A saturation screen for the cis-acting regulatory DNA of Hox genes of *Ciona intestinalis*

David N. Keys[1+], Byung-in Lee[2+], Anna Di Gregorio[1], Naoe Harafuji[1], J. Chris Detter[2], Mei Wang[2], Orsalem Kahsai[2], Sylvia Ahn[2], Andre Arellano[2], Quin Zhang[2] Stephan Trong[2], Sharon Doyle[2], Noriyuki Satoh[3], Yutaka Satou[3], Hidetoshi Saiga[4], Allen Christian[5], Dan Rokhsar[2], Trevor Hawkins[2], Mike Levine[1], and Paul M.Richardson[2*].

[+] These authors contributed equally to the work presented in this paper.
[1] Division of Genetics & Development, Department of Molecular and Cell Biology, University of California at Berkeley, Koshland Hall, Berkeley, California 94720.
[2] US Department of Energy Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, California, 94598.
[3] Department of Zoology, Graduate School of Science, Kyoto University, Sakyo-ku, Kyoto 606-8502, Japan.
[4] Department of Biological Science, Tokyo Metropolitan University, Hachiohji 192-0397, Tokyo, Japan.
[5] Biology and Biotechnology Research Program, Lawrence Livermore National Laboratory, 7000 East Avenue, PO Box 808, L-452, Livermore, CA. 94551
* Corresponding author

David N. Keys- dkeys@uclink4.berkeley.edu
Byung-in Lee- Lee110@llnl.gov
Anna Di Gregorio- annadg@uclink4.berkeley.edu
Naoe Harafuji- harafuji@uclink.berkeley.edu
Chris Detter- detter2@llnl.gov
Mei Wang- mwang@lbl.gov
Orsalem Kahsai- ojkahsai@lbl.gov
Sylvia Ahn- ahn2@llnl.gov
Andre Arellano- arellano4@llnl.gov
Quin Zhang- QZhang@lbl.gov
Stephan Trong- trong1@llnl.gov
Sharon Doyle- sadoyle@lbl.gov
Noriyuki Satoh- satoh@ascidian.zool.kyoto-u.ac.jp
Yutaka Satou- yutaka@ascidian.zool.kyoto-u.ac.jp
Hidetoshi Saiga- hidetoshi@c.metro-u.ac.jp
Allen Christian- christian4@llnl.gov
Dan Rokhsar- dsrokhsar@lbl.gov
Trevor Hawkins- tlhawkins@lbl.gov
Mike Levine- mlevine@uclink4.berkeley.edu
Paul Richardson- pmrichardson@lbl.gov

A screen for the systematic identification of cis-regulatory DNA within large (>100kb) genomic domains was performed using the simple chordate *Ciona intestinalis*. Randomly generated 3kb DNA fragments from Bacterial Artificial Chromosomes containing 2 groups of *Hox* genes were inserted into a vector upstream of a minimal promoter and *lacZ* reporter gene. 221 resultant fusion genes were separately electroporated into fertilized eggs and their regulatory activities were monitored in larvae. At least 21 separable cis-regulatory modules were found, including 8 new *Hox* enhancers that direct localized patterns of *CiHox2*, *CiHox3*, *CiHox4*, *CiHox5*, *CiHox12* and *CiHox13*.

Currently, many protein coding regions can be assigned a potential function based on simple sequence inspection. This is due to both the large number of proteins that have been experimentally characterized and the evolutionary conservation of the motifs essential to their functions. In contrast, it is not yet possible to infer gene expression patterns based on sequence analysis, nor is it known whether coordinately regulated genes share a common "language" or employ any type of defined cis-regulatory code. Since a large percentage of eukaryotic genomic DNA is utilized to mediate gene regulation, unlocking such a code, if it exists, represents one of the greatest challenges in revealing the function of complex genomes. In humans, only 3% of the genome corresponds to protein coding sequences (1, 2). There may be upwards of 1000 MB of cis-regulatory DNA (3, 4, 5). Despite this, fewer than 100 cis-regulatory DNAs have been characterized in the context of transgenic metazoans (6). This is primarily due to the difficulties and complexities involved in current methods (transgenic flies, mice, worms, etc). Until a larger data set of cis-regulatory DNA sequences can be produced, there is little hope of bringing the tools of bioinfomatics to bear on cracking the cis-regulatory code.

Here we assess the feasibility of conducting large scale searches for cis-regulatory DNA using a high-throughput screening method. This method is made possible through several experimental advantages afforded by *Ciona intestinalis*, a protochordate. The first such advantage is that plasmids containing coding regions linked to native or heterologous cis-regulatory DNAs are accurately transcribed, both spatially and temporally, when transformed into *C. intestinalis* embryos. There is no need for germline transformation since transgenes must persist for fewer than 20 hours and less than 12 rounds of cleavage before embryos develop into swimming tadpole larvae containing phylotypic chordate tissues. The second major advantage of working with *C. intestinalis* is that it is possible to simultaneously transform hundreds of embryos with recombinant plasmid DNA using simple electroporation techniques (7, 8). Thus, it is practical to systematically screen large numbers of DNA fragments for cis-regulatory activity without the logistical difficulties of maintaining large populations of transgenic strains. The primary goals of the experiments described in this report are to identify cis-regulatory modules in the *C. intestinalis Hox* cluster, determine whether they are genuine enhancers and to assess the possibility of extending this high-throughput methodology to large scale screens.

In this study, we chose to screen DNA isolated from the *C. intestinalis Hox* complex. There are three reasons for targeting *Hox* cis-regulatory DNAs. First, the strong conservation of the homeobox DNA-binding motifs allows the unambiguous identification of *Hox* genes. Second, *Hox* genes have been cloned in a broad spectrum of phyla and with few exceptions are found to be linked in single complexes with few or no intervening non-*Hox* genes. Third, and most importantly, the expression patterns of *C. intestinalis Hox* genes were either known, *CiHox3* (9) and *CiHox5* (10), or could be predicted based on studies in related organisms. Thus, it was considered likely that *bona fide* enhancers could be easily recognized.

Two BACs containing *CiHox* genes were isolated. The BACs, designated MRD and MRE, were found to contain 5 putative *CiHox* genes and a minimum of 10 additional genes. (Fig. 1A). The BAC designated MRD contains putative *CiHox2*, *CiHox3* and *CiHox4*. These share the same orientation and appear to have no intervening genes. The MRE BAC contains *CiHox12* and *CiHox13* aligned in a divergent orientation with no intervening gene. Regions containing the predicted *CiHox4*, *CiHox12* and *CiHox13* genes matched homeodomain containing DNAs previously isolated from a *C. intestinalis* genomic library (11) and the entire sequence of previously isolated CiHox3 cDNA (9) was present. Although the BAC sequences do not overlap, examination of their chromosomal positions by FISH analysis shows that these two DNAs come from neighboring regions on a single chromosome (Fig. 1B).

The organization of the *C. intestinalis Hox* genes differs from *Hox* complexes seen in other higher metazoans in several significant ways. First, other Chordate complexes exhibit tight linkage of their *Hox* genes and rarely contain large numbers of intervening non-transcription factors. Additionally, in most Chordates and non-chordate examples, linked *Hox* genes share the same transcriptional orientation. Because of the similarities between the previously described *Hox* complexes, the most parsimonious explanation of the exceptional structure of the *C. intestinalis Hox* complex is that it arose after its lineage diverged from the higher chordates. These evolutionary changes in the *Hox* complex could have been closely associated with body plan simplification that may have occurred within the urochordates. In this context, it will be interesting to see how the structure of the *C. intestinalis Hox* complex compares with those from other .

A library of random ~3kb genomic DNA fragments from MRD and MRE was created in a plasmid containing a *lacZ* reporter gene under the control of a minimal *C.intestinalis forkhead* (*fkh*) basal promoter (12, 13). Recombinant plasmids were separately electroporated into developing embryos and tested for localized patterns of LacZ activity. Because the *fkh* basal promoter is insufficient to induce expression, LacZ activity is indicative of cis-regulatory activity from the insert DNA.

Using this strategy, the two BACs that contained *CiHox* genes were systematically tested for cis-regulatory activity in *C. intestinalis* embryos (Fig. 1A). 221 subclones were individually electroporated into *C. intestinalis* embryos. 29 of these plasmids exhibited specific LacZ staining patterns in tadpole-stage embryos (Fig 1A). Many of these positives come from overlapping regions of the genome and give identical expression patterns, suggesting that they share a common cis-regulatory activity. Taking these into account, it appears that 21 independent cis-regulatory DNAs were identified. Careful inspection of overlaps between fragments with and without cis-regulatory activity reveals several DNA fragments that would have been expected to give LacZ activity but did not. It is not currently known whether these negatives are due to biologically relevant differences in the fragments or are experimental artifacts. This indicates that, while it is clearly capable of discovering DNAs with cis-regulatory activity, this screening method will not elucidate all such modules. As was to be expected based on the presence of non-*Hox* genes in this survey, many of the positive fragments are positioned closest to non-*Hox* genes. It is difficult to assess whether these cis-regulatory DNAs represent authentic enhancers as these flanking genes are uncharacterized. Therefore, the remainder of this analysis will focus on the putative *Hox* enhancers.

A typical *Hox* complex contains 9 or 10 linked genes that are sequentially expressed along the anterior-posterior axis in metazoan embryos. Cephalochordates (i.e., *Amphioxus*) and vertebrates exhibit localized patterns of *Hox* gene expression in the neural tube and derivatives of the mesoderm (usually paraxial mesoderm such as somites). In contrast, arthropods exhibit localized expression in the ventral nerve cord, visceral mesoderm and epidermis. Despite intensive analysis of *Hox* gene regulation, only ~20 *Hox* enhancers have been previously characterized in vertebrates and fruit flies. Within the 100kb that contain the two *CiHox* gene groupings, 11 independent cis-regulatory DNAs were identified. A twelfth enhancer was identified 5' of the previously described (10) *CiHox5* gene.

Expression patterns mediated by some of these 12 potential *Hox* enhancers (Fig. 2) are distinct from those seen in vertebrates. Three of the putative enhancers direct muscle expression, which is not seen in either vertebrates or arthropods. Muscle expression of *Hox* genes may be a situation that is unique to urochordates. Alternatively, the muscle staining may be spurious or associated with regulation of flanking non-*Hox* genes as discussed bellow.

The endogenous expression patterns of *CiHox3, CiHox4*, *CiHox5* and *CiHox12* have been characterized by *in situ* hybridization, *CiHox4* and *CiHox12* as part of this study, *CiHox3* and *CiHox5* previously (9, 10). It would appear that authentic enhancers have been identified for three of these 4 genes, *CiHox4*, *CiHox5*, and *CiHox12*. At 18 hours post-fertilization, *CiHox4* is expressed in the trunk lateral cells (Fig. 3A,B). These clusters of mesoderm cells flank the junction between the posterior cerebral vesicle and the anterior neural tube and give rise to multiple adult tissues including blood (14). At least one other *Hox* gene, *CiHox5*, is also known to be expressed in these cells (10). The xni178 DNA, which is located 3' of the *CiHox4* transcription unit, activates *lacZ* transcription in the same domain (Fig. 2; 3C, D), suggesting that the enhancer for this gene has been identified.

CiHox5 is expressed in the lateral cells of the nerve cord (10). The expression pattern extends from the boundary between the trunk and tail through the anterior fourth of the tail. Although not part of the random screen of the two BACs, a DNA fragment from the 5' of the *CiHox5* transcription unit, xow730, was found to activate the same pattern of expression (Fig. 2), indicating that it includes the authentic enhancer for *CiHox5*.

In early tailbud embryos, *CiHox12* is expressed in two different domains in the posterior tail, the epidermis and the neural tube (Fig. 3E, F, G), with the CNS expression pattern extending further to the anterior than the epidermal pattern. As with *CiHox5* (10), the CNS expression of *CiHox12* is restricted to the lateral cells (Fig. 3G). The *CiHox12* locus is flanked by two cis-regulatory DNAs. The first of these, xne165, activates *lacZ* expression in the epidermis of the posterior third of the tail, while the second, xne345, drives expression in the posterior half of the neural tube (Fig. 2; 3H, I). Thus the composite *CiHox12* expression pattern appears to be accounted for by two separate enhancers.

CiHox3 has been show to be expressed in hatched tadpoles in the visceral ganglion (CNS at the trunk/tail junction) (9). The present survey identified 4 DNA fragments flanking the *CiHox3* locus that exhibit cis-regulatory activity. However, none of these directed expression in the CNS. Specifically, the overlapping DNAs xni337 and xni213 both produce expression in the tail muscles, and probably represent a single module (Fig.2). The overlapping xni291 and xni012 DNAs direct expression in the tail muscles and epidermis, respectively (Fig.2). While there is no *a priori* reason to believe the muscle expression represents authentic *Hox* enhancer activity, the epidermal expression driven by xni012 exhibits restricted expression along the anterior-posterior axis, which is typical of arthropod *Hox* genes. DNA fragments that overlap xni291 and xni012 were previously shown to direct expression in the visceral ganglion, which is an authentic site of *CiHox3* expression (9). There are several potential explanations for the distinct expression patterns observed for xni291 and xni012. First, the heterologous *Ci-fkh* promoter that was used in this screen might fail to mediate authentic patterns of expression in response to certain enhancers. Second, the muscles might represent a hotspot for spurious expression using these methods. Previous work has identified the CNS cells ventral to the otolith and ocellar cells (pigment sensory organs) as such a hotspot (Harafuji, *et al*. submitted). A third possibility is that the muscle enhancers are authentic and regulate some of the neighboring non-*Hox* genes, such as *Nebulin*, whose vertebrate orthologs are known to express in muscle (15). Under this model, these muscle enhancers would normally be prevented from activating the *Hox* genes by repressor elements located elsewhere. In this regard, we note that no muscle enhancers were obtained with DNA fragments from the MRE BAC.

There is currently no *in situ* localization data for *CiHox2* or *CiHox13*, so it is not certain whether the remaining fragments containing cis-regulatory activity direct authentic components of their endogenous expression patterns.

In total, 8 of the putative enhancers identified in this screen direct sequential patterns of expression along the anterior-posterior axis of the CNS and epidermis and appear to be authentic *Hox* enhancers (Fig. 4). The sequential patterns of expression in the nerve cord are similar to those observed in other chordates. In contrast, co-linear expression in the epidermis has not been observed in other members of this phyla. *Drosophila* embryos exhibit localized expression in the epidermis, however these patterns are not directed by dedicated enhancers, but are produced by "stripe" enhancers that regulate expression in both the epidermis and CNS. Thus, the regulation of *C. intestinalis Hox* genes in the epidermis might represent a specific adaptation of urochordates that is not generally seen in metazoans. Alternatively, Hox expression in both the epidermis and CNS could be the basal state, and the vertebrate and arthropod lineages have retained different aspects of this activity since their divergence.

In addition to identifying previously undescribed *Hox* enhancers, this study demonstrates the feasibility of systematic screens for tissue-specific enhancers in *C. intestinalis*. Given the estimated size of the *C. intestinalis* genome (150-160 megabases, (16, Rokhsar, *et al.* unpublished), the identification of as many as 21 cis-regulatory modules in an interval of 368kb (~1 per 20kb) suggests that there may be as many as 8000 such modules which could be found using this method. The identification of even a small percentage of these will provide a rich data pool for detailed characterization of cis-regulatory mechanisms using both experimental and bioinfomatic approaches.

Fig. 1. Predicted structure of two *Hox* containing BACs of *C. intestinalis* genomic DNA aligned against clones screened for cis-regulatory activity. A BAC library made from *C. intestinalis* collected in Maizuru, Japan was hybridized with probes prepared from cosmids that including putative *CiHox2*, *CiHox3*, *CiHox4* and *CiHox12*. (**A**) Two *Hox* containing BACs, designated MRD and MRE, were isolated. FISH was performed with standard techniques (17) on cells disassociated from 15hr embryos using labeled DNA from MRD and MRE. Detection of overlapping signals indicates that the BACs co-localize on the same chromosome. (**B**) The BACs were sequenced and assembled using the hierarchical shotgun sequencing strategy (18, 19, 20) with ABI3700 DNA DNA sequencers. Gene models were predicted using Genscan (21) and by Blast alignments to known genes (22). The BACs contain the indicated *CiHox* genes, but do not overlap. For each BAC, a plasmid library was built such that random ~3kb inserts were placed 5' of a basal promoter driving the *lacZ* marker gene (13). 221 clones were picked and their positions on the BACs were identified (orientations of the inserts relative to the basal promoter in the vectors are indicated by dots). Plasmids were individually electroporated into single cell *C. intestinalis* embryos (24). 29 of the clones, shown in red, were found to direct distinct, repeatable expression of the *lacZ* reporter gene.

Fig. 2. Cis-regulatory activity from genomic fragments neighboring *Hox* genes. 15 fragments which were tightly linked to *Hox* loci exhibited distinct domains of LacZ activity, indicating that the cloned fragments contained cis-regulatory activity. Specific

tissues and domains of expression are represented diagrammatically: CNS = blue, epidermis = green, gut = yellow, muscle = orange, trunk lateral cells = red. Note, status of linkage of *CiHox5* to other *Hox* genes is not currently known.

Fig. 3. Characterization of native *CiHox4* and *CiHox12* transcription by whole mount *in situ* hybridization. Digoxigenin-labeled RNA antisense probes synthesized from exon regions of *CiHox4* (**A**,**B**) and *CiHox12* (**E**, **F**, **G**) were hybridized to *C. intestinalis* embryos as described previously (7). (**A**, **B**) In late tailbud embryos (18 hours post fertilization at 15°), *CiHox4* transcription is detected in the trunk lateral cells, a domain which is identical to that driven by the xni178 construct (**C**, **D**). (**E**, **F**, **G**) In early tailbud embryos (**E**, **G**: 10 hours post fertilization at 15°; **F**: 12 hours post fertilization at 15°), *CiHox12* transcript is detected in both the posterior neural tube (blue arrows) and the posterior tail epidermis (green arrows). (**G**) Optical cross section at the approximate position of the line shown in (**E**) shows that the *CiHox12* CNS expression is only detected in the 2 lateral cells of the neural tube. (**G'**) Schematic representation of the staining shown in G. These two patterns are a summation of the domains driven by the xne345 (**H**, posterior neural tube, blue arrow) and xne165 (**I**, posterior tail epidermis, green arrow).

Fig. 4. Cis-regulatory domains tightly linked to *Hox* genes exhibit characteristic nested anterior/posterior expression in both the CNS and epidermis. (**A**) xni333 , which is internal to *CiHox2*, drives expression in the neural tube just anterior to the trunk/tail junction. (**B**) xow730, which is 5' to *CiHox5*, drives expression in the neural

tube to the posterior of the trunk/tail junction. (**C**) xne345, which lies between and 5' to both *CiHox12* and *CiHox13*, drives expression in the neural tube in the posterior tail. (**D**) xni200, which is internal to *CiHox2*, drives expression in the epidermis of the posterior trunk. (**E**) xni012, which is in the 5'-flank of *CiHox3*, drives expression in the anterior tail epidermis. (**F**) xne165, which is 3' of *CiHox12*, drives expression in the posterior tail epidermis.

Notes

(1) E. S. Lander, L. M. Linton, B. Birren, C. Nusbaum, M. C. Zody, J. Baldwin, K. Devon, K. Dewar, M. Doyle, W. FitzHugh, *et al*., *Nature* **409**, 860 (2001).

(2) J. C. Venter, M. D.Adams, E. W. Myers, P. W. Li, R. J. Mural, G. G. Sutton, H. O. Smith, M. Yandell, C. A. Evans, R. A. Holt, *et al*., *Science* **291**, 1304 (2001).

(3) L. Duret & P. Bucher, *Curr. Opin. Struct. Biol*. **7**, 399 (1997).

(4) R. C. Hardison, *Trends in Genetics* **16**, 369 (2000).

(5) P. Onyango, W. Miller, J. Lehoczky, C. T. Leung, B. Birren, S. Wheelan, K. Dewar, A. P. Feinberg, *Genome Research* **10**, 1697 (2001).

(6) E. H. Davidson, *Genomic Regulatory Systems: Development and Evolution* Academic Press, San Diego. (2001).

(7) J. C. Corbo, M. Levine, R. W. Zeller, *Development* **124**, 589 (1997).

(8) A. Di Gregorio, M. Levine, *Differentiation* In press (2002)

(9) A. Locascio, F. Aniello, A. Amoroso, M. Manzanares, R. Krumlauf, M. Branno, *Development* **126**, 4737 (1999).

(10) M. Gionti, F. Ristoratore, A. Di Gregorio, F. Aniello, M. Branno, R. Di Lauro, *Development, Genes & Evolution* **207**, 515 (1998).

(11) A. Di Gregorio, A. Spagnuolo, F. Ristoratore, M. Pischetola, F. Aniello, M. Branno, L. Cariello, R. Di Lauro, *Gene* **156**, 253 (1995).

(12) A. Di Gregorio, J. C. Corbo, M. Levine, *Developmental Biology* **229**, 31 (2001).

(13) A plasmid designed to screen for cis-regulatory activity, pCES (plasmid *Ciona* Enhancer Screen), was built in the p72-1.27 plasmid (7). The *C. intestinalis fkh* basal promoter, transcription start site, native initiator codon and lacZ reporter gene were isolated by means of PCR from the –2.6kb *Ci-fkh/lacZ* plasmid (12). The reaction used the Pfu polymerase (Stratagene) and the oligonucleotides FKH5A: CGCGGATCCCCATGGTCAGAGAAACAGATTTGTAACAGG, which contains a terminal BamHI site in it's 5' tail, and lacZ3A: GCTACCCGGGCCGAGCTCAGAAAAAATGACTGC. Both the resultant fragment and p72-1.27 were digested with BamHI and EcoRI (internal to the *lacZ* coding region) and ligated together to produce pCES. DNA from the MRD and MRE BACs was sheared by Hydroshear (Genemachines, San Carlos, CA) to generated random DNA fragments approximately 3.0kb in size. These fragments

were end-filled then ligated into the library vector upstream of the basal promoter. Colonies from this library were randomly picked and both ends of the inserts were sequenced using the TempliPhi Rolling Circle process (Detter, *et al*. submitted). The *fkh* basal promoter was chosen since the native *fkh* gene is known to be expressed in a large number of tissues (and is therefore unlikely to contain tissue-specific silencers) and it has been shown to produce faithful patterns of expression when combined with several heterologous enhancers (12).

(14) H. Nishida, N. Satoh, *Developmental Biology* **132**, 355 (1989).

(15) R. Horowitsa, L. Ganga, Q. Z. Jiana, A. Herreraa, *Advances in Biophysics* **33**, 143 (1996).

(16) M. W. Simmen, S. Leitgeb, V. H. Clark, S. J. M. Jones, A. Bird, *Proc. Natl. Acad. Sci. USA* **95**, 4437 (1998).

(17) A. T. Christian, H. E. Garcia, J. D. Tucker, *Mammalian Genome* **10**, 628 (1999).

(18) S. Anderson, *Nucleic Acids Research* **9**, 3015 (1981).

(19) R. C. Gardner, A. J. Howarth, P. Hahn, M. Brown-Luedi, R. J. Shepherd, J. Messing, *Nucleic Acids Research*. **9**, 2871 (1981).

(20) P. L. Deininger, *Anal Biochem.* **129**, 216 (1983).

(21) C. Burge, S. Karlin, *Journal of Molecular Biology* **268**, 78 (1997).

(22) W. Gish, D. J. States *Nature Genetetics* **3**, 266 (1993).

(23) J. C. Corbo, A. Erives, A. Di Gregorio, A. Chang, M. Levine, *Development* **124**, 2335 (1997).

(24) Adult *C. intestinalis* were collected from Pillar Point Harbor in San Mateo County, California under scientific permit of the State of California Department of Fish and Game. Electroporation, fixation and LacZ staining were carried out as previously described (23). Each electroporation utilized aliquots containing 100ug of experimental plasmid and 25ug of a control plasmid, -3.5*CiBra>GFP* (7). Embryos were allowed to develop for 18 hours post-fertilization at 15°C. Because of the intracellular stability of LacZ protein, staining in a given cell is indicative of *lacZ* transcription at some time during that cell's ontogeny, but is not necessarily indicative of active transcription at the time of fixation. Fixed embryos were prescreened for percentage of fully developed animals and expression of the control plasmid. Controls performed using plasmid without insert indicate that there are low levels of background expression in the mesenchyme (undifferentiated cells in the posterior trunk which contribute to post-metamorphosis mesodermal structures) and notochord. Additionally, the

ventral sensory vesicle (midbrain) appears to be an expression “hotspot” which exhibits expression with many inserts, but may not represent native enhancer activity (Harafuji, *et al*, submitted). Because of the clonal nature of plasmid incorporation in *Ciona* electroporation, individual embryos do not necessarily exhibit all of the expression domains seen with a given plasmid.