransferase mutant hamster cell line. *sqv-2* encodes a protein similar to human galactosyltransferase II. A recombinant SQV-2 fusion protein had galactosyltransferase II activity with substrate specificity similar to that of human galactosyltransferase II. We conclude that *C. elegans* SQV-6 and SQV-2 likely act in concert with other SQV proteins to catalyze the stepwise formation of the proteoglycan core protein linkage tetrasaccharide GlcA β 1,3Gal β 1,3Gal β 1,4Xyl β -O-(Ser), which is common to the two ma-

jor types of glycosaminoglycans in vertebrates, chondroitin and heparan sulfate. Our results strongly support a model in which *C. elegans* vulval morphogenesis and zygotic cytokinesis depend on the expression of glycosaminoglycans.

Glycosaminoglycans (GAGs)¹ are important in animal development, and defects in GAGs are responsible for certain human disorders. For example, mutations in the *Drosophila melanogaster* genes *tout-velu* (1) and *sulfateless* (2), which encode homologs of heparan sulfate co-polymerase and heparan sulfate *N*-deacetylase/*N*-sulfotransferase, respectively, cause zygotic lethality and defects in segmentation. Mutations in the mouse *tout-velu* homolog EXT1 disrupt gastrulation and the generation of mesoderm (3), while mutations in human EXT1 and EXT2 have been associated with hereditary multiple exostoses (reviewed in Ref. 4). Mutations in the human galactosyltransferase I have been associated with a progeroid variant of the connective-tissue disorder Ehlers-Danlos syndrome (EDS) (5–7). EDS is a group of heritable disorders characterized by hyperelasticity of the skin and hypermobile joints. *Tout-velu*, EXT-1, EXT-2, and *Sulfateless* affect the biosynthesis of heparan sulfate specifically, while galactosyltransferase I deficiency affects the biosynthesis of both chondroitin and heparan sulfate.

The backbones of chondroitin and heparan sulfate consist of repeating disaccharide units: GlcA β 1,3GalNAc β 1,4 for chondroitin and GlcA β 1,4GlcNAc α 1,4 for heparan sulfate (reviewed in Ref. 8). Their polymerization occurs on a tetrasaccharide primer (GlcA β 1,3Gal β 1,3Gal β 1,4Xyl β -) that is linked to the protein core of a proteoglycan. The addition of these four sugars is catalyzed stepwise in the lumen of the Golgi apparatus and requires three nucleotide sugars, UDP-Xyl, UDP-Gal, and UDP-GlcA, and four glycosyltransferases.

Eight *sqv* (squashed vulva) genes were genetically identified in a screen for *Caenorhabditis elegans* mutants defective in vulval morphogenesis (9). All *sqv* mutants fail to form a large fluid-filled vulval extracellular space and have a reduced separation of the anterior and posterior halves of the vulva from the early to middle phases of L4 larval development. Strong mutant alleles of all eight *sqv* genes also cause maternal-effect lethality. Most progeny of mothers homozygous for a strong *sqv* mutant allele arrest at the one-cell stage (9). The nuclei of the arrested progeny divide normally, but the extrusion of the polar bodies and the initiation of cytokinesis are impaired (10). These mutant eggs fail to form the normal fluid-filled extracellular space between the membrane of the egg and the eggshell. We have postulated that the *sqv* genes control the biosynthesis of GAGs that are secreted and become hydrated to form fluid-filled extracellular spaces (10, 11).

The molecular identification of five *sqv* genes has led to a model implicating the biosynthesis of chondroitin and/or heparan sulfate in *C. elegans* development. *sqv-1*, -3, -4, -7, and -8 encode UDP-GlcA decarboxylase (10), galactosyltransferase I

* This work was supported by National Institutes of Health (NIH) Grant GM24663 (to H. R. H.) and NIH Grant GM33063 (to J. D. E.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The nucleotide sequence(s) reported in this paper has been submitted to GenBank™/EBI Data Bank with the accession number(s) AY241927 and AY241928.

[§] The on-line version of this article (available at <http://www.jbc.org>) contains a Supplementary Materials and Methods and Refs. 1–4.

¶ Investigator of the Howard Hughes Medical Institute. To whom correspondence and reprint requests should be addressed: Howard Hughes Medical Inst., Dept. of Biology, Massachusetts Inst. of Technology, Rm. 68-425, 77 Massachusetts Ave., Cambridge, MA 02139. Tel.: 617-253-4671; Fax: 617-253-8126; E-mail: horvitz@mit.edu.

¹ The abbreviations used are: GAG, glycosaminoglycan; Gal, galactose; GalNAc, *N*-acetylgalactosamine; GlcA, glucuronic acid; GlcNAc, *N*-acetylglucosamine; Xyl, xylose; EDS, Ehlers-Danlos syndrome; ORF, open reading frame; CHO, Chinese hamster ovary; UTR, untranslated region; CMV, cytomegalovirus.

(12), UDP-glucose dehydrogenase (13), UDP-GlcA/UDP-Gal/UDP-GalNAc transporter (14), and glucuronosyltransferase I (12), respectively. *sqv-3* was used to identify the human galactosyltransferase I, which has been implicated in the progeroid variant of EDS (7, 15). In this paper, we show that *sqv-6* encodes the xylosyltransferase that adds Xyl to the protein core, thus initiating GAG biosynthesis. *sqv-2* encodes a galactosyltransferase that adds the second Gal residue to the linkage tetrasaccharide.

EXPERIMENTAL PROCEDURES

C. elegans Maintenance—Strains were cultured as described (16) and were grown at 20–22 °C unless indicated otherwise.

Molecular Biology—Standard molecular biological techniques were used (17). The sequences of all PCR-amplified DNAs used for cloning were confirmed to exclude unintended mutations. Oligonucleotide sequences used for amplification or mutagenesis of DNA are shown in Supplementary Materials and Methods.

Rescue of C. elegans sqv-2 and sqv-6 Mutants—For germ line rescue, we injected cosmids carrying genomic DNA into *sqv-2*(n2821) and *sqv-6*(n2845) *unc-60*(e677)/*unc-34*(s138) animals with the dominant roller marker pRF4, as described by Mello *et al.* (18). Rol lines were established, and Rol animals and Unc-60 Rol animals were examined for rescue of the *sqv-2* and the *sqv-6* mutant phenotype, respectively. We injected *sqv-2*(n2821) hermaphrodites with plasmids containing the *sqv-2* open reading frame (ORF) under the control of the *C. elegans* heat-shock promoters (19) and pRF4 as the coinjection marker. We injected *sqv-6*(n2845)/*nT1*(n754) hermaphrodites with plasmids containing the *sqv-6* ORF under the control of the *C. elegans* heat-shock promoters (19) and pRF4. Rol lines were established, and Rol (non-Unc) animals were examined for rescue of the *sqv-2* and *sqv-6* mutant phenotype following induction of *sqv-2* and *sqv-6* expression by 30 min of heat-shock treatment at 33 °C.

SQV-2 Galactosyltransferase II Assay—A sequence encoding amino acids 25–330 of SQV-2, thus lacking the presumptive transmembrane domain at the amino terminus, was cloned into pDEST-CMV-protA. This plasmid was designed to express a secreted fusion protein containing protein A and SQV-2 amino acids 25–330. COS7 cells were transiently transfected with pDEST-CMV-protA-*sqv-2* using LipofectAMINE (Invitrogen) according to the manufacturer's instructions. After 72 h of incubation, the fusion protein was recovered from the cell culture supernatant by affinity chromatography using IgG-agarose (20). Galactosyltransferase II activity was assayed as described by Bai *et al.* (21).

Rescue of the Xylosyltransferase Defect in Chinese Hamster Ovary (CHO) pgsA-745 Cells by sqv-6—The xylosyltransferase-deficient CHO pgsA-745 cells (22) were transfected with *sqv-6* ORF, which was cloned into pcDNA3.1. Stable transfectants were selected with 400 µg/ml geneticin (Invitrogen). Several drug-resistant colonies were isolated and screened by flow cytometry for *sqv-6* expression based on binding of biotinylated FGF-2 as described (23). Incorporation of ³⁵SO₄ into GAG chains of wild-type CHO or pgsA-745 cells with or without *sqv-6* was assayed essentially as described by Bame and Esko (24), labeling cells overnight at 30 °C with 50 µCi/ml ³⁵SO₄ (PerkinElmer Life Sciences).

SQV-6 Xylosyltransferase Assay—Cell extracts of wild-type CHO, pgsA-745, and *sqv-6* or empty vector stable transfectants of pgsA-745 were prepared as described previously (22). Xylosyltransferase activity was assayed essentially as described (22) by incubating 25 µg crude cell extract with 50 µg of soluble silk acceptor and 6 × 10⁵ cpm UDP-[1-³H]xylose (PerkinElmer Life Sciences, 8.9 Ci/mmol) at 26 °C for 5 h. Product formation was dependent on the addition of silk. The concentration of substrate was saturating.

RESULTS AND DISCUSSION

Molecular Identification of sqv-2—*sqv-2* was previously mapped to the left of *lin-31* on LGII (25). We further mapped *sqv-2* to an interval between *sup-9* and *lin-31* (see Supplementary Materials and Methods). We assayed 27 cosmids in this interval for the ability to rescue the *sqv-2* mutant phenotype, but none showed rescuing activity (Fig. 1A).

We examined the DNA sequence corresponding to the gaps between the cosmids in this interval and found a predicted gene, Y110A2AL.14, that is weakly similar to galactosyltransferases. Because all previously cloned *sqv* genes are implicated

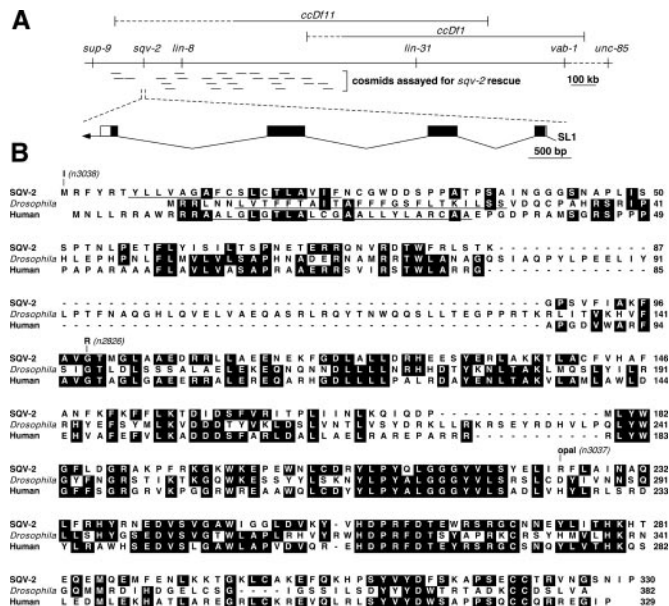


FIG. 1. SQV-2 is similar to galactosyltransferase II. A, genetic and physical maps showing *sqv-2*. The dashed horizontal lines depicting *ccDf1* and *ccDf1* indicate the possible extents of the left end points of these deletions (C. Spike and R. Herman (University of Minnesota), personal communication; E. Davison and H. R. Horvitz, unpublished observations). Short solid lines represent cosmid clones that were assayed in germ line transformation experiments. Below is the structure of the *sqv-2* gene as deduced from genomic and cDNA sequences. Solid boxes indicate exons, and open boxes indicate untranslated sequences. The trans-spliced leader SL1 is indicated, and the arrow indicates the poly(A) tail. B, alignment of SQV-2, the *Drosophila* homolog, and human galactosyltransferase II. Identities between at least two proteins are shaded in black. The predicted transmembrane domains are underlined. The three *sqv-2* mutant alleles are indicated. The numbers on the right indicate amino acid positions.

TABLE I

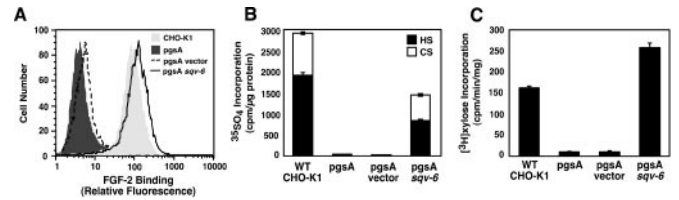
The SQV-2 fusion protein has acceptor specificity consistent with its being galactosyltransferase II

Galactosyltransferase activity was assayed *in vitro* using UDP-[³H]galactose together with various acceptor substrates. No substrate controls ranged from 339 to 357 cpm. The range for all substrates shown as "0" activity was 137–619 cpm. The range for the substrate Galβ1,4Xylβ-O-Bn was 141,000–142,000 cpm. Bn, benzyl; NM, naphthalenemethanol; C₁₀, O-decenyl (CH₂)₈CH=CH₂

Acceptor substrates	Enzyme activity
	pmol/h/ml medium
Monosaccharides (5 mM)	
Xylβ-O-Bn	0
Xylβ-O-naphthol	0
Galβ-O-NM	0
GalNAcα-O-Bn	0
GlcNAcβ-O-NM	0
Disaccharides (5 mM)	
Galβ1,4Xylβ-O-Bn	2660
Galβ1,3GalNAcα-O-NM	1
Galβ1,3Galβ-O-NM	3
Galβ1,4GlcNAcβ-O-NM	0
Galβ1,3GlcNAcβ-O-NM	0
GlcNAcβ1,3Galβ-O-NM	6
Manα1,6Manα-O-C ₁₀	0

in the biosynthesis of chondroitin and/or heparan sulfate, we suspected that *sqv-2* also encodes a protein involved in GAG biosynthesis. Specifically, it seemed plausible that Y110A2AL.14 encodes the galactosyltransferase II involved in the formation of the protein core linkage tetrasaccharide and that had not been identified molecularly in any organism at the time.

We identified three molecular lesions corresponding to three of the four identified alleles of *sqv-2* in the ORF of



The diagram illustrates the biosynthesis of SQV (Sulfated Quinovic Acid) and its incorporation into proteoglycans. The process begins in the cytosol with the conversion of UDP to UDP-SQV-4. SQV-4 is then transported into the Golgi apparatus, where it is converted to UDP-SQV-7. SQV-7 is further processed through a series of steps involving SQV-1, SQV-2, SQV-3, SQV-6, and SQV-8, each releasing a UDP derivative. The final product, SQV-8, is then incorporated into the core structure of chondroitin and heparan sulfate proteoglycans. The diagram shows the repeating disaccharide units of these proteoglycans, with SQV-8 being a key component. A legend identifies the symbols used for the various sugars: glucose (white triangle), glucuronic acid (black triangle), galactose (black circle), GalNAc (white square), GlcNAc (black square), and xylose (white triangle).

Legend:

- ▲ glucose
- ◆ glucuronic acid
- galactose
- GalNAc
- GlcNAc
- ▼ xylose

Proteoglycan Structures:

- chondroitin proteoglycans:** A repeating unit consisting of a disaccharide (glucose and glucuronic acid) linked to a core structure (GlcNAc and galactose) via a sulfate group (SQV-8).
- heparan sulfate proteoglycans:** A repeating unit consisting of a disaccharide (glucose and glucuronic acid) linked to a core structure (GlcNAc and galactose) via a sulfate group (SQV-8).

Molecular Identification of *sqv-6*—*sqv-6* was previously mapped to the left of the polymorphism *stP3* on LGV (9). We

We determined the sequences of two cDNA clones, yk94e4 and yk292g2 (see Supplementary Materials and Methods), that correspond to Y110A2AL.14. The yk292g2 clone contains 990 bases of ORF, 17 bases of 5'-untranslated region (UTR), and 121 bases of 3'-UTR. The 5' end contains three bases that correspond to the sequence of 5' SL1 *trans*-spliced leader, which is found at the 5' end of many *C. elegans* transcripts (26). The 3' end contains a poly(A) sequence. The longest ORF in

further mapped *sqv-6* to the left of the cosmid W07B8 and within about 0.2 map units of *unc-34* (see Supplementary Materials and Methods). We assayed 11 cosmids to the right of *unc-34* for the ability to rescue the *sqv-6* mutant phenotype, but none showed rescuing activity (Fig. 2A).

We examined the DNA sequences in the gaps in the cosmid coverage near the cosmid W07B8 and *unc-34* and found a gene, Y50D4C.d, that is similar to two human xylosyltransferases (27). Using DNA from the only allele of *sqv-6*, *n2845*, we identified in the ORF of Y50D4C.d an amber nonsense mutation causing the elimination of the last 42 amino acids of the predicted protein product (Fig. 2B).

We determined the sequence of PCR-amplified cDNA and 5'-rapid amplification of cDNA ends products corresponding to Y50D4C.d (see Supplementary Materials and Methods). We found that this cDNA contains a 5' SL1 *trans*-spliced leader, 23 bases of 5'-UTR, and 2418 bases of ORF, including two additional 5' exons not in Y50D4C.d. The longest ORF in this cDNA including the additional exons is predicted to encode a protein of 806 amino acids. Expression of this ORF under the control of the *C. elegans* heat-shock promoters (19) prior to the start of vulval morphogenesis rescued the *sqv-6* vulval morphogenesis defect in all animals ($n = 13$) and the maternal-effect lethality of the progeny of *sqv-6* homozygotes generated by $+/sqv-6$ heterozygous parents for three of 13 *sqv-6* homozygotes studied.

sqv-6 Encodes a Protein Similar to Xylosyltransferases—Of the 806 amino acids of the SQV-6 protein, 182 (23%) and 193 (24%) are identical to human xylosyltransferases I and II, respectively (Fig. 2B). Both the predicted SQV-6 protein and the human xylosyltransferase II contain a putative transmembrane domain near the amino terminus and are likely to be type II transmembrane proteins. Neither the start codon nor a presumptive transmembrane domain has been defined for human xylosyltransferase I (27).

sqv-6 Can Correct a Xylosyltransferase Defect in CHO Cells—We tested the ability of *sqv-6* to act as a xylosyltransferase by testing its ability to complement GAG-deficient CHO mutant cells lacking this enzymatic activity (22). Mutant pgsA-745 cells were transiently transfected with a plasmid containing *sqv-6* under the control of a cytomegalovirus (CMV) promoter. These cells showed partial rescue of the defect, as assayed by the restored abilities to incorporate $^{35}\text{SO}_4$ into GAGs (16–27% that of the wild type) and to bind biotinylated FGF-2, which binds cell surface heparan sulfate as assayed by flow cytometry (data not shown). From these transiently transfected cells, we obtained a clonal cell line stably expressing *sqv-6*. This cell line showed full restoration of FGF-2 binding to heparan sulfate on the cell surface (Fig. 3A). Stable expression of *sqv-6* in pgsA-745 cells enhanced the incorporation of $^{35}\text{SO}_4$ into GAGs to ~50% of wild-type levels, compared with 1% for the untreated mutant or mutant transfected with empty vector (Fig. 3B). The $^{35}\text{SO}_4$ incorporation into GAGs was similar in wild-type CHO cells and pgsA-745 cells transfected with *sqv-6*: 30–40% was released by treatment with chondroitinase ABC and 55–65% by a heparin lyase mixture in both cells, indicating that the composition of chondroitin and heparan sulfate was comparable. Expression of *sqv-6* also resulted in restoration of xylosyltransferase activity, as measured by the transfer of xylose from UDP-xylose to a soluble silk acceptor (22),

whereas pgsA-745 cells transfected with empty vector had virtually no activity (Fig. 3C).

The sqv-2 and sqv-6 Genes Act in the C. elegans Chondroitin and Heparan Sulfate Biosynthesis Pathway—Our findings indicate that *sqv-2* and *sqv-6* encode galactosyltransferase II and xylosyltransferase, respectively. With the previously identified *sqv-3* galactosyltransferase I and *sqv-8* glucuronosyltransferase I, all four *C. elegans* genes responsible for the biosynthesis of the proteoglycan core protein linkage tetrasaccharide of chondroitin and heparan sulfate have now been defined (Fig. 4). Three previously identified genes, *sqv-4* UDP-glucose dehydrogenase, *sqv-1* UDP-GlcA decarboxylase and *sqv-7* UDP-GlcA/UDP-Gal/UDP-GalNAc transporter, act in earlier steps of GAG biosynthesis. All *sqv* genes identified to date affect the biosynthesis of both chondroitin and heparan sulfate. Based upon these observations, we conclude that in *C. elegans* early embryonic cytokinesis and epithelial invagination during vulval development depend on the expression of GAGs.

Acknowledgments—We thank Beth Castor for help with DNA sequence determination, Ewa Davison and Ignacio Perez de la Cruz for communicating the locations of the *lin-8* and *sup-9* loci, respectively, and Mark Alkema and Melissa Harrison for helpful suggestions concerning this manuscript.

REFERENCES

- Bellaiche, Y., The, I., and Perrimon, N. (1998) *Nature* **394**, 85–88
- Lin, X., and Perrimon, N. (1999) *Nature* **400**, 281–284
- Lin, X., Wei, G., Shi, Z., Dryer, L., Esko, J. D., Wells, D. E., and Matzuk, M. M. (2000) *Dev. Biol.* **224**, 299–311
- Zak, B. M., Crawford, B. E., and Esko, J. D. (2002) *Biochim. Biophys. Acta* **1573**, 346–355
- Quentin, E., Gladen, A., Roden, L., and Kresse, H. (1990) *Proc. Natl. Acad. Sci. U. S. A.* **87**, 1342–1346
- Almeida, R., Levery, S. B., Mandel, U., Kresse, H., Schwientek, T., Bennett, E. P., and Clausen, H. (1999) *J. Biol. Chem.* **274**, 26165–26171
- Okajima, T., Fukumoto, S., Furukawa, K., and Urano, T. (1999) *J. Biol. Chem.* **274**, 28841–28844
- Esko, J. D., and Selleck, S. B. (2002) *Annu. Rev. Biochem.* **71**, 435–471
- Herman, T., Hartwig, E., and Horvitz, H. R. (1999) *Proc. Natl. Acad. Sci. U. S. A.* **96**, 968–973
- Hwang, H.-Y., and Horvitz, H. R. (2002) *Proc. Natl. Acad. Sci. U. S. A.* **99**, 14218–14223
- Herman, T., and Horvitz, H. R. (1999) *Proc. Natl. Acad. Sci. U. S. A.* **96**, 974–979
- Bulik, D. A., Wei, G., Toyoda, H., Kinoshita-Toyoda, A., Waldrup, W. R., Esko, J. D., Robbins, P. W., and Selleck, S. B. (2000) *Proc. Natl. Acad. Sci. U. S. A.* **97**, 10838–10843
- Hwang, H.-Y., and Horvitz, H. R. (2002) *Proc. Natl. Acad. Sci. U. S. A.* **99**, 14224–14229
- Berninsone, P., Hwang, H. Y., Zemtseva, I., Horvitz, H. R., and Hirschberg, C. B. (2001) *Proc. Natl. Acad. Sci. U. S. A.* **98**, 3738–3743
- Okajima, T., Yoshida, K., Kondo, T., and Furukawa, K. (1999) *J. Biol. Chem.* **274**, 22915–22918
- Brenner, S. (1974) *Genetics* **77**, 71–94
- Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY
- Mello, C. C., Kramer, J. M., Stinchcomb, D., and Ambros, V. (1991) *EMBO J.* **10**, 3959–3970
- Stringham, E. G., Dixon, D. K., Jones, D., and Candido, E. P. (1992) *Mol. Biol. Cell* **3**, 221–233
- Wei, Z., Swiedler, S. J., Ishihara, M., Orellana, A., and Hirschberg, C. B. (1993) *Proc. Natl. Acad. Sci. U. S. A.* **90**, 3885–3888
- Bai, X., Zhou, D., Brown, J. R., Crawford, B. E., Hennen, T., and Esko, J. D. (2001) *J. Biol. Chem.* **276**, 48189–48195
- Esko, J. D., Stewart, T. E., and Taylor, W. H. (1985) *Proc. Natl. Acad. Sci. U. S. A.* **82**, 3197–3201
- Bai, X., Wei, G., Sinha, A., and Esko, J. D. (1999) *J. Biol. Chem.* **274**, 13017–13024
- Bame, K. J., and Esko, J. D. (1989) *J. Biol. Chem.* **264**, 8059–8065
- Herman, T., and Horvitz, H. R. (1997) *Cold Spring Harbor Symp. Quant. Biol.* **62**, 353–359
- Krause, M., and Hirsh, D. (1987) *Cell* **49**, 753–761
- Gottling, C., Kuhn, J., Zahn, R., Brinkmann, T., and Kleesiek, K. (2000) *J. Mol. Biol.* **304**, 517–528