THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. 280, NO. 46, pp. 38787–38794, November 18, 2005

Distinct Regulatory Elements Mediate Similar Expression Patterns in the Excretory Cell of *Caenorhabditis elegans*[*] [S]

Received for publication, May 25, 2005, and in revised form, August 23, 2005 Published, JBC Papers in Press, September 13, 2005, DOI 10.1074/jbc.M505701200

Zhongying Zhao[‡§1], Li Fang[§], Nansheng Chen[¶], Robert C. Johnsen[§], Lincoln Stein[¶], and David L. Baillie[§]

From the [‡]Department of Genome Science, University of Washington, Seattle, Washington 98195-7730, the [§]Department of Molecular Biology and Biochemistry, Simon Fraser University, Burnaby, British Columbia V5A 1S6, Canada, and the [¶]Cold Spring Harbor Laboratory, Cold Spring Harbor, New York 11724

Identification of *cis*-regulatory elements and their binding proteins constitutes an important part of understanding gene function and regulation. It is well accepted that co-expressed genes tend to share transcriptional elements. However, recent findings indicate that co-expression data show poor correlation with co-regulation data even in unicellular yeast. This motivates us to experimentally explore whether it is possible that co-expressed genes are subject to differential regulatory control using the excretory cell of *Caenorhabditis elegans* as an example. Excretory cell is a functional equivalent of human kidney. Transcriptional regulation of gene expression in the cell is largely unknown. We isolated a 10-bp excretory cell-specific *cis*-element, Ex-1, from a *pgp-12* promoter. The significance of the element has been demonstrated by its capacity of converting an intestine-specific promoter into an excretory cell-specific one. We also isolated a cDNA encoding an Ex-1 binding transcription factor, DCP-66, using a yeast one-hybrid screen. Role of the factor in regulation of *pgp-12* expression has been demonstrated both *in vitro* and *in vivo*. Search for occurrence of Ex-1 reveals that only a small portion of excretory cell-specific promoters contain Ex-1. Two other distinct *cis*-elements isolated from two different promoters can also dictate the excretory cell-specific expression but are independent of regulation by DCP-66. The results indicate that distinct regulatory elements are able to mediate the similar expression patterns.

Given the increasing number of genome sequences, a major challenge is to understand the non-coding sequences of the genomes. The temporal and spatial expression pattern of genes is controlled by a variety of short *cis*-acting DNA elements (*cis*-elements) that act as binding sites for transcription factors. These *cis*-elements have frequently been identified within the non-coding sequences of different genomes (1–4). Therefore, identification of *cis*-regulatory elements and their binding proteins constitute an important part of deciphering the role of non-coding sequences. Investigation of the regulation of tissue-specific expression began over a dozen years ago in *Caenorhabditis elegans* (5), and has been greatly facilitated by the use of reporter genes such as green fluorescent protein (GFP)[2] (6), but progress is limited because of the complexity and labor intensity of the analysis. In *C. elegans*, *cis*- and *trans*-regulation of tissue-specific gene expression is known only for a few tissues such as intestine (7), excretory duct cell (8), another cell of excretory system along with the excretory cell, and muscle (1). Many of such investigations usually have been done by searching for over-represented sequences within a subset of promoters of co-expressed genes. However, recent investigations in yeast demonstrated poor correlation between co-expression and co-regulation (9, 10), suggesting a substantial number of co-expressed genes might not be co-regulated. To test this hypothesis, we use *C. elegans* excretory cell as an example.

Excretory cell, also called the excretory canal cell, is the largest mononucleate tubular cell in *C. elegans*, making it an excellent model for investigating the morphogenesis of biological tube. It functions in part as a kidney, excreting saline fluid via the duct and pore to maintain the salt balance of the animal and probably to remove metabolites (11, 12). Its unique morphology makes it straightforward and unambiguous to identify the cell *in vivo*. Mutation of several genes has previously been reported to affect morphogenesis of the excretory cell. For example, disruption of the POU homeobox gene, *ceh-6*, leads to inappropriate canal structures in *C. elegans* (13). Loss of function for *unc-53* blocks the progression of canal development (14). Mutations in a mucin encoding gene, *let-653*, result in an extremely large canal lumen (15). Several other mutants have also been isolated, which show enlarged canal lumens termed cysts (12). However, regulatory mechanisms that control tissue-specific expression in the excretory cell have not been reported. Here, we have chosen the *C. elegans* excretory cell as a model to address two questions: what is the regulatory mechanism, *i.e. cis*-element and its binding protein, which controls the tissue-specific expressions in the cell; and whether the excretory cell-specific genes share the same regulatory mechanisms.

EXPERIMENTAL PROCEDURES

Strains—All strains were maintained and cultured using standard techniques. The following strains were used: N2 (Bristol); KR3532, *dpy-5*(*e907*); VC26 *pgp-12*(*gk19*) X; VC924 *dcp-66*(*gk370*) I/hT2[*bli-4*(*e937*)*let-?*(*q782*)qIs48] (I;III); BC06405, *dpy-5*(*e907*), sEX970[*dpy-5*(+)+rCes-*pgp-12*-260-GFP + pCes361]; BC06392, *dpy-5*(*e907*), sEX958[*dpy-5*(+) + *rCes-pgp-12*::DsRed (a red fluorescence protein derived from coral) + pCeh361]; BC10210, *dpy-5*(*e907*), sIs10089[*dpy-5*(+) + rCes-*pgp-12*-GFP+pCes361]; BC06288, *dpy-5*(*e907*), sEX906[*dpy-5*(+) + rCes-Cb-*pgp-12*-GFP+pCes361]; BC06293, *dpy-5*(*e907*), sEX911[*dpy-5*(+) + rCes-*pgp-12*-228-GFP+pCes361]; BC6418, *dpy-5*(*e907*), sEX982[*dpy-5*(+) + rCes-*pgp-4-DsRed*+pCes361]; BC10036, *dpy-5*(*e907*), sEX870[*dpy-5*(+) + rCes-*haf-4*-GFP+pCes361]; BC12036, *dpy-5*(*e907*), sEX12036[*dpy-5*(+) + rCes-*pgp-14*-DsRed+pCes361]; BC10364, *dpy-5*(*e907*), sEX10364[*dpy-5*(+) + rCes-*pmp-1*-DsRed+pCes361]; BC06295, *dpy-5*(*e907*), sEX913[*dpy-5*(+) + rCesZK470.5-GFP+pCes361]; BC06417, *dpy-5*(*e907*), sEX981[*dpy-5*(+) + rCesC18C4.2-DsRed+pCes361]; BC12778, *dpy-5*(*e907*), sEX12778[*d-*

* This work was supported by NSERC Canada, Genome British Columbia, and Genome Canada. 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.

[S] The on-line version of this article (available at http://www.jbc.org) contains supplemental Figs. S1–S5 and Tables 1–3.

[1] To whom correspondence should be addressed. Tel.: 206-221-4648; Fax: 206-685-7301; E-mail: zhao@gs.washington.edu.

[2] The abbreviations used are: GFP, green fluorescence protein; EMSA, electrophoresis mobility shift assay.

py-5(+) + rCes-dcp-66-GFP+pCes361]; BC12781, *dpy-5*(*e907*), sEX-12781[*dpy-5*(+) + rCesR05D11.3-GFP+pCes361]; BC10257, *dpy-5*(*e907*), sEX10257[*dpy-5*(+) + rCes-pgp-3-GFP+pCes361]; BC13573, *dpy-5*(*e907*), sEX13573[*dpy-5*(+) + rCes-nas-31-GFP+pCes361].

Construction of Reporter Genes and Site-directed Mutagenesis—Promoter::reporter fusion constructs were built by the fusion PCR technique as previously described (16). Site-directed deletions were performed by similar fusion techniques. The two primer sets were designed to amplify fragments upstream and downstream of the sequence to be deleted. The reverse primer used to amplify the upstream region contains an overhang that is complementary to the forward primer used to amplify the downstream region. The two fragments were fused together by PCR using the forward primer to amplify the upstream fragment and reverse primers to amplify the downstream fragment. The resulting PCR products were sequenced to confirm the correctness of the deletions.

DNA Transformation and Microscopy—DNA transformation and microscopy were performed as described (17). The DNA concentrations for the GFP construct and *dpy-5* rescuing plasmid in the injection mixture were 10 and 100 ng/μl. Pictures were taken for 20 worms each for at least 3 independent lines using the same exposure time (1.05 s). GFP intensity was measured for at least three independent strains using public domain NIH Image program version 1.60.[3] Only cell bodies were measured. Chromosomal integration of the *pgp-12* transgenic strain was performed as described (19). The transgenic worms for *pgp-12*::*GFP* were exposed to low dose x-ray irradiation (1500 R) to induce double-stranded breaks in chromosomes. The resulting F2 worms were screened for integrants.

Yeast One-hybrid Screening—Yeast one-hybrid screen was performed according to the manufacturer's instruction (MATCHMAKER One-Hybrid System; Clontech). Briefly, three tandem copies of the Ex-1 element were synthesized as two reverse complementary strands with EcoRI and XbaI overhangs at the ends. Annealing of the two reverse complementary single-stranded DNAs yielded an EcoRI and XbaI adaptor. The resulting adaptor was cloned upstream of the HIS3 reporter plasmid pHISi digested with EcoRI and XbaI. The reporter plasmid was then integrated into yeast strain YM4271 to generate the reporter strain. Background growth of this reporter strain was ablated by titrating with 30 mM 3-aminotrizole (Sigma). Screening was carried out by transforming worm cDNA library RB2 (A gift from Dr. Robert Barstead) into the Ex-1-integrated pHISi-1 reporter strain using the LiOAc strategy. Around 2 million clones were screened. Plasmids were retrieved from the yeast colonies growing in His-Leu-agar plates with 30 mM 3-aminotrizole. The cDNA inserts were sequenced and then compared against Wormbase records. To confirm yeast one-hybrid screening results, the resulting positive clones were transformed back into the Ex-1-integrated YM4721 strain. The yeast colonies from each transformation were then streaked onto His-Leu-agar plates with 30 mM 3-aminotrizole. The growth of the streaked yeast colonies was compared with control (YM7271 with Ex-1 reporter integrated within) after 3 days incubation at 30 °C.

Electrophoretic Mobility Shift Assays (EMSA)—Primers were designed so that the *dcp-66* cDNA sequence isolated from the one-hybrid screen was amplified as template for protein production. *dcp-66* cDNA encoding 122 amino acids that contains the predicted zinc finger domain was used as template to produce protein using the TnT QuikChange Transcription/Translation System from Promega. The resulting protein was quantified using the Pierce Micro BCA™ Protein Assay Kit (catalog number 23227) with Test Tube Procedure and SDS-polyacrylamide gel analysis. The three tandem Ex-1 or a single copy of Ex-2 or Ex-3 were labeled by the Pierce Biotin 3′ End DNA Labeling Kit (catalog number 89818). The labeling efficiency was determined as described by the manufacturer. The labeled Ex-1, Ex-2, or Ex-3 were annealed with its reverse complementary strand to form a double-stranded oligo. EMSA was performed using the Pierce LightShift™ Chemiluminescent EMSA Kit (catalog number 20148X) following manufacturer's specifications.

RNA Interference—RNA interference by feeding was done using feeding library provided by Ahringer laboratory (20) as described. Five GFP expressing L4 worms were transferred onto plates seeded with RNA interference bacteria or control bacteria (containing empty RNA interference vector) at room temperature. After 2 days, 30 GFP expressing L2 or L3 worms were transferred onto three replica plates seeded with the same bacteria as mentioned above and checked for GFP expression after incubation for another day. Loss of GFP expression was calculated as average percentage from the three plates.

Genetic Analysis of the dcp-66 Deletion Mutant—The deletion allele *dcp-66*(*gk370*) was isolated by the *C. elegans* Reverse Genetics Core Facility, which is part of the International *C. elegans* Gene Knock-out Consortium, using a PCR screen of a mutant collection of TMP/UV-mutagenized animals. Homozygous *dcp*-66(*gk370*) worm are inviable and most of them die as small, disintegrated adults. Allele *gk370* was balanced by pharyngeal GFP marked translocation hT2 in the strain VC924.

Outcrossing was performed as follows: Four VC924 worms were crossed with four N2 male at 20 °C. Four F1 self-progeny without GFP (gk370 heterozygote) were mated with six F1 male progeny with GFP. GFP marked F2 crossing progeny were transferred into individual plates. The presence of *gk370* in the F2 crossing progeny was followed by segregation of inviable small and disintegrated adults. The process was repeated another three times and the resulting outcrossed strains were used for subsequent analysis.

The outcrossed VC924 worms were mated with N2 male. 40 cross progeny without GFP were transferred onto one fresh plate and allowed to lay eggs for 3 h at 20 °C and then picked off the plate. The laid eggs were scored and phenotypes were examined every 12 h.

BC10210 worms carrying the integrated transgene *pgp-12*::*GFP* were crossed to N2 male and the resulting F1 GFP male progeny were crossed into the outcrossed VC924 hermaphrodites. The progeny with GFP in the excretory cell but not in the pharynx were picked up for phenotypic analysis. Similar crossings were done between the outcrossed VC924 and BC6148, BC10257, or BC13573, which carries transgenes *pgp-4*::*DsRed*, *pgp-3*::*GFP*, or *nas-31*::*GFP*, respectively.

Rescue analysis to verify the genetic deletion was performed by the generation of transgenic strains carrying both the 9307-bp genomic regions of the *dcp-66* and *pgp-12*::*GFP* construct. The region was amplified from N2 genomic DNA containing 3188 bp upstream of the translation start and 1882 bp downstream of the translation stop codon using forward primer, gaatggttttcaagagaaaggcta, and reverse primer, tggagagagtatgttcactcaattaca. The resulting transgenic worms were crossed to N2 male, and F1 GFP male progeny were crossed to the outcrossed VC924 worms. The progeny with GFP in the excretory cell but not pharynx were plated and the presence of the *gk370* allele was followed as mentioned above.

Bioinformatic Analysis—Promoters of eight genes that are exclusively expressed in the excretory cell were scanned for 10-mers that show over 80% identity to the Ex-1 sequence. The resulting hits were used to formulate the Ex-1 profile and generated a sequence logo by

[3] Scion (rsb.info.nih.gov/nih-image).

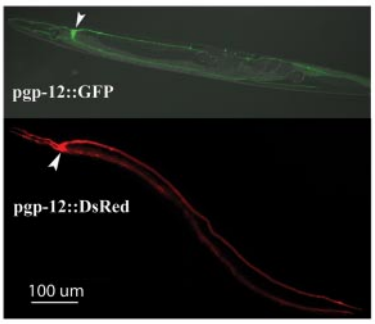

FIGURE 1. ***pgp-12* promoter drove GFP or DsRed expressions exclusively in the excretory cell.** Photographs were taken on transgenic animals containing *pgp-12::GFP* (*top panel*) or *pgp-12::DsRed* (*bottom panel*). Excretory cell bodies are indicated with *arrows*. *Scale bar* for 100 μm is shown.

WebLogo (21). The resulting profile was used to scan 82 excretory non-exclusive and 223 non-excretory cell gene promoters (19) for hits with over 80% identity to Ex-1 sequence.

RESULTS

Promoters of pgp-12 and its Caenorhabditis briggsae Ortholog Drive GFP/DsRed Expression in the Excretory Cell—To identify *cis*-regulatory elements that direct specific expressions in the excretory cell, we chose Ce-pgp-12, a 3410-bp excretory cell exclusive promoter from *pgp*-12 (17) and fused it to GFP and DsRed to generate transgenic animals. As expected, the promoter drove both GFP and DsRed expression exclusively in the excretory cell at all developmental stages (Fig. 1). The dual reporter genes were used to exclude the possibility that the resulting expression was the artifact of reporter rather than the result of the promoter used. To determine whether the observed pattern is conserved between *C. elegans* and its sister species, *C. briggsae*, which has been diverged from the former around 100 million years (22), we generated the fusion construct between GFP and a 1160-bp *C. briggsae* orthologous region of Ce-pgp-12, Cb-pgp-12. The resulting construct was introduced into *C. briggsae* AF16 strain. The same expression pattern was observed as that of Ce-pgp-12 (data not shown). To further confirm the conserved regulatory mechanisms between the two species, we reciprocally introduced fusion constructs from one species into the other, and observed the indistinguishable expression patterns in either species, indicating the conserved regulatory mechanisms for the expression of the two orthologous genes.

A 10-bp Element, Ex-1, Is Critical for Driving Excretory Cell-specific Expression—To locate the *cis*-element that drives the excretory cell-specific expression, we performed a series of truncations on the 3410-bp Ce-pgp-12 to determine definitive regions responsible for excretory cell-specific expression. As a result, we were able to identify a 10-bp element, termed hereafter as Ex-1 with the sequence ccatacatta, deletion of which caused complete loss of the expression driven by the promoter (Fig. 2*A*). A search of the TransFac data base (23) indicates that Ex-1 is a novel *cis*-element. Alignment of Ce-pgp-12 and its *C. briggsae* orthologous region, Cb-pgp-12, showed perfect conservation of Ex-1. The flanking sequences, however, are not well conserved (Fig. 2*B*). Site-directed removal of Ex-1 but not Ex-L or Ex-R could completely abolish GFP expression. Therefore, Ex-1 is the most critical element for the observable expressions. Ex-L and Ex-R also contribute to the expression because removal of them caused decreased expression (Fig. 2*A*).

For further confirmation that Ex-1 drives tissue-specific expression in the excretory cell, we examined whether Ex-1 could confer excretory cell expression on a promoter not usually expressed in this cell, *i.e.* ectopic expression. A 286-bp intestine-specific promoter, Ce-haf-4, drives moderate GFP expression in the intestine (TABLE ONE). A single copy of Ex-1 fused to the 5′ end of the promoter does not cause any changes in its expression pattern. However, Ex-L-Ex-1-Ex-R joined at the 5′ end of 286-bp Ce-haf-4 successfully change expression patterns of the promoter, *i.e.* from intestine-specific expression to expression in both intestine and excretory cell (TABLE ONE and supplemental Fig. S1). Interestingly, the reversed complementary sequence for Ex-L-Ex-1-Ex-R joined on the 5′ end of Ce-haf-4 did not cause any changes in the expression patterns, indicating that Ex-1 and its flanking sequences work in an orientation dependent manner. In addition, the Ex-L-Ex-1-Ex-R joined at the 5′ end of the 498-bp Ce-haf-4 cannot change expression patterns of the promoter, suggesting the three elements work in a distance-dependent manner (TABLE ONE). Ex-L-Ex-1-Ex-R are the minimum sequences that can cause ectopic expressions in the excretory cell. Strikingly, the three elements together with its upstream 48-bp sequence from Ce-pgp-12 were able to make the 286-bp Ce-haf-4 drive GFP expression only in the excretory cell but not in the intestine when they were joined at the 5′ end the promoter. The Ex-L-Ex-1-Ex-R sequence was also tested on the 5′ end of two other promoters: a 1724-bp pharynx-specific promoter Ce-pgp-14 and a 1193-bp intestine-specific promoter Ce-pmp-1. No changes were observed in expression patterns of the two promoters (TABLE ONE), again implying that there is a requirement for distance to the start codon for Ex-1 and its flanking sequences to function properly.

Given the significance of the 48-bp sequence in directing excretory cell-specific expression, we asked whether it can cause ectopic expression in the excretory cell by itself. Interestingly, the 286-bp Ce-haf-4 promoter did not give observable expression in both the excretory cell and the intestine when the 48-bp sequence was placed at its 5′ end (TABLE ONE), indicating that the 48-bp sequence function as a suppressor for the Ce-haf-4-mediated intestine expression. However, inclusion of the 48-bp sequence at the 5′ end of the 498-bp Ce-haf-4 did not affect the expression. To map the essential part of the 48-bp sequence responsible for the observed inhibition, we split the sequence into two 24-bp fragments and fused each of them at the 5′ end of the Ce-haf-4, neither of them can significantly inhibit Ce-haf-4-mediated intestine expression. The central 24 bp of the 48-bp sequence fused at the 5′ end of Ce-haf-4 only caused modest suppression, indicating that most parts of the 48-bp sequence are involved in such inhibition. Scanning the 48-bp sequence against TransFac (23) did not give any hits. Alignment of *C. elegans* and *C. briggsae pgp-12* promoters showed little conservation within the region. Joining the 48-bp *C. briggsae* sequence upstream of Ex-L within Cb-pgp-12 at the 5′ end of Ce-haf-4 did not alter Ce-haf-4-mediated expression (TABLE ONE). We fused the 48-bp sequence along with Ex-L-Ex-1-Ex-R to the 5′ end of two other promoters, *i.e.* pharynx-specific Ce-pgp-14 and intestine-specific promoter Ce-pmp-1. No obvious expression changes were observed as opposed to expression without the grafted sequence. The results show that the 48-bp sequence work as a suppressor for Ce-haf-4 at a certain distance to the start codon. It is unknown whether the 48-bp sequence serves as a binding site for suppressors or interferes with the binding of intestine-specific transcription factors to the Ce-haf-4.

Identification of Transcription Factor That Binds Ex-1 Using One-hybrid Screening—To identify the Ex-1 binding transcription factors, we performed a yeast one-hybrid screen (Clontech) using three tandem Ex-1s as a bait. Two positive clones were isolated from the screen. One of the inserts contained an in-frame cDNA sequence for R05D11.3. However, no substantial DNA binding domain was found within the

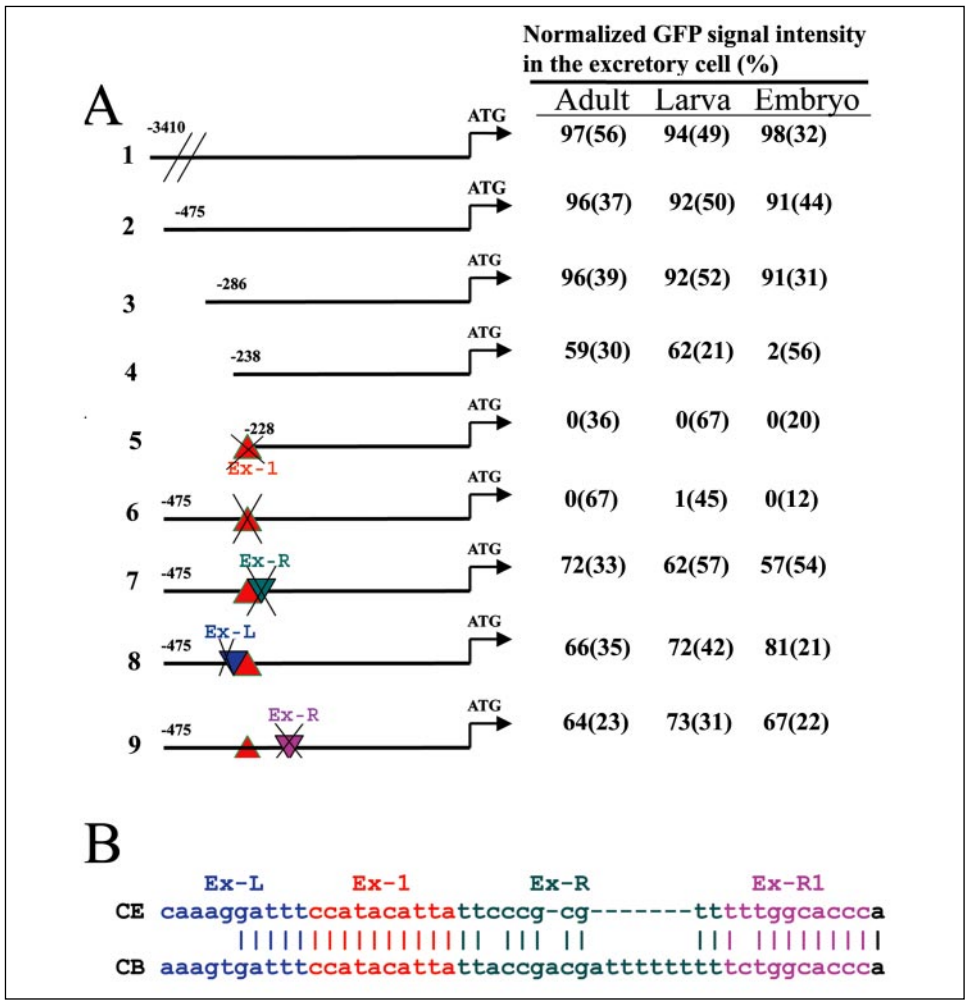

FIGURE 2. **Ex-1 is a critical element for excretory cell-specific expression within the *pgp-12* promoter.** *A*, deletion of a 10-bp motif, termed Ex-1 (*red triangle*), eliminated GFP expression at all developmental stages (*5*). Site-directed deletion of Ex-1 within 475-Ce-pgp-12 also significantly decreased GFP expression at all stages (*6*). Deletion of Ex-1 flanking sequences, Ex-L (*blue triangle*), Ex-R (*green triangle*), or Ex-R1 (*pink triangle*) decreased but did not eliminate GFP expression at any stage. The *crosses* denote deletion of the element. The GFP signal was normalized to that of integrated strain containing the 3410-bp *pgp-12* promoter fused to GFP. *Numbers* shown are % of expression level with sample size in the *parentheses*. *B*, alignment of the 500-bp Ce-pgp-12 (500-Ce-pgp-12) and its 500-bp *C. briggsae* orthologous region (500-Cb-pgp-12) using NCBI BLAST 2 Sequences with "Filter" off, word size 7, Expect 30, and gap extension 1. The 10-bp element, Ex-1 (*red*), Ex-L (*green*), Ex-R (*blue*), and Ex-R1 (*pink*) are highlighted.

protein. A fusion construct of the R05D11.3 promoter with a GFP gave exclusive expression in the intestine but not in the excretory cell (data not shown), suggesting that it may be a false positive cDNA clone. The other insert matches a cDNA sequence of *dcp-66*. It shows protein similarity to the deacetylase complex protein (24). The protein also contains a putative glucocorticoid receptor-like zinc finger DNA-binding domain (residues 432–501) identified by scanning the InterPro data base (25), suggesting that it is a DNA-binding protein. PSORT II (psort.ims.u-tokyo.ac.jp) predicted the nuclear localization of DCP-66 (73.9%, data not shown). Alignment of DCP-66 amino sequences with its homologs from other organisms revealed two conserved fragments. The alignment close to the amino terminus seems conserved only among vertebrates, whereas the one near the carboxyl terminus shows high conservation across all species and contains a putative zinc finger domain (supplemental Fig. S2).

dcp-66 Is Also Expressed in Excretory Cell and Binds Ex-1 in Vitro—To confirm the interaction of DCP-66 with Ex-1, we generated *promoter*::*GFP* transgenic animals for the gene. The 2525-bp promoter of the gene drove GFP expression in the excretory cell in adult, larva, and embryo. Weak expression was observed in the excretory cell, pharynx, vulva, and posterior neurons in adult worms (Fig. 3*A*). Strong expression in the excretory cell and weak expression in the pharynx were seen in larva (Fig. 3*B*). Embryonic expression was also seen in the excretory cell (Fig. 3*C*). Overlapping expression patterns of *dcp-66* and *pgp-12* suggest that the former regulate the later expression by binding of Ex-1 within its promoter. To examine whether the DCP-66 binds Ex-1 *in vitro*, we performed EMSA using biotin-labeled three tandem Ex-1 sequences as a probe. As shown in Fig. 3*D*, the 122 amino acids from DCP-66 that encodes the predicted zinc finger domain binds the Ex-1 element *in vitro*. Taken together, we demonstrate that DCP-66 binds Ex-1 both *in vivo* and *in vitro*.

DCP-66 Specifically Up-regulates pgp-12 Expression—To examine the regulatory activity of DCP-66 on Ce-pgp-12mediated excretory cell-specific expression *in vivo*, we crossed the integrated transgene, *pgp-12*::*GFP* into the homozygous *dcp-66* (*gk370*) mutant (see "Experimental Procedures"). Deletion of the gene resulted in significant suppression of *pgp-12* expression in the excretory cell (Fig. 4, *A* and *B*). The strain containing integrated *pgp-12*::*GFP* gains weak GFP expression in the pharynx. Compared with the wild type strain (Fig. 4*A*), loss of *dcp-66* significantly inhibited expressions in the excretory cells with little effect on GFP expression in the pharynx (Fig. 4*B*). To determine whether it is the loss of *dcp-66* that caused the decreased expression, we used a 9307-bp fragment spanning the N2 genomic region for *dcp-66* to rescue the homozygous mutant strain containing *pgp-12*::*GFP* (see "Experimental Procedures"). Three independent rescued lines were generated. Apparently increased *pgp-12*::*GFP* expression was observed in all the rescued lines as opposed to that of the *dcp-66* mutant (Fig. 4*C*, TABLE TWO). To test whether the loss of *dcp-66* can also down-regulate expression of other excretory cell-specific genes whose promoters do not contain the Ex-1 element, we chose *pgp-4* that also shows exclusive expression in the excretory cell (17) but does not contain any Ex-1 element. It is interesting to note that deletion of *dcp-66* does not affect the expression of *pgp-4*::*DsRed* (Fig. 4*D*), suggesting that *dcp-66* specif-

TABLE ONE

Examination of ectopic expression in the excretory cell mediated by Ex-1

Constructs	Fluorescence intensity (%)[a]		
	Intestine	Excretory cell	Pharynx
Ce-haf-4[b]	100 (37)[c]	0 (28)	ND[d]
Ex-1+Ce-haf-4	100 (54)	0 (41)	ND
Ex-L-Ex-1-Ex-R +Ce-haf-4	67 (52)	65 (55)	ND
48[e]+ Ex-L-Ex-1-Ex-R+Ce-haf-4	0 (18)	89 (63)	ND
48+Ce-haf-4	2 (63)	0 (26)	ND
24-L[f]+Ce-haf-4	94 (43)	0 (18)	ND
24-R[g]+Ce-haf-4	97 (32)	0(21)	ND
24-C[h]+Ce-haf-4	72 (28)	0 (17)	ND
Cb-48[i]+Ce-haf-4	96 (37)	0 (23)	ND
Ex-R-Ex-1-Ex-L[j]+Ce-haf-4	89 (26)	0 (15)	ND
Ex-L-Ex-1-Ex-R+Ce-haf-4 (498)	94 (30)	0 (27)	ND
48+Ex-L-Ex-1-Ex-R+Ce-haf-4 (498)	91 (52)	0 (17)	ND
48+Ce-haf-4 (498)	93 (53)	0 (24)	ND
Ce-pgp-14[b]	ND	0 (10)	100 (52)
Ex-L-Ex-1-Ex-R+ Ce-pgp-14	ND	0 (19)	95 (41)
48+Ex-L-Ex-1-Ex-R+ Ce-pgp-14	ND	0 (27)	99 (33)
Ce-pmp-1[b]	100 (20)	0 (19)	ND
Ex-L-Ex-1-Ex-R+Ce-pmp-1	89 (33)	0 (12)	ND
48+Ex-L-Ex-1-Ex-R+Ce-pmp-1	92 (37)	0 (21)	ND

[a] Fluorescence of transgenic animals containing Ce-haf-4 and its derivatives was normalized against that of strains BC10036 and BC10210 for intestine and excretory cell, respectively; fluorescence of transgenic animals containing Ce-pgp-14 and its derivatives was normalized against that of the strain BC06392 and BC12306 for excretory cell and pharynx, respectively; fluorescence of transgenic animals containing Ce-pmp-1 and its derivatives was normalized against that of the strain BC06392 and BC10364 for excretory cell and intestine, respectively.
[b] 286-, 1724-, and 1193-bp promoters were used for Ce-haf-4, Ce-pgp-4, and Ce-pmp-1, respectively, except otherwise indicated within the parentheses.
[c] Numbers shown are % of fluorescence intensity with the sample size in parentheses.
[d] ND, not determined.
[e] A 48-bp sequence upstream of Ex-L in Ce-pgp-12.
[f] A 24-bp sequence of 5′ end of the 48-bp sequence.
[g] A 24-bp sequence of 3′ end of the 48-bp sequence.
[h] A 24-bp sequence of central part of the 48-bp sequence.
[i] A 48-bp sequence upstream of Ex-1 in Cb-pgp-12.
[j] It is a reverse complementary sequence for Ex-L-Ex-1-Ex-R.

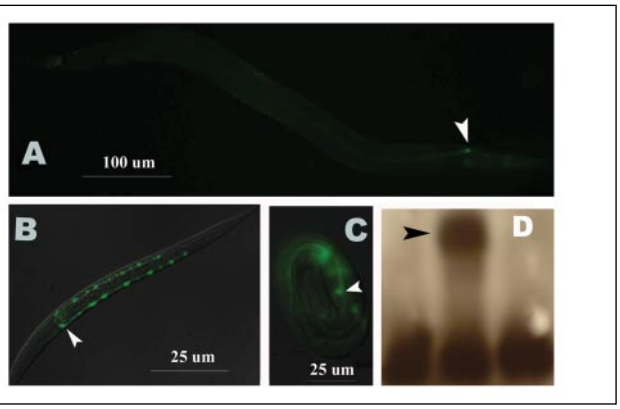

FIGURE 3. **dcp-66 is expressed in excretory cell at all stages and binds Ex-1 *in vitro*.** *A*, weak expression of *dcp-66* in the adult excretory cell (*arrowhead*), pharynx bulbs, posterior neurons, and vulvae. *B*, strong expression of *dcp-66* in the excretory cell (*arrowhead*) but weak in the pharynx in larvae. *C*, embryonic expression of *dcp-66* in the excretory cell (*arrowhead*). *D*, DCP-66 binds Ex-1 *in vitro* as determined by EMSA as described under "Experimental Procedures." *Left lane*, free labeled Ex-1; *middle lane*, labeled Ex-1 plus reticulate lysate, shifted bands are indicated by the *arrowhead*; *right lane*, competitive binding with ×500 molar excess of unlabeled Ex-1. Excretory cell bodies are indicated by *arrowheads*. *Scale bars* are shown in individual pictures.

ically up-regulated *pgp-12* expression via Ex-1. Given that DCP-66 interacts with Ex-1 both *in vivo* and *in vitro* and is expressed in the excretory cell at all stages, we conclude that DCP-66 is a novel transcription factor that up-regulates excretory cell-specific gene expression by binding to Ex-1. The homozygous deletion mutant for *dcp-66* is inviable and 18% of them arrest during embryogenesis and 82% die as small adults. Disintegrated tissues such as desiccated germline and shrunken intestine are frequently observed in adult, which might be caused by improperly regulated osmolarity, a major role of the excretory cell (11). Deletion of *dcp-66* also resulted in defective vulva (supplemental Fig. S3).

Promoters of Many Excretory Cell-specific Genes Do Not Contain Ex-1 Element—Given the mapped *cis*-regulatory element, we asked whether the genes co-expressed in the excretory cell are regulated by the same element. We define those as co-expressed genes if their promoters drive GFP expression exclusively or non-exclusively in the excretory cell from larvae to the adult stage. As an initial search for occurrence of Ex-1 in other excretory cell-specific 1-kb promoters, we scanned promoters of eight co-expressed genes whose promoters drive exclusive GFP expression in the excretory cell (19) for hits with over 80% identity to the Ex-1 sequence. Totally eight hits were found in four of the eight promoters, including Ce-pgp-12 and three other promoters (supplemental Fig. 4*A* and supplemental Table 1). Four of the co-expressed gene promoters do not produce any hit, implying they may not be regulated by binding of DCP-66 to Ex-1. This is supported by the observation that RNA interference against *dcp-66* showed a substantial inhibition of GFP expression driven by the promoters containing the Ex-1 hit, *i.e. pgp-12*, F48E8.3, DH11.3, and Y8G1A.2, but no significant inhibition for those whose promoters do not contain Ex-1 hits (data not shown), indicating that independent regulatory pathways exist for control of excretory cell-specific expression. We used the eight hits to generate the Ex-1 profile

and produced a sequence logo for this motif (supplemental Fig. 4*B*) using WebLogo (21). The resulting profile was used to scan a subset of 82 excretory non-exclusive (drive GFP expression in both the excretory cell and in other tissues) and 223 non-excretory cell gene promoters (drive GFP in posterior neurons but not in the excretory cell) (19). Surprisingly, only 10 of the 82 (12.2%) excretory non-exclusive promoters contain positive hits, suggesting again that there are other *cis*-regulatory elements independent of Ex-1 that are able to drive the expression of these excretory cell genes (supplemental Table 2). Notably, Ex-1 is significantly over-represented in excretory cell-specific promoters compared with non-excretory cell ones (12.2 *versus* 4.0%, $p < 0.05$, Fisher's Exact Test) (supplemental Fig. 4*A*, supplemental Table 3).

Different cis-Elements Can Mediate Similar Expression Patterns in Excretory Cell—To map the *cis*-elements from promoters that direct excretory cell-specific expression but do not contain Ex-1, we chose two such promoters for further analysis. One is for *pgp-3*, an excretory cell non-exclusive gene (expressed in both excretory cell and intestine (17)) with similar function annotation to *pgp-12*. The other one is *nas-31*, an excretory cell exclusive gene encoding putative astacin-like protein. The two genes were chosen to examine the regulatory elements for the shared expression patterns between related or unrelated genes. We used similar strategies as those used for *pgp-12* to map the critical elements in the two promoter sequences and were able to identify a 26- (Ex-2) and 25-bp element (Ex-3) within the promoters of *pgp-3* and *nas-31* with the following sequences, respectively: CTCACAAAATATAAATATGGTAATTC and GTTTCGAAAGTTCATCACCCCCAAC (Fig. 5, *A* and *B*). Deletion of the two individual elements abolished the GFP expression driven by each promoter. The two elements are strongly conserved between *C. elegans* and *C. briggsae*. They are not found in the promoter of *pgp-12*. Reciprocal introduction of *promoter*::*GFP* constructs for both *pgp-3*::*GFP* and *nas-31*::*GFP* into the two species yields the same expression patterns, indicating that regulatory mechanisms of excretory cell-specific expression between the two species are conserved. To examine the significance of the two elements in directing excretory cell-specific expression, we joined the individual Ex-2 or Ex-3 at the 5′ end of the 286-bp *haf-4* promoter followed by GFP. Both of the two fragments are able to yield ectopic expression in the excretory cell (Fig. 5, *A* and *B*). These results demonstrate the significant roles of the two elements in directing excretory cell-specific expression. Scanning of the two elements against the TransFac data base did not give any hits, implying the two elements are novel excretory cell-specific *cis*-elements. Both of the two elements show no obvious similarity to Ex-1 and sequences similar to the two elements are not found within the *pgp-12* promoter, indicating different *cis*-elements are able to mediate excretory cell-specific expression. To test whether excretory cell-specific expression of *pgp-3* or *nas-31* is under control of DCP-66, we crossed animals containing transgenes *pgp-3*::*GFP* or *nas-31*::*GFP* into the *dcp-66* null mutant VC924. The results show that loss of DCP-66 does not significantly reduce the expression level of the two transgenes (TABLE TWO). The slightly decreased GFP levels might be the consequences of *gk370* allele. In agreement with this observation, DCP-66 did not bind Ex-2 or Ex-3 *in vitro* (supplemental Fig. 5). Taken together, we conclude that Ex-1, Ex-2, and Ex-3 are all able to mediate similar expression in the excretory cell but through independent regulatory pathways.

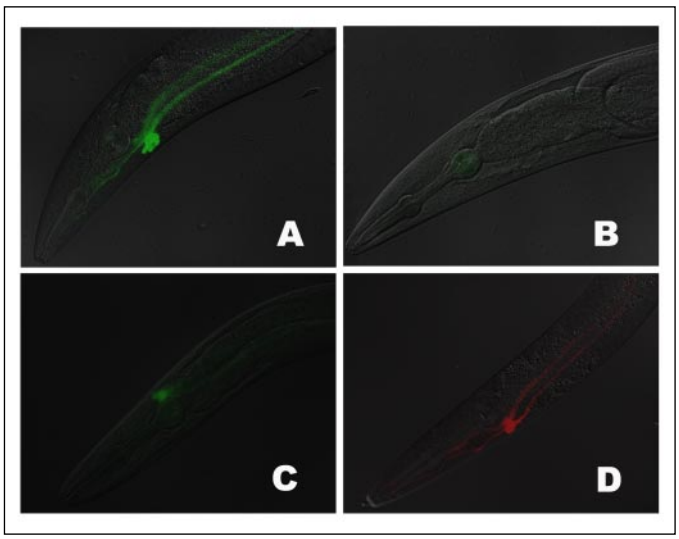

FIGURE 4. **Loss of *dcp-66* specifically suppresses the gene expressions in excretory cell.** *A*, *dpy-5*(*e907*), sIs10089 [*pgp-12*::*GFP*+pCes361(dpy-5(+))]. Expression of integrated *pgp-12*::*GFP* in the excretory cell (strong) and pharynx (faint); *B*, *dpy-5*(*e907*), *dcp-66* (*gk370*), sEX2020 [*pgp-12*::*GFP*+pCes361(dpy-5(+))]. Loss of *dcp-66* suppresses the expression of *pgp-12*::*GFP* in the excretory cell but not in the pharynx; *C*, *dpy-5*(*e907*), *dcp-66*, sEX2022 [*pgp-12*::*GFP*+pCes361(dpy-5(+))+p-*dcp-66*(genomic DNA). *pgp-12*::*GFP* expression in the excretory cell is recovered after rescue]; *D*, *dpy-5*(*e907*), *dcp-66*, sEX2020 [*pgp-4*::*DsRed*+pCes361(dpy-5(+))]. Loss of *dcp-66* does not affect the expression of *pgp-4*::*DsRed*. *Scale bars* represent 50 μm.

TABLE TWO

Effect of *dcp-66* on the expression of excretory cell specific transgenes

Transgenes *pgp-12*::*GFP*, *pgp-3*::*GFP*, and *nas-31*::*GFP* were crossed into *dcp-66 null* mutant allele *gk370* and the promoter lengths used here are 3410, 1050, and 1012 bp, respectively. Fluorescence intensity was measured for at least 50 worms for each strain and normalized to GFP intensity driven by the 3410-bp *pgp-12* promoter.

Transgene	Elements[a]	Genotype	Fluorescence intensity
			%
pgp-12::*GFP*	Ex-1	*dpy-5*(*e907*), sEX11257[*rCes-pgp-12*::*GFP*+pCes361]	100 (26)[b]
pgp-12::*GFP*	Ex-1	*dcp-66*(*gk370*), sEX11258[*rCes-pgp-12*::*GFP*+pCes361]	17 (54)[b]
pgp-12::*GFP*	Ex-1	*dcp-66*(*gk370*), sEX11259[*rCes-pgp-12*::*GFP*+pCes361,*rCes-dcp-66-wild-type-genomic-DNA*]	62 (43)[b]
pgp-3::*GFP*	Ex-2	*dpy-5*(*e907*), sEX11260[*rCes-pgp-3*::*GFP*+pCes361]	100 (52)[c]
pgp-3::*GFP*	Ex-2	*dcp-66*(*gk370*), sEX11261[*rCes-pgp-3*::*GFP*+pCes361]	67 (33)[c]
nas-31::*GFP*	Ex-3	*dpy-5*(*e907*), sEX11262[*rCes-nas-31*::*GFP*+pCes361]	100 (39)[d]
nas-31::*GFP*	Ex-3	*dcp-66*(*gk370*), sEX11263[*rCes-nas-31*::*GFP*+pCes361]	79 (41)[d]

[a] Element contained in the transgene promoters.
[b] Fluorescence was normalized against that of the transgenic strain carrying GFP fused to 3410-bp promoters for *pgp-12*.
[c] Fluorescence was normalized against that of the transgenic strain carrying GFP fused to 1050-bp promoters for *pgp-2*.
[d] Fluorescence was normalized against that of the transgenic strain carrying GFP fused to 1012 bp promoters for *nas-31*.

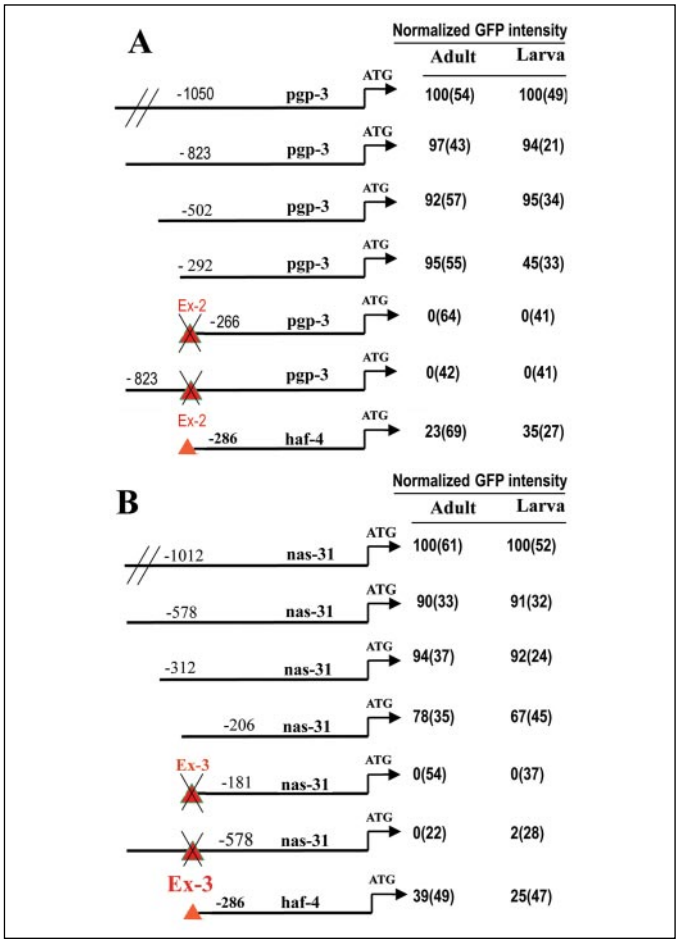

FIGURE 5. **Mapping of excretory cell-specific *cis*-elements from *pgp-3* and *nas-31* promoters (only expression in excretory cell shown).** *A*, the 1050-bp *pgp-3* promoter drives strong GFP expression in larvae and medium GFP expression in adult. The truncated version of the promoter containing the 823- or 502-bp sequence from the start codon did not significantly alter the expression. Decreased GFP expression was observed in larvae with the truncated 292-bp promoter. Truncation of another 26 bp, termed Ex-2, caused complete loss of expression in both adult and larvae. Site-directed removal of the Ex-2 within the 502-bp promoter also abolished the expression. *Numbers* shown are % of the expression level against that of 1050-bp promoter with the sample size in *parentheses*. *B*, a 1012-bp *pgp-3* promoter drives strong GFP expression in both larvae and adult. A 25-bp element, termed Ex-3, was mapped within the *nas-31* promoter using a similar strategy as described in *A*. *Numbers* shown are % of expression level against that of 1012-bp promoter with the sample size in the *parentheses*. Both Ex-2 and Ex-3 can cause ectopic expression in the excretory cell when joined at the 5′ end of 286-bp haf-4 promoter.

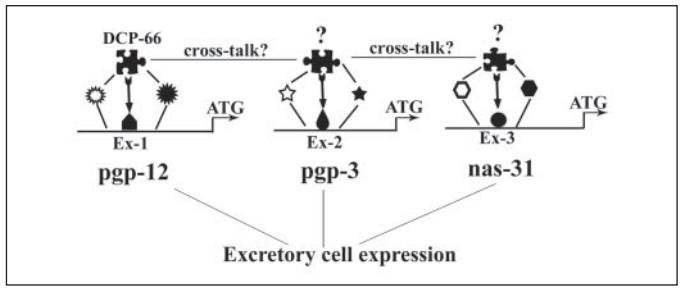

FIGURE 6. **Excretory cell-specific gene expressions mediated by distinct *cis*-regulatory elements.** Promoters of *pgp-12*, *pgp-3*, and *nas-31* all drive GFP expressions in excretory cell. Three distinct *cis*-elements, Ex-1, Ex-2, and Ex-3, were identified at −249, −267, and −189 each in the respective promoters. Putative transcription factors (*question mark*) that bind Ex-2 and Ex-3 are shown. Potential co-factors and cross-talks between transcription factors are also indicated.

DISCUSSION

With the availability of complete genome sequences, huge amounts of functional genomic data have been generated to determine functional links between genes. Gene co-expression is such a type of link that has been frequently used to determine the frequency of co-regulation to predict or infer the cellular function of unknown genes (9, 10). However, recent investigations show that large scale transcription factor binding sites by ChIP-on-chip experiments reveal very limited correlation between co-expression and the number of shared transcription factor binding sites (27, 28). This discrepancy may not be uncommon. For example, a terminal differentiation gene may be regulated by its upstream transcriptional factor, which is also expressed in the same tissue. The factor itself is subject to regulation by its own upstream components within the same transcriptional cascade. Many of the components within such a cascade could be found in the same tissue in certain time windows. However, it is unlikely for these genes share transcriptional elements. Problems could arise if such components are simply treated as co-expressed gene sets to search for *cis*-regulatory elements. Understanding the mechanisms for such discrepancy could improve the accuracy for functional prediction of genes based on co-expression data. Here we demonstrated that distinct *cis*-regulatory elements or regulatory pathways could mediate similar expression patterns in the excretory cell (Figs. 2, 5, and 6), which provides at least in part an explanation for the above discrepancy. Ex-1 can be found in only 12.2% of 82 excretory cell-specific promoters (supplemental Fig. 4*A*), indicating that 87.8% of excretory cell-specific expression is mediated by different *cis*-regulatory elements such as Ex-2 and Ex-3. Using a subset of the 82 excretory cell-specific promoters as the co-expression dataset will have a relatively low signal to noise ratio, *i.e.* 12.2%, which may prevent a *cis*-element discovery program from identifying biologically relevant *cis*-elements. This phenomenon might not be uncommon for other expression patterns. Poor performance of co-expression data by itself may also explain the observed discrepancy. Combination of expression data and comparative genomic data with functional annotations or subcellular localizations is reported to increase the likelihood to cluster co-regulated genes (29). It may increase the possibility to cluster the real co-expressed genes if data are collected from independent data sources, such as those from microarray, localization, and common transcription factor binding sites (10). Many of the co-expressed genes may not be truly co-regulated but only spuriously or superficially co-expressed. Poor correlation between co-expression and co-regulation may be even more evident in multicellular organisms because of the additional dimension (spatial) involved. However, we do not preclude the possibility that many of the carefully selected co-expression datasets do reflect the consequence of co-regulation, such as shared *cis*-elements and transcription factors. Therefore, caution should be exercised when choosing subsets of co-expressed genes for characterization of their *cis*-regulatory sites and binding factors. On the other hand, the presence of the same *cis*-regulatory element may not necessarily define the same expression patterns. For example, as shown in supplemental Fig. 4*A*, around 5% of the 223 neuronal but not excretory cell-specific promoters also contain elements with over 80% identity to Ex-1. It is possible that these putative elements cannot be recognized by DCP-66 or the presence of inhibitory elements around the Ex-1 that suppresses the Ex-1 function. It is well accepted that transcription factors usually function within a protein complex consisting of multiple components. Each component may bind its own cognate DNA element or work as the regulatory binding partner without interacting with chromosomal DNA. A specific expression pattern is determined by a combination of these components. A protein factor may works as an activator in one complex and suppressor in the other depending on the composition of the complex. This might be the case for DCP-66. Human and *Xenopus*

homologs of DCP-66 were identified as a component of a gene silencing complex (24, 30). However, we showed that DCP-66 works as a *trans*-activator in *C. elegans*. This can be partially explained by the fact that the protein shows significant conservation in the DNA binding domain but not in the other domain at the amino terminus that is only conserved among vertebrate homologs (supplemental Fig. 2). It is possible that the DCP-66 can play dual roles as both suppressor and activator, such as p53 does in cell growth and death decisions (31, 32). Given the significance of Ex-L and Ex-R in *pgp-12* expression (Fig. 2*A*) and inability of single Ex-1 to yield ectopic expression (TABLE ONE), we proposed that DCP-66 work synergistically with other partners that might bind elements Ex-L and/or Ex-R (Fig. 6). Further research needs to be done to test the two possibilities. It should also be noted that loss of DCP-66 did not completely eliminate the excretory cell-specific expression driven by the *pgp-12* promoter, suggesting some other factors may be involved in such a regulation. Identification of Ex-1, Ex-2, and Ex-3 as distinct elements in directing excretory cell-specific expression as well as differential responses of *pgp-12*::*GFP*, *pgp-3*::*GFP*, and *nas-31*::*GFP* expressing strains to loss of *dcp-66* indicate that there exist multiple pathways that regulate excretory cell-specific expression. Selective cross-talk between these pathways could exist to coordinate the similar expression patterns, which need to be further investigated (Fig. 6).

DCP-66 does not seem involved in the morphogenesis of the excretory cell because loss of the gene did not cause the obvious morphological changes in the cell (Fig. 4*D*). Neither *pgp-3* nor *pgp-12* are involved in morphogenesis of the excretory cell. They are recently duplicated ABC transporters and may be involved in xenobiotic resistance based on their homologs in human (33). Similar functional annotation and expression patterns between *pgp-3* and *pgp-12* suggest that the two related genes are possibly subject to subfunctionalization that is often seen in duplicated genes (34). *nas-31* encodes an astacin-like protein that might be involved in morphogenesis and pattern formation of excretory cell (35).

DCP-66 was also found in vulvae, pharynx, and posterior neurons, in addition to the excretory cell. Incomplete overlapping expression patterns of *dcp-66* and *pgp-12* suggest its possible regulatory roles in multiple tissues. These roles are consistent with previous observations (18, 36). Because the excretory cell is a functional equivalent of vertebrate kidney, the regulatory element and its binding factor identified here will provide a new entry point into investigating mechanisms of kidney development and its related diseases. *pgp-12* is a member of the ABCB subfamily of the ABC transporter genes in *C. elegans* whose homologs in human are multiple drug resistance proteins (33). Identification of *pgp-12* regulators will provide insights into modulation of multiple drug resistance that is a big challenge during chemotherapy.

Acknowledgments—We thank Robert Barstead for the C. elegans cDNA library and Wanyuan Ao, John Tyson, and Chris Beh for technical support during the yeast one-hybrid screening. We are also indebted to Ann Rose, Allan Mah, and Colin Thacker for the dpy-5 selection marker and strains. VC26 (gk19) and VC924 (gk370) were provided by the C. elegans Knockout Consortium.

REFERENCES

1. GuhaThakurta, D., Schriefer, L. A., Waterston, R. H., and Stormo, G. D. (2004) *Genome Res.* **14,** 2457–2468
2. Elgar, G. (2004) *Semin. Cell Dev. Biol.* **15,** 715–719
3. Vilaboa, N., Bermejo, R., Martinez, P., Bornstein, R., and Cales, C. (2004) *Nucleic Acids Res.* **32,** 6454–6467
4. Cui, M., and Han, M. (2003) *Dev. Biol.* **257,** 104–116
5. Spieth, J., MacMorris, M., Broverman, S., Greenspoon, S., and Blumenthal, T. (1988) *Dev. Biol.* **130,** 285–293
6. Chalfie, M., Tu, Y., Euskirchen, G., Ward, W. W., and Prasher, D. C. (1994) *Science* **263,** 802–805
7. Gilleard, J. S., Shafi, Y., Barry, J. D., and McGhee, J. D. (1999) *Dev. Biol.* **208,** 265–280
8. Wang, X., and Chamberlin, H. M. (2004) *Nat. Genet.* **36,** 231–232
9. Yeung, K. Y., Medvedovic, M., and Bumgarner, R. E. (2004) *Genome Biol.* **5,** R48
10. Snel, B., van Noort, V., and Huynen, M. A. (2004) *Nucleic Acids Res.* **32,** 4725–4731
11. Nelson, F. K., and Riddle, D. L. (1984) *J. Exp. Zool.* **231,** 45–56
12. Buechner, M., Hall, D. H., Bhatt, H., and Hedgecock, E. M. (1999) *Dev. Biol.* **214,** 227–241
13. Bürglin, T. R., and Ruvkun, G. (2001) *Development* **128,** 779–790
14. Stringham, E., Pujol, N., Vandekerckhove, J., and Bogaert, T. (2002) *Development.* **129,** 3367–3379
15. Jones, S. J., and Baillie, D. L. (1995) *Mol. Gen. Genet.* **248,** 719–726
16. Zhao, Z., Fang, L., Johnsen, R. C., and Baillie, D. L. (2004) *Biochem. Biophys. Res. Commun.* **323,** 104–111
17. Zhao, Z., Sheps, J. A., Ling, V., Fang, L., and Baillie, D. L. (2004) *J. Mol. Biol.* **344,** 409–417
18. Fraser, A, G., Kamath, R. S., Zipperlen, P., Martinez-Campos, M., Sohrmann, M., and Ahringer J. (2000) *Nature* **408,** 325–330
19. McKay, S. J., McKay, S. J., Johnsen, R., Khattra, J., Asano, J., Baillie, D. L., Chan, S., Dube, N., Fang, L., Goszczynski, B., Ha, E., Halfnight, E., Hollebakken, R., Huang, P., Hung, K., Jensen, V., Jones, S. J., Kai, H., Li, D., Mah, A., Marra, M., McGhee, J., Newbury, R., Pouzyrev, A., Riddle, D. L., Sonnhammer, E., Tian, H., Tu, D., Tyson, J. R., Vatcher, G., Warner, A., Wong, K., Zhao, Z., and Moerman, D. G. (2003) *Cold Spring Harbor Symp.* **68,** 159–170
20. Fire, A., Xu, S., Montgomery, M. K., Kostas, S. A., Driver, S. E., and Mello, C. C. (1998) *Nature* **391,** 806–811
21. Crooks, G. E., Hon, G., Chandonia, J. M., and Brenner, S. E. (2004) *Genome Res.* **14,** 1188–1190
22. Stein, L. D., Bao, Z., Blasiar, D., Blumenthal, T., Brent, M. R., Chen, N., Chinwalla, A., Clarke, L., Clee, C., Coghlan, A., Coulson, A., D'Eustachio, P., Fitch, D. H., Fulton, L. A., Fulton, R. E., Griffiths-Jones, S., Harris, T. W., Hillier, L. W., Kamath, R., Kuwabara, P. E., Mardis, E. R., Marra, M. A., Miner, T. L., Minx, P., Mullikin, J. C., Plumb, R. W., Rogers, J., Schein, J. E., Sohrmann, M., Spieth, J., Stajich, J. E., Wei, C., Willey, D., Wilson, R. K., Durbin, R., and Waterston, R. H. (2003) *PLoS Biol.* **1,** E45
23. Wingender, E., Chen, X., Fricke, E., Geffers, R., Hehl, R., Liebich, I., Krull, M., Matys, V., Michael, H., Ohnhauser, R., Pruss, M., Schacherer, F., Thiele, S., and Urbach, S. (2001) *Nucleic Acids Res.* **29,** 281–283
24. Wade, P. A., Gegonne, A., Jones, P. L., Ballestar, E., Aubry, F., and Wolffe, A. P. (1999) *Nat. Genet.* **23,** 62–66
25. Mulder, N. J., Apweiler, R., Attwood, T. K., Bairoch. A., Barrell, D., Bateman, A., Binns, D., Biswas, M., Bradley, P., Bork, P., Bucher, P., Copley, R. R., Courcelle, E., Das, U., Durbin, R., Falquet, L., Fleischmann, W., Griffiths-Jones, S., Haft, D., Harte, N., Hulo, N., Kahn, D., Kanapin, A., Krestyaninova, M., Lopez, R., Letunic, I., Lonsdale, D., Silventoinen, V., Orchard, S. E., Pagni, M., Peyruc, D., Ponting, C. P., Selengut, J. D., Servant, F., Sigrist, C. J., Vaughan, R., and Zdobnov, E. M. (2003) *Nucleic Acids Res.* **31,** 315–318
26. Kim, S. K., Lund, J., Kiraly, M., Duke, K., Jiang, M., Stuart, J. M., Eizinger, A., Wylie, B. N., and Davidson, G. S. (2001) *Science* **293**, 2087–2092
27. Lee, T. I., Rinaldi, N. J., Robert, F., Odom, D. T., Bar-Joseph, Z., Gerber, G. K., Hannett, N. M., Harbison, C. T., Thompson, C. M., Simon, I., Zeitlinger, J., Jennings, E. G., Murray, H. L., Gordon, D. B., Ren, B., Wyrick, J. J., Tagne, J. B., Volkert, T. L., Fraenkel, E., Gifford, D. K., and Young, R. A. (2002) *Science* **298,** 799–804
28. Qian, J., Lin, J., Luscombe, N. M., Yu, H., and Gerstein, M. (2003) *Bioinformatics* **19,** 1917–1926
29. Allocco, D. J., Kohane, I. S., and Butte, A. J. (2004) *BMC Bioinformatics* **5,** 18
30. Feng, Q., Cao, R., Xia, L., Erdjument-Bromage, H., Tempst, P., and Zhang, Y. (2002) *Mol. Cell. Biol.* **22,** 536–546
31. el-Deiry, W. S. (1998) *Semin. Cancer Biol.* **8,** 345–357
32. Johnson, R. A., Ince, T. A., and Scotto, K. W. (2001) *J. Biol. Chem.* **276,** 27716–27720
33. Sheps, J. A., Ralph, S., Zhao, Z., Baillie, D. L., and Ling, V. (2004) *Genome Biol.* **5,** R15
34. Lynch, M., and Force, A. (2000) *Genetics* **154,** 459–473
35. Mohrlen, F., Hutter, H., and Zwilling, R. (2003) *Eur. J. Biochem.* **270,** 909–4920
36. Simmer, F., Moorman, C., van der Linden, A. M., Kuijk, E., van den Berghe, P. V., Kamath, R. S., Fraser, A. G., Ahringer, J., and Plasterk, R. H. (2003) *PLoS Biol.* **1,** E12