Genomic structure and functional control of the *Dlx3-7* bigene cluster

Kenta Sumiyama*, Steven Q. Irvine*, David W. Stock†, Kenneth M. Weiss‡, Kazuhiko Kawasaki‡, Nobuyoshi Shimizu§, Cooduvalli S. Shashikant¶, Webb Miller‖, and Frank H. Ruddle*,**

*Department of Molecular, Cellular, and Developmental Biology, Yale University, New Haven, CT 06511; †Department of Biology, University of Colorado, Boulder, CO 80309; Departments of ‡Anthropology, ¶Dairy and Animal Science, and ‖Computer Science, Pennsylvania State University, University Park, PA 16802; and §Department of Molecular Biology, Keio University School of Medicine, Tokyo 160, Japan

Contributed by Frank H. Ruddle, November, 1, 2001

The *Dlx* genes are involved in early vertebrate morphogenesis, notably of the head. The six *Dlx* genes of mammals are arranged in three convergently transcribed bigene clusters. In this study, we examine the regulation of the *Dlx3-7* cluster of the mouse. We obtained and sequenced human and mouse P1 clones covering the entire *Dlx3-7* cluster. Comparative analysis of the human and mouse sequences revealed several highly conserved noncoding regions within 30 kb of the *Dlx3-7*-coding regions. These conserved elements were located both 5′ of the coding exons of each gene and in the intergenic region 3′ of the exons, suggesting that some enhancers might be shared between genes. We also found that the protein sequence of *Dlx7* is evolving more rapidly than that of *Dlx3*. We conducted a functional study of the 79-kb mouse genomic clone to locate cis-element activity able to reproduce the endogenous expression pattern by using transgenic mice. We inserted a *lacZ* reporter gene into the first exon of the *Dlx3* gene by using homologous recombination in yeast. Strong *lacZ* expression in embryonic (E) stage E9.5 and E10.5 mouse embryos was found in the limb buds and first and second visceral arches, consistent with the endogenous *Dlx3* expression pattern. This result shows that the 79-kb region contains the major cis-elements required to direct the endogenous expression of *Dlx3* at stage E10.5. To test for enhancer location, we divided the construct in the mid-intergenic region and injected the *Dlx3* gene portion. This shortened fragment lacking *Dlx7*-flanking sequences is able to drive expression in the limb buds but not in the visceral arches. This observation is consistent with a cis-regulatory enhancer-sharing model within the *Dlx* bigene cluster.

The mammalian *Dlx* family consists of six genes with homeoboxes related to that of *Drosophila Distal-less* (*Dll*). The mammalian genes take the form of bigene clusters. The paired genes, termed *Dlx2–1*, *Dlx5–6*, and *Dlx3–7*, are organized in an inverted, convergently transcribed manner (1–3). It has been proposed that a single *Dlx* gene duplicated to form an ancestral inverted bigene cluster during the evolution of the early chordates and subsequent duplications gave rise to multiple bigene clusters (4). This hypothesis is based on phylogenetic clustering of *Dlx2, 3, 5* to the exclusion of *Dlx1, 6, 7*. It is likely that the three *Dlx* clusters underwent duplication together with the Hox gene clusters, because *Dlx2–1*, *5–6*, and *3–7* are closely linked to the Hox D, A, and B clusters, respectively (1, 3, 5). There is a certain degree of spatial and temporal expression overlap between the linked *Dlx* genes around embryonic stage (E) E10.5 (6). Models are proposed to explain this overlapping pattern, including cis-regulatory sharing between paired genes (7). In this report, we will focus on the *Dlx3–7* bigene cluster with the ultimate goal of testing the shared enhancer hypothesis of bigene expression control.

The mouse *Dlx3* gene is first expressed weakly in the rostral ectoderm, anterior to the neural plate during the head fold stage (6). *Dlx3* is expressed by E8.5 in the ectoplacental cone and chorionic plate, developing into a high level of expression by day E10.5 in the labyrinthine layer of the embryonic placenta (8). *Dlx3* is expressed in the first and second visceral arches and fronto-nasal ectoderm at E9.5 (9). Later in development, *Dlx3* is also expressed in the external respiratory epithelium of the nares, in whisker follicles, taste bud primordia, dental and mammary gland epithelia, apical ectodermal ridge (AER) of limb buds, genital primordia, and in several additional sites of epithelial-mesenchymal interaction (8, 9). By the end of embryonic development, *Dlx3* is down-regulated, except in the skin, where it is transcribed in stratified epidermis and in the matrix cells of the hair follicles (10). The *Dlx3* pattern of expression is different in some respects in comparison with the bigene clusters *Dlx2–1* and *Dlx5–6* in that *Dlx3* is not expressed in the central nervous system and is largely expressed in recently acquired structures common to the mammals. *Dlx7* has an overlapping expression pattern with the *Dlx3* gene in the visceral arches and limb buds before E10.5 (ref. 11 and unpublished data).

The *Dlx* genes provide an interesting model for the regulation of clustered genes because their arrangement in pairs is the simplest case of gene clustering. The fact that all of the mammalian clusters have the same genomic arrangement suggests that the clusters have not undergone extensive rearrangement since their formation and may help elucidate the fate of cis-regulatory elements after gene duplication. The *Dlx3–7* cluster is of special interest to us because its expression pattern lacks the complication of CNS expression. In this report, we provide detailed information on the structural organization of the *Dlx3–7* bigene cluster, identify conserved sequence elements in the coding and noncoding domains, and by means of functional transgenic analysis test the enhancer activity for *Dlx3–7* expression control.

Materials and Methods

P1 Artificial Chromosome (PAC) Clones and Sequencing. The clones P1–972 and P1–1490 were isolated from Incyte Genomics (Palo Alto, CA) mouse and human PAC libraries, respectively, by PCR screening of pooled libraries by using primers for the 3′ end of the *Dlx3/DLX3* homeodomain (3, 12).

Nucleotide sequences were primarily determined by using the shotgun method, and unsequenced gaps were filled by primer walking (see ref. 13 for methods). The sequence was confirmed by comparing a restriction map deduced from genomic sequence with an experimentally constructed restriction map.

Cluster Alignments and Sequence Comparisons. Alignments were computed with PIPMAKER (14) (available at http://bio.cse.

Abbreviations: E, embryonic stage; AER, apical ectodermal ridge; PAC, P1 artificial chromosome.

Data deposition: The sequences reported in this paper have been deposited in the GenBank database (accession nos. AF452637 and AF452638).

**To whom reprint requests should be addressed. E-mail: frank.ruddle@yale.edu.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

 www.pnas.org/cgi/doi/10.1073/pnas.012584999

psu.edu/pipmaker/). The BLAST 2 sequences program accessed through the National Center for Biotechnology Information website (15) was also used for identifying local sequence conservation.

Construction of the Mouse *Dlx3-lacZ* Reporter Gene. The 80-kb insert of the P1–972 was captured into the pPAC-ResQ vector by using homologous recombination in yeast (16) to make construct *Dlx37*-A8. In brief, 50 ng of linearized pPAC-ResQ vector and 400 ng of circular PAC P1–972 were cotransformed into yeast strain Y724 (17). More than 60% of the 200 clones obtained were identified as recombinant by hybridization tests by using the *Dlx3* 5′ upstream region as probe. Yeast genomic DNA from recombinant clones was isolated by using the PureGene kit (Gentra Systems), transformed into *Escherichia coli* DH10B cells by electroporation (16), and checked for rearrangement by restriction analysis.

The *lacZ* reporter gene was inserted in-frame into the first exon of the mouse *Dlx3* gene by using another round of homologous recombination in yeast. A targeting vector (pLZFSV-*Dlx3*AB) was generated in which a *lacZ-Ura3* cassette (18) was flanked by PCR generated recombinogenic ends (Fig. 4*a*) from 5′ of the *Dlx3*-coding region [primers KL001 (5′-GGT AAC AAC AAA GAG GGT TGA GA-3′) and KL 002 (5′-GAA GGA GCC GCT CAT GCT-3′)], and portions of exon 1 and the downstream intron [primers KL003 (5′-CAC GCG GCC GCC AGC ATC CTC ACC GAC ATC T-3′) and KL 004 (5′-GCG GCG GCC GCA AGC TTC CTC TTC GGA GTG TCG TTC A-3′)]. The resulting pLZFSV-*Dlx3*AB was linearized and transformed into yeast harboring construct *Dlx37*-A8. One recombinant clone (construct *Dlx37-lacZ*-79kb) was transformed into *E. coli* DH10B and verified by restriction analysis which confirmed the *lacZ* insertion and lack of genomic rearrangement (Fig. 4*b*). The 5′ integration site was sequenced to confirm proper integration. Construct *Dlx37-lacZ*-19kb was subsequently made by *Hin*dIII digestion of construct *Dlx37-lacZ*-79kb.

Production of Transgenic Mice. Inserts were released from the 9.8-kb vector by digestion of 150 μg with I-*Sce* I (Boehringer Mannheim) and purified by ultra centrifugation on a 10–40% sucrose gradient. Concentration of dialyzed DNA was adjusted to 4 ng/μl and injected into pronuclei of one-cell mouse embryos as described previously (19). Injected embryos were transferred into the oviduct of pseudopregnant CD1 mice, dissected at appropriate stages, and stained for β-galactosidase activity.

***In Situ* Hybridization.** Sense and antisense RNA probes were transcribed from a *Dlx3* partial cDNA (clone 62895–57). This 757-bp probe contains the most 3′ 27 bp of the homeobox, a further 300 bp of coding sequence, and 468 bp of 3′ untranslated region (positions 538-1295 of GenBank databank accession no. S81932). The protocol for *in situ* hybridization was modified from that of T. Sanders and C. Ragsdale (personal communication) and is available from the authors on request. Control embryos were hybridized with sense probe verify the specificity of the antisense probe signal.

Results

Genomic Organization of the *Dlx3–7* Bigene Cluster. We obtained human and mouse P1/PAC clones covering the entire *Dlx3–7* bigene clusters including flanking regions. The clones were completely sequenced: the human clone termed hP1–1490 measures 75,022 nt, whereas the mouse clone termed mP1–972 measures 78,651 nt.

The general genomic organization of the human and mouse *Dlx3–7* bigene clusters was determined by comparisons of genomic and cDNA sequences. The *Dlx7* and *Dlx3* genes in both species are transcribed convergently and contain three exons in agreement with our earlier studies (3). Thus, we can divide the bigene clusters into three domains: an intergenic 3′ region shared by both genes, a flanking 5′ upstream region unique to the *Dlx7* gene, and a flanking 5′ region unique to the *Dlx3* gene (Figs. 1 and 2). The intergenic regions measure 17 kb in both species. At least 7 kb of sequence upstream of the *Dlx3* translation start site and at least 37 kb upstream of the *Dlx7* translation start site are captured in both the human and mouse clones. A close comparison of the human and mouse clones showed no other genes in the P1 clones, although several expressed sequence tag sequences show similarity in the first 20 kb upstream of *Dlx7*.

We compared the mouse *Dlx3-7* bigene cluster to its human paralogs, namely, *DLX2–1* and *DLX5–6* (data not shown). The general genomic organization is similar among the clusters. All show convergent translation, three exons, and 5′-flanking and 3′-intergenic domains. The *DLX2–1* and *DLX5–6* intergenic regions both measure 10 kb. Sequence similarities are found only in the homeodomains, whereas other coding and noncoding domains show low sequence conservation. The absence of sequence similarity in the noncoding domains is interesting, considering the paralogous nature of the bigene clusters and their overlapping patterns of expression during development.

Protein Sequence Comparisons Between *Dlx7* and *Dlx3*. The nucleotide sequences of coding regions of the mouse and human exons was subjected to comparative analysis to determine the extent of functional evolutionary constraints on the *Dlx7* and *Dlx3* proteins. We calculated synonymous and nonsynonymous substitution rates for coding regions of both *Dlx7* and *Dlx3* genes between human and mouse in each exon by using the method of Nei and Gojobori (20). The result shows that the *Dlx3* gene coding sequence is highly conserved, whereas the *Dlx7* gene is much less conserved in all three exons (Table 1). Note that a nonsynonymous to synonymous ratio (d_N/d_S) of one is predicted for neutrally evolving genes. Both genes in all three exons show ratio

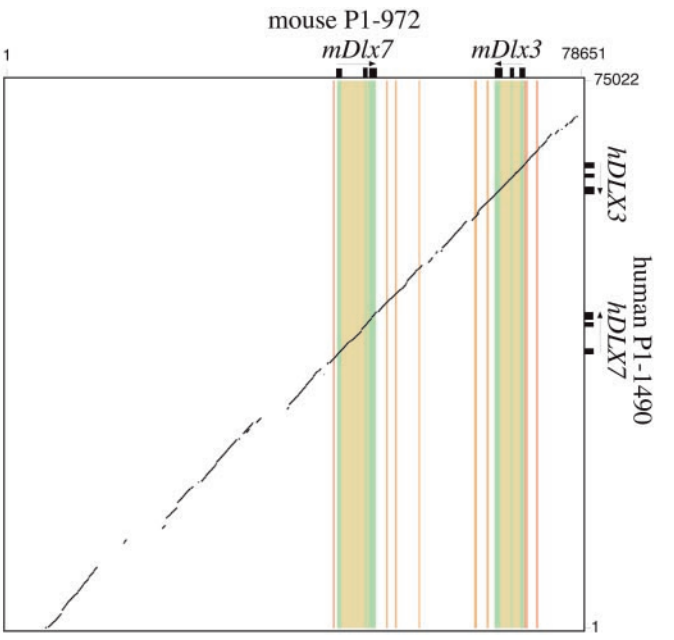

Fig. 1. Dot plot analysis between mouse *Dlx3–7* cluster (P1–972: horizontal line) and human *DLX3–7* cluster (P1–1490: vertical line). Black boxes indicate exons. Exon and intron positions of mouse *Dlx3* and *Dlx7* are color coded as green and yellow, respectively. *Dlx7*- and *Dlx3*-flanking conservation motifs (F7–1, F3–1,2) are indicated in red. Intergenic conservation motifs (from I37–1 to 5) are shown in orange.

Table 1. Genetic distances of synonymous and nonsynonymous substitutions between human and mouse *Dlx3* and *Dlx7* genes computed by using Nei and Gojobori's method

	Dlx3			*Dlx7*		
Region	$d_S \pm$ SE	$d_N \pm$ SE	d_N/d_S	$d_S \pm$ SE	$d_N \pm$ SE	d_N/d_S
Exon 1	0.34 ± 0.076	0.0084 ± 0.0060	0.025	0.62 ± 0.13	0.15 ± 0.030	0.24
Exon 2	0.56 ± 0.15	0.0 ± 0.0	0.0	0.65 ± 0.17	0.090 ± 0.026	0.13
Exon 3	0.39 ± 0.084	0.016 ± 0.0078	0.040	0.27 ± 0.080	0.13 ± 0.029	0.48

values significantly less than one, indicating that *Dlx3* and *Dlx7* are under functional constraint. However, all three exons of *Dlx7* have ratio values much higher than *Dlx3*, indicating that the *Dlx7* protein is evolving more rapidly than the *Dlx3* protein. We will consider the implications of this situation in the *Discussion*.

Noncoding Genomic Domain Comparisons Between the Mouse and Human *Dlx3–7* Bigene Clusters. We carried out several types of homology plot analyses between the human and mouse *Dlx3–7* bigene clusters to identify putative cis-regulatory elements. We predict that cis-regulatory elements will be highly conserved unless the gene expression patterns have changed dramatically during species divergence. First, we performed dot plot analysis between mouse and human (Fig. 1) and observed similarities in most regions except for the mouse 12- to 21-kb region (Fig. 2) equivalent to the human map position 8–15 kb. Genomic rearrangement may have taken place in this region.

Second, we used percentage identity plot analysis to compare mouse (reference sequence) and human genomic sequences to find shared conserved motifs (Fig. 2).

An interesting aspect of the sequence comparison analysis is the striking conservation of a 30-kb region at mouse map position 44–74 kb that includes the *Dlx7* and *Dlx3* genes (Fig. 2). There are three strong matches of ≈85% similarity between human and mouse in the flanking domains (Fig. 2; red color code). There are also five conserved noncoding regions in the *Dlx3–7* intergenic region (Fig. 2, orange color code, and Fig. 3). Element I37–1 is the most highly conserved element with 90% identity between human and mouse. Element I37–5 appears to be a partial duplication of I37–1 with 80% identity >40 bp (Fig.

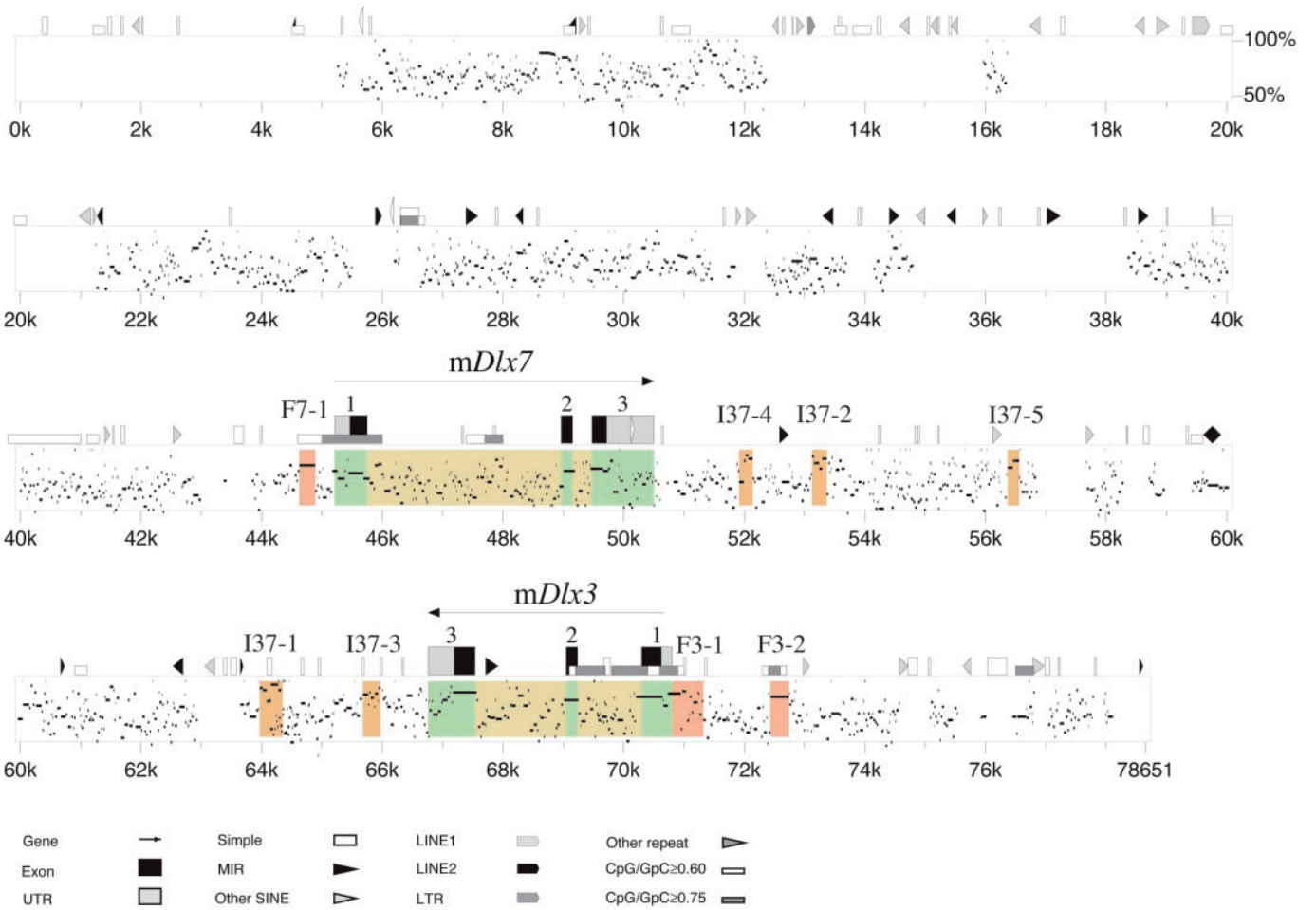

Fig. 2. Percentage identity plot of mouse *Dlx3–7* cluster (horizontal line) to human *DLX3–7* cluster. Vertical axis shows percentage of sequence identity from 50% (bottom) to 100% (top). Large black and gray boxes on the plot represent coding and untranslated region, respectively. Small rectangles and triangles are repetitive elements (see percentage identity plot legend). Exon and intron positions are color coded as green and yellow, respectively. *Dlx7* and *Dlx3* flanking conservation motifs (F7–1, F3–1, 2) are indicated in red. Intergenic conservation motifs (from I37–1 to 5) are shown in orange.

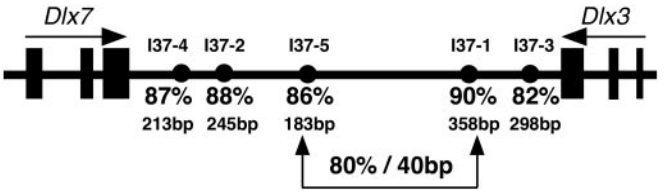

Fig. 3. Diagram showing locations and information of five highest sequence similarities in the human and mouse *Dlx3–7* intergenic region (from I37–1 to 5). Numbers below are length of conservation and the percentage of sequence identity within the alignment between human and mouse. A 40-bp region with 80% sequence identity is found between I37–1 and I37–5, as noted.

3). Element I37–1 also shows similarity with an element in the zebrafish *dlx3–7* intergenic domain (M. Ekker, unpublished data). The three remaining conserved elements, I37–2, I37–3, and I37–4 have identities of 88%, 82%, and 87%, respectively, between human and mouse (Fig. 3). All of these noncoding genomic regions are more highly conserved than the exon 1 coding domains (Fig. 2).

Functional Properties of Intergenic Domain Conserved Elements. We carried out functional characterization of the intergenic putative control elements to explore their specific regulatory features. The P1–972/79-kb insert was transferred into a yeast-bacterial shuttle vector, pClasper, by homologous recombination in yeast cells (17) (Fig. 4*a*). Homologous recombination was also used to insert a *lacZ* reporter gene together with a *ura3* yeast selection marker in-frame into the first exon of the *Dlx3* gene. The transfer of the large insert into pClasper and its subsequent modification into a reporter construct was successfully accomplished without apparent rearrangement (Fig. 4*b*). The resulting *Dlx37-lacZ*-79-kb reporter was excised from pClasper and injected into the pronuclei of mouse zygotes to produce transgenic animals. Two types of transgenic mice were studied: "transient" transgenic animals sampled after injection at various developmental stages, and "stable" transgenic animals obtained from stably transformed lines derived from independent founder transformants. The stable transformants are advantageous in that they can be analyzed reproducibly. We analyzed four transient embryos and five independently derived stable transformed lines.

The expression pattern of the *lacZ* reporter in the transient and stable embryos was largely similar but differed in some instances in signal intensity (Fig. 5). Reporter expression is found in fore and hind limb buds and tail bud at stage E9.5 (Fig. 5*bB*). Expression is also seen in the first and second visceral arches. Expression is observed in the AER of the limb buds and some ventral ectodermal cells on the fore and hind limb buds at stage E10.5. Strong expression is found in the mesenchyme of the distal caudal portion of the mandibular process of the first visceral arch and the distal lateral portion of the second visceral arch (Fig. 5*bE*). Both E9.5 and E10.5 embryos show similar reporter expression patterns in limbs and visceral arches in comparison to the endogenous patterns of expression as determined by *in situ* hybridization (Fig. 5 *bA* and *bD*). We conclude from these results that a considerable proportion of the endogenous expression of the *Dlx3* gene is reconstituted by the *Dlx37-lacZ*-79-kb reporter construct, implying that critical control elements necessary for the appropriate expression of *Dlx3* in the visceral arches and limb buds are present in the construct.

Other regions show differences between reporter and endogenous patterns of expression. First, expression in frontonasal ectoderm and rostral midline ectoderm of the first arch is detected in *in situ* hybridization experiments (Fig. 5 *bA* and *bD*) but not in transgenics (Fig. 5 *bB* and *bE*). Second, some transgenic embryos show *LacZ* expression on the dorsal midline,

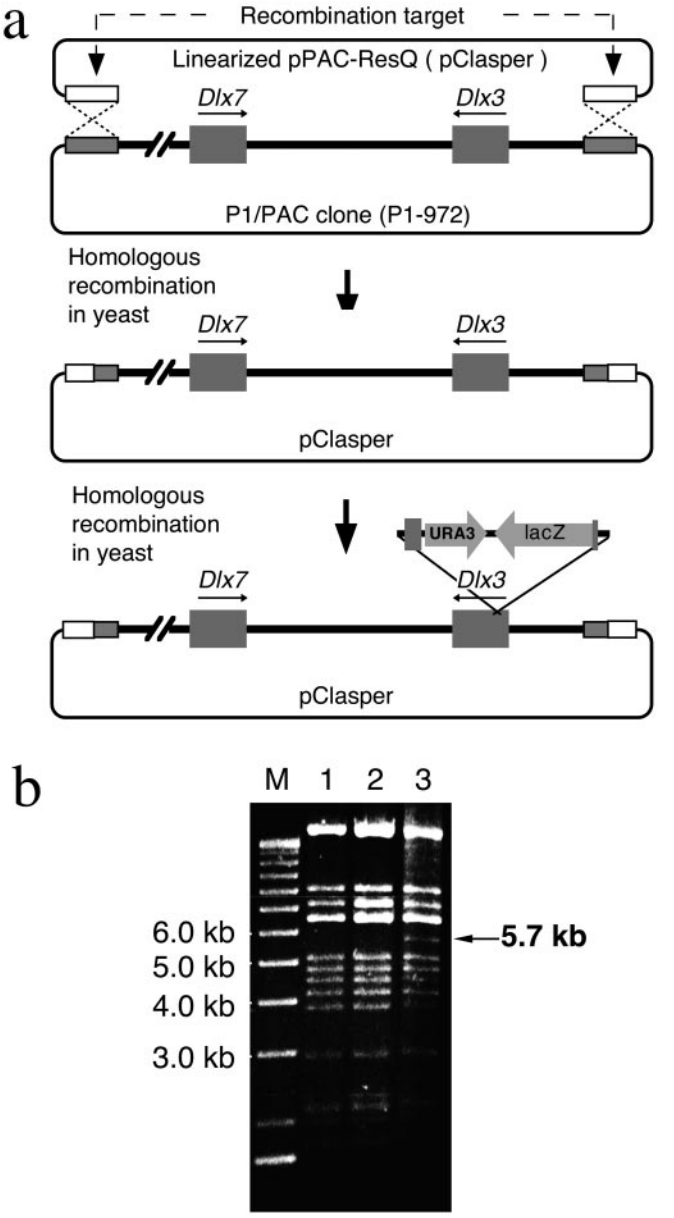

Fig. 4. Procedure diagram to make the *Dlx37-lacZ*-79kb construct. (*a*) Linearized pPAC-ResQ yeast-bacteria shuttle vector captures P1–972 insert by homologous recombination in yeast. Further homologous recombination makes *lacZ-URA3* insertion after eighth amino acid from first methyonine. (*b*) *Eco*RI restriction enzyme digestion patterns of original P1–972 (lane 1), pPAC-ResQ with P1–972 insert (lane 2), and with *lacZ-URA3* insertion (lane 3). A 5.7-kb fragment is a diagnostic fragment for *lacZ-URA3* insertion (lane 3, indicated by arrow).

which is not detected by *in situ* hybridization. One transient embryo and one stable transgenic line show pronounced expression on the dorsal midline and in the cranio-facial region. Interestingly, the stable transformant animals in this line show a neural tube closure defect with 50% penetrance. We have not determined the reason for this anomaly; however, it may result from the over- or misexpression of *Dlx7*, which is not disabled in the construct.

In an effort to discover specific control elements in the *Dlx3-7* bigene cluster, we initiated experiments in which subregions of the *Dlx37-lacZ*-79kb reporter construct are tested for expression properties in transgenic mice. In an initial experiment, we have

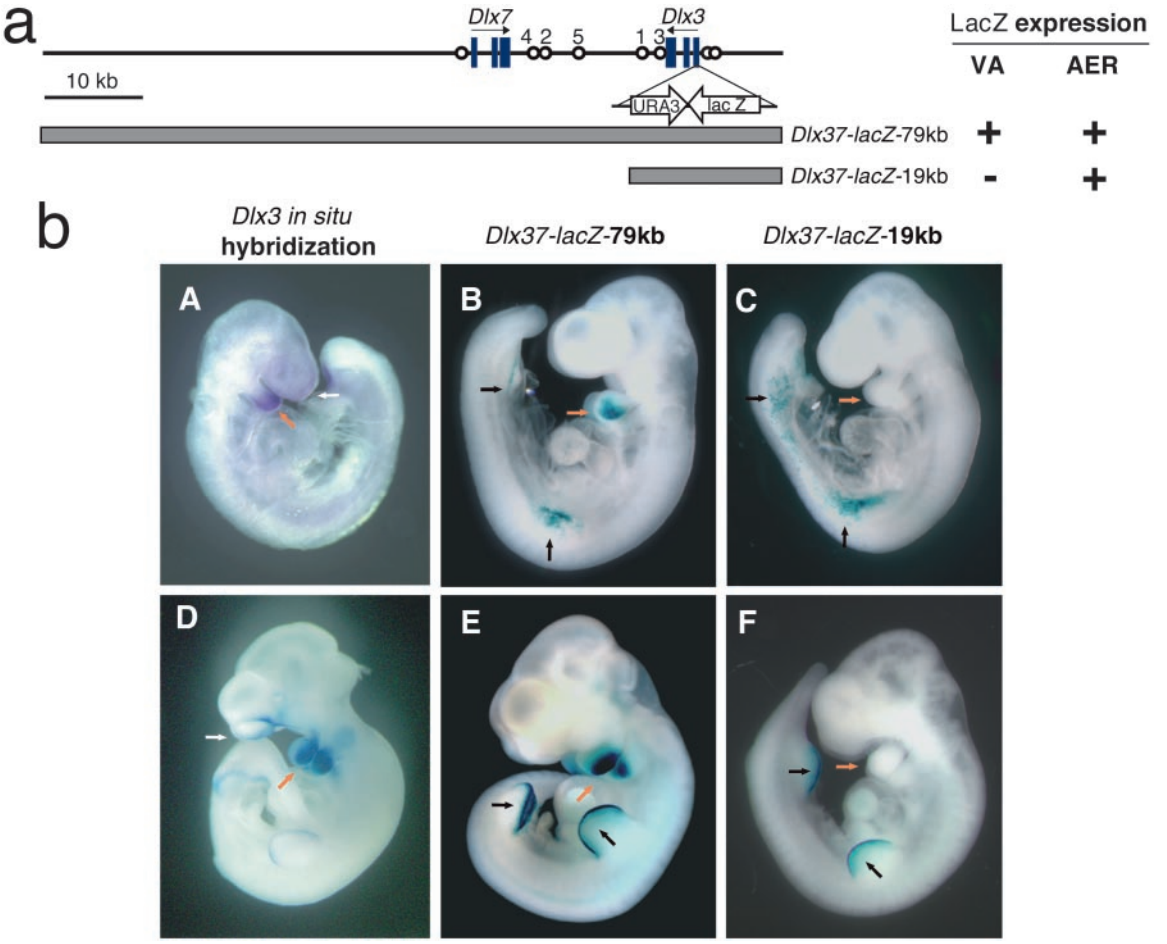

Fig. 5. (*a*) Diagram of transgenes and summary of reporter expression patterns in transgenic animals. Blue boxes and open circles indicate exons and conserved motifs, respectively. *Dlx37-lacZ*-79kb has the full-length of P1–972, whereas *Dlx37-lacZ*-19kb is truncated at the middle of the intergenic region, retaining 7 kb upstream and 4 kb downstream of *Dlx3*. (*b*) *In situ* hybridization of *Dlx3* (*A* and *D*) and β-galactosidase staining showing transgene expression (*B*, *C*, *E*, and *F*). (*A–C*) Stage E9.5. (*D–F*) Stage E10.5. (*B* and *E*) β-galactosidase staining of *Dlx37-lacZ*-79kb transgenic embryos. Note staining in visceral arches (indicated by red arrow) and limbs (black filled arrow) detected in *in situ* hybridization. (*C* and *F*) β -galactosidase staining in *Dlx37-lacZ*-19kb transgenic embryos. Note loss of staining in visceral arches (red arrow). Frontonasal ectoderm expression detected in *in situ* hybridization (white open arrow) is not detected in transgenic animals.

tested a subconstruct, *Dlx37-lacZ*-19kb that lacks *Dlx7* and its flanking elements. This reporter construct contains the *Dlx3*-coding exons with *lacZ* insertion and intergenic conserved elements I37–1 and I37–3 (Fig. 5*a*). We obtained three transient and two stable transgene lines with this construct. Once again, the expression patterns between the independent founders are highly consistent (Fig. 5 *bC* and *bF*). Expression is seen in the AER of both fore and hind limb buds at stages E9.5 and E10.5. Moreover, the pattern is highly similar in this respect to that of the *Dlx37-lacZ*-79kb construct. However, there is a striking difference in that there is no detectable expression of the *Dlx3* reporter in the first and second visceral arches. This result provides strong initial evidence that the conserved elements proximal to *Dlx7* are necessary for visceral arch expression of *Dlx3*.

Discussion

The *Dlx3–7* bigene cluster is one of three *Dlx* clusters in mammals. These clusters share a convergent transcriptional orientation and are linked to the Hox clusters, suggesting that they are all derived from cluster duplications subsequent to an initial tandem duplication event (4).

Nonsynonymous to synonymous rate analysis, between human and mouse protein sequences, shows that *Dlx3* is more evolutionarily conserved than *Dlx7* (Table 1). There are two possible explanations. In the first, *Dlx3* has retained more of its original functions since the two genes duplicated, whereas *Dlx7* has been evolving more free of functional constraints. In the second, *Dlx7* is evolving under directional selection. Some evidence points to important differences in function of the two genes despite partially overlapping patterns of expression. First, *Dlx7* appears to have an important role in hematopoeisis (3), which has not been found for *Dlx3*. Both genes, on the other hand, are expressed in the placenta, but *Dlx7* cannot compensate for a loss-of-function *Dlx3* mutation, which causes embryonic death because of placental failure (8). The divergence in the *Dlx7* protein sequence between human and mouse may relate to as yet unknown changes in protein function.

A comparison between human and mouse *Dlx3–7* bigene clusters reveals striking conservation in noncoding regions, including the 5′ upstream promoter regions of both genes and the 3′ downstream intergenic region common to both. We have focused on the intergenic region because we hypothesize that this region might be the site of critical shared enhancers produced in the tandem duplication that created the bigene cluster (7). There are five putative control elements that show a high level of sequence similarity between human and mouse in the intergenic region (Fig. 3). Moreover, element IG73–1

shows a sequence match to an element in the corresponding zebrafish *dlx3–7* bigene cluster. The zebrafish has additional *Dlx7* gene, termed *dlx8*, in comparison to mammals. One of these, designated *dlx8*, is actually an ortholog of the mammalian *Dlx7* gene. Therefore, it is possible that additional conserved elements between mammals and zebrafish might exist between the mammalian *Dlx3–7* cluster and the zebrafish *dlx8* gene as a consequence of cis-element subfunctionalization between duplicated gene clusters (21).

We have made sequence comparisons, *inter se*, among the three *Dlx* bigene clusters found within mammalian species. We find significant similarity only in the homeodomains but none in other coding or the noncoding regions. We find this anomalous because all of the clusters are derived from a common ancestor and the expression patterns of all genes partially overlap. How might expression patterns be conserved, in part, while control sequences undergo major modification after cluster duplication? We postulate two interrelated processes. In the first, termed "fine-tuning", the transcription factor binding elements that themselves cluster in modules or enhanceosomes undergo loss, gain, or modification randomly, but some of these changes are advantageous, and are selected or drift to fixation, resulting in novel adaptive patterns of expression (22). In the second process, termed "compensation", sequence changes occur randomly and are not immediately eliminated by purifying selection because of redundancy or buffering in the overall enhancer module. Constant expression patterns are maintained by balancing selection for subsequent compensatory changes, as proposed by Kreitman and coworkers (23–26). We hypothesize that both these processes are occurring in the *Dlx* bigene clusters. Thus, sequence modifications are constantly introduced, erasing the similarity in the sequences as originally inherited from the common ancestral cluster. Fine tuning in control elements results in functionally important changes in individual gene or cluster expression patterns while compensation maintains the overall similarities in the overlapping patterns of expression between the genes.

We report here initial transgenic experiments to directly ascertain the function of noncoding conserved elements. A reporter construct consisting of a 79-kb genomic fragment with a *LacZ* insertion in exon 1 of the *Dlx3* gene is capable of reconstituting expression in tail bud, limbs, and visceral arches in a manner closely resembling endogenous *Dlx3* expression in stage E9.5 and E10.5 embryos (Fig. 5). This finding supports the view that most of the cis-elements required for proper *Dlx3* expression are contained in the construct. In the visceral arches, the *Dlx37-lacZ*-79kb construct can initiate mesenchymal expression in the distal tip of the arches at E9.5 where postmigratory neural crest cells are present. At stage E10.5, reporter expression is restricted to the caudal portion of the mandibular process in conformity with endogenous expression (9). In the fore and hind limb buds, expression is detected at E9.5. At this developmental stage, the reporter is expressed in a broad area of lateral epithelium and weakly in the AER. Expression is restricted to the AER later at E10.5. Reporter expression is absent in the fronto-nasal process although *Dlx3* transcripts are detected by *in situ* hybridization. This may be explained by the absence of certain control elements in our construct. In particular, the *Dlx3* 5′-flanking portion of *Dlx37-lacZ*-79kb is relatively short, being only 7 kb in length.

The *Dlx* bigene clusters are notable in that the two genes have overlapping patterns of expression ≈E10.5. Ellies *et al.* (6) proposed that the similarities in expression pattern within a cluster are due to enhancer sharing between the linked genes, and they put forward three hypothetical models of the original tandem gene duplication to explain similarity in expression patterns. In model A, a single control element upstream of one gene can regulate both genes similarly. In model B, a single control element located in the intergenic region regulates both genes similarly. In model C, separate elements located upstream of each gene regulate both similarly. Models A and B are consistent with enhancer sharing between *Dlx7* and *Dlx3*. Our transgene results with the construct *Dlx37-LacZ*-19kb rule out model C because visceral arch expression of *Dlx3* is lost when cis-elements proximal to *Dlx7* are deleted. This result argues in favor of enhancer sharing. We will reexamine this question by transgenic experiments in which the putative enhancer elements are deleted singly and in combination.

We thank Marc Ekker for sharing zebrafish sequence data and Günter Wagner for insightful comments on the manuscript. This research was supported by National Institutes of Health Grant GM 09966 and National Science Foundation Grant IBN-9905403 (to F.H.R.); National Science Foundation/Sloan Postdoctoral Fellowship in Molecular Evolution (to S.Q.I.); and Japan Society for the Promotion of Science Postdoctoral Fellowship for Research Abroad (to K.S.).

1. Simeone, A., Acampora, D., Pannese, M., D'Esposito, M., Stornaiuolo, A., Gulisano, M., Mallamaci, A., Kastury, K., Druck, T., Huebner, K., *et al.* (1994) *Proc. Natl. Acad. Sci. USA* **91,** 2250–2254.
2. McGuinness, T., Porteus, M., Smiga, S., Bulfone, A., Kingsley, C., Qiu, M., Liu, J., Long, J., Xu, D. & Rubenstein, J. (1996) *Genomics* **35,** 473–485.
3. Nakamura, S., Stock, D. W., Wydner, K. L., Bollekens, J. A., Takeshita, K., Nagai, B. M., Chiba, S., Kitamura, T., Freeland, T. M., Zhao, Z., *et al.* (1996) *Genomics* **38,** 314–324.
4. Stock, D. W., Ellies, D. L., Zhao, Z., Ekker, M., Ruddle, F. H. & Weiss, K. M. (1996) *Proc. Natl. Acad. Sci. USA* **93,** 10858–10863.
5. Rossi, E., Faiella, A., Zeviani, M., Labeit, S., Floridia, G., Brunelli, S., Cammarata, M., Boncinelli, E. & Zuffardi, O. (1994) *Genomics* **24,** 34–40.
6. Quint, E., Zerucha, T. & Ekker, M. (2000) *J. Exp. Zool.* **288,** 235–241.
7. Ellies, D. L., Stock, D. W., Hatch, G., Giroux, G., Weiss, K. M. & Ekker, M. (1997) *Genomics* **45,** 580–590.
8. Morasso, M. I., Grinberg, A., Robinson, G., Sargent, T. D. & Mahon, K. A. (1999) *Proc. Natl. Acad. Sci. USA* **96,** 162–167.
9. Robinson, G. W. & Mahon, K. A. (1994) *Mech. Dev.* **48,** 199–215.
10. Beanan, M. & Sargent, T. (2000) *Dev. Dyn.* **218,** 545–553.
11. Zhao, Z., Stock, D. W., Buchanan, A. V. & Weiss, K. M. (2000) *Dev. Genes Evol.* **210,** 270–275.
12. Weiss, K. M., Ruddle, F. H. & Bollekens, J. (1995) *Connect. Tissue Res.* **32,** 35–40.
13. Kawasaki, K., Minoshima, S., Nakato, E., Shibuya, K., Shintani, A., Asakawa, S., Sasaki, T., Klobeck, H.-G., Combriato, G., Zachau, H. G., *et al.* (2001) *Eur. J. Immunol.* **31,** 1017–1028.
14. Schwartz, S., Zhang, Z., Frazer, K. A., Smit, A., Riemer, C., Bouck, J., Gibbs, R., Hardison, R. & Miller, W. (2000) *Genome Res.* **10,** 577–586.
15. Tatusova, T. A. & Madden, T. L. (1999) *FEMS Microbiol. Lett.* **174,** 247–250.
16. Bhargava, J., Shashikant, C., Carr, J., Bentley, K., Amemiya, C. & Ruddle, F. (1999) *Genomics* **56,** 337–339.
17. Bradshaw, M. S., Bollekens, J. A. & Ruddle, F. H. (1995) *Nucleic Acids Res.* **23,** 4850–4856.
18. Bradshaw, M. S., Shashikant, C. S., Belting, H.-G., Bollekens, J. A. & Ruddle, F. H. (1996) *Proc. Natl. Acad. Sci. USA* **93,** 2426–2430.
19. Shashikant, C. S., Bieberich, C. J., Belting, H. G., Wang, J. C., Borbely, M. A. & Ruddle, F. H. (1995) *Development (Cambridge, U.K.)* **121,** 4339–4347.
20. Nei, M. & Gojobori, T. (1986) *Mol. Biol. Evol.* **3,** 418–426.
21. Force, A., Lynch, M., Pickett, F. B., Amores, A., Yan, Y. & Postlethwait, J. (1999) *Genetics* **151,** 1531–1545.
22. Tautz, D. (2000) *Curr. Opin. Genet. Dev.* **10,** 575–579.
23. Kreitman, M. & Ludwig, M. (1996) *Semin. Cell Dev. Biol.* **7,** 583–592.
24. Ludwig, M. Z. & Kreitman, M. (1995) *Mol. Biol. Evol.* **12,** 1002–1011.
25. Ludwig, M. Z., Patel, N. H. & Kreitman, M. (1998) *Development (Cambridge, U.K.)* **125,** 949–958.
26. Ludwig, M. Z., Bergman, C., Patel, N. H. & Kreitman, M. (2000) *Nature (London)* **403,** 564–567.