

Conservation of sequence and structure flanking the mouse and human β -globin loci: The β -globin genes are embedded within an array of odorant receptor genes

MICHAEL BULGER*, J. HIKKE VON DOORNINCK†, NORIKO SAITOH†, AGNES TELLING*, CATHERINE FARRELL‡, M. A. BENDER*, GARY FELSENFELD‡, RICHARD AXEL†, AND MARK GROUDINE*§¶

*Division of Basic Sciences, Fred Hutchinson Cancer Research Center, 1100 Fairview Avenue North, Seattle, WA 98109; †Center for Neurobiology and Behavior, Howard Hughes Medical Institute, College of Physicians and Surgeons, Columbia University, New York, NY 10032; ‡Laboratory of Molecular Biology, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, MD 20892; and §Department of Radiation Oncology, University of Washington Medical School, Seattle, WA 98195

Contributed by Gary Felsenfeld, February 24, 1999

ABSTRACT In mouse and human, the β -globin genes reside in a linear array that is associated with a positive regulatory element located 5' to the genes known as the locus control region (LCR). The sequences of the mouse and human β -globin LCRs are homologous, indicating conservation of an essential function in β -globin gene regulation. We have sequenced regions flanking the β -globin locus in both mouse and human and found that homology associated with the LCR is more extensive than previously known, making up a conserved block of ≈ 40 kb. In addition, we have identified DNaseI-hypersensitive sites within the newly sequenced regions in both mouse and human, and these structural features also are conserved. Finally, we have found that both mouse and human β -globin loci are embedded within an array of odorant receptor genes that are expressed in olfactory epithelium, and we also identify an olfactory receptor gene located 3' of the β -globin locus in chicken. The data demonstrate an evolutionarily conserved genomic organization for the β -globin locus and suggest a possible role for the β -globin LCR in control of expression of these odorant receptor genes and/or the presence of mechanisms to separate regulatory signals in different tissues.

In mammals, the multiple β -globin genes reside in a linear array that reflects their developmental order of expression. Examination of naturally occurring deletions of the human β -globin locus, and experiments in transgenic mice, have defined a locus control region (LCR) that acts as a powerful erythroid-specific positive regulatory element (1–3). In humans, the LCR has been defined as a 20- to 25-kb region beginning ≈ 10 kb upstream of the embryonic β -globin gene (≈ 50 kb upstream of the adult β -globin gene). The LCR includes five major DNaseI-hypersensitive sites (HSs) that map to regions that contain binding sites for both general and erythroid-specific transcription factors. The binding of such factors is believed to alter chromatin structure to generate the local HSs within the LCR and also to propagate more long-range structural changes including the formation of an extensive nuclease-sensitive region along the entire locus. The HS sequences are highly conserved among mammalian species from human to mouse, suggesting that they perform an essential function in regulating gene expression from the β -globin locus (1).

The naturally occurring Hispanic deletion, which removes HS2–5 of the human LCR and 25 kb of DNA further 5', results in a complete loss of β -globin gene expression and the abrogation of the nuclease-sensitive structure in erythroid cells

(4). Recently, targeted deletions of HS1–6 of the mouse β -globin LCR and HS2–5 of the human LCR were accomplished (5, 6) (see Fig. 1). These deletions were associated with complete loss of β -globin gene expression from the human locus and a significant reduction of transcription from the mouse locus. Nuclease sensitivity within both β -globin loci was maintained, however, which suggests that important regulatory sequences, in addition to the LCR, contribute to this function.

Regions possessing functions analogous to those of the β -globin LCR have been identified within a number of gene loci, including the chicken lysozyme gene (7), the human adenosine deaminase gene (8), the human Ig heavy chain locus (9), the T cell receptor α/δ locus (10), and many others. Within multiple-gene arrays, such as the β -globin locus, the TCR α/δ locus, and the red-green opsin locus, it has been suggested that LCRs may also function to ensure that only one gene is expressed at any given time (10–12). However, it has been shown that the developmental pattern of expression of specific β -globin genes is maintained in the absence of the LCR (5, 13). Thus, transcription of a single gene from within a linked array may represent a decision made independent of, but amplified by, the LCR.

In this study, we have sequenced 30.6 kb of genomic DNA 5', and 54.5 kb of DNA 3', to the previously published β -globin sequences in mouse, along with 36 kb of DNA 5' to the published sequences within the human β -globin locus (Fig. 1). The new sequence reveals that the regions of conservation associated with the mouse and human β -globin LCRs extend considerably farther than previously recognized. We have also identified conserved DNaseI-hypersensitive sites in both mouse and human. The data suggest that the β -globin LCR may be larger and more complex than has been thought.

In addition, we have found that the β -globin locus is surrounded by genes encoding olfactory receptors expressed in mouse olfactory epithelium. In mice, the family of odorant receptor genes consists of about 1,000 members distributed in large linked arrays present on each of the autosomes (14–16). Despite this large number, each individual olfactory neuron expresses only a single receptor gene (17). Little is known of how this pattern of gene regulation is accomplished. One model suggests that an LCR could be responsible for the activation of a specific array and may also govern the expression of one gene from within the cluster. However, LCRs have not yet been identified within the multiple loci encoding

Abbreviations: LCR, locus control region; HS, hypersensitive sites; PIP, percent identity plot; ORG, olfactory receptor genes.

Data deposition: The sequences reported in this paper have been deposited in the GenBank database (accession nos. AF133300, AF128269, AF137131, and AF137396).

¶To whom reprint requests should be addressed. e-mail: markg@fhcrc.org.

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.

PNAS is available online at www.pnas.org.

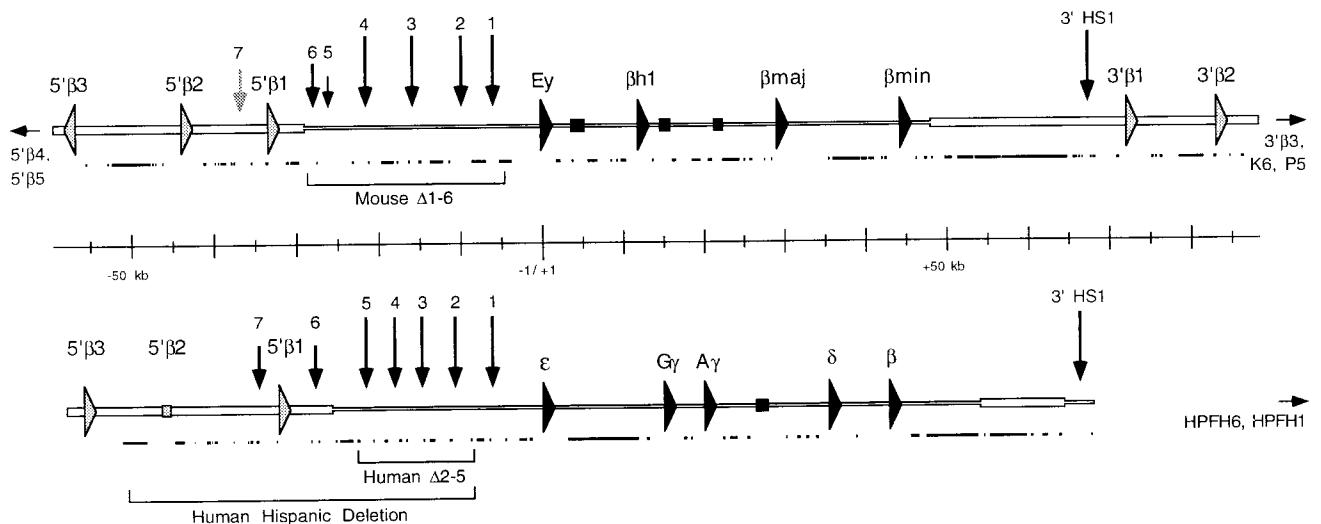


FIG. 1. Mouse (*Upper*) and human (*Lower*) β -globin loci. The cap sites of the human ϵ -globin gene and of the mouse *Ey* gene are set as +1. Sequences reported here are represented by a wider horizontal bar. Globin genes are represented by filled arrows, olfactory receptor ORFs by shaded arrows, and pseudogenes by filled or shaded boxes. HSs are shown as vertical arrows. The lengths of the arrows illustrate the relative intensities of the HS that they represent; the arrow indicating mouse 5' HS7 is lighter to indicate that this site is still hypothetical. The lower, filled lines for both mouse and human loci designate repetitive elements. Horizontal arrows pointing outward from these sequences indicate the presence of additional olfactory receptor ORFs located beyond the regions shown here. Brackets beneath the sequences show the extent of deletions within the β -globin locus that are discussed in the text.

olfactory receptor genes. Our results raise the possibility that the β -globin LCR may also play a role in regulating transcription of some odorant receptor genes, and/or that elements in or near the LCR may insulate regulatory elements of the two gene families.

MATERIALS AND METHODS

Clones containing β -globin-flanking sequences were obtained from several sources. For mouse 5' and 3' sequences, clones used included a 15-kb insert from a 129Sv mouse library derived from AK-7 embryonic stem (ES) cells (18) and inserts derived from two P1 clones isolated from a 129Sv mouse library (Genome Systems). Human sequences 5' of the LCR were derived from a λ phage library (19) and from a P1 clone (Genome Systems, St. Louis). Smaller inserts were subcloned into pBluescript KS II (+) or SK II (+) vectors by standard methods. Sequencing was done by using dye terminators and universal or custom-synthesized primers on an ABI 377XL automated sequencer. Both strands of all regions were sequenced at least once. Repetitive elements were identified by using the RepeatMasker2 web server (<http://ftp.genome.washington.edu/cgi-bin/RepeatMasker>).

The generation of percent identity plots (PIPs) was as described (20). Phylogenetic analysis of olfactory receptor sequences was performed by using the programs CLUSTALX and NJTREE-UNROOTED. A full description of the receptor genes used in generating the tree is available on request.

DNaseI-HS mapping was performed as described (4). The probes used were restriction enzyme fragments as described in the figure legend. *In situ* hybridization was performed as described (21) using full-length odorant receptor clones as templates to synthesize digoxigenin-labeled cRNA probes.

RESULTS AND DISCUSSION

Extending the β -Globin LCR: Additional 5' HSs. The previously published sequences of the β -globin loci of mouse and human reveal that whereas the globin gene coding regions are conserved, the intergenic regions are divergent, with limited exceptions (data not shown). In contrast, the LCR, located 5' of the globin genes, is composed of regions of high

sequence similarity between mouse and human along nearly all of the nonrepetitive DNA sequences, including the five conserved DNaseI HSs (1). Our new sequencing allows us to determine how far this conservation actually extends. By using PIP analysis, significant sequence homology is evident for ≈ 20 kb upstream of the previously published sequences (Fig. 2A). Of interest, the homology ends near the 5' breakpoint for the naturally occurring Hispanic deletion within the human β -globin locus. No additional conservation is seen along the >10 kb of available sequence farther 5' of this point in human and mouse.

Recently, we sequenced and characterized the most 5' DNaseI-HSs of the mouse LCR (18). In that study, we identified the mouse sequence homologous to human 5' HS5, but this region was associated with a very weak HS. A more intense HS was observed farther 5', corresponding to a sequence with no known homologue in human, and was termed 5' HS6. The newly determined human sequence includes a region 5' of HS5 with homology to mouse HS6 (Fig. 2A). To determine whether a homologue of mouse 5' HS6 exists in human, we have mapped HS in this region by DNaseI digestion of nuclei from human fetal liver (Fig. 3A). Our analysis reveals two HSs located approximately 6 and 12 kb 5' of human HS5. The site closest to 5' HS5 maps to the human sequence that is homologous to mouse 5' HS6, and thus we term this site human 5' HS6. The other site, which we term 5' HS7, maps to another region in human that shows homology to mouse (Figs. 2A and 3A). Neither site is evident in a DNaseI digestion series of nuclei from human fetal brain (data not shown), indicating that these sites are not indicators of a chromatin structure present in all cell types.

The observed degree of sequence and structural conservation suggests that erythroid regulatory sequences 5' of the β -globin locus may be more extensive than previously thought and may effectively double the known size of the LCR. This could provide an explanation for recent experiments involving deletions of the mouse and human LCRs. A targeted deletion limited to 5' HS2–5 of the human LCR results in a drastic loss of globin gene expression, but does not abrogate generalized nuclease sensitivity (6). Moreover, the remaining 5' HS1 is retained in a hypersensitive state in erythroid cells. In contrast, the naturally occurring Hispanic deletion within the human

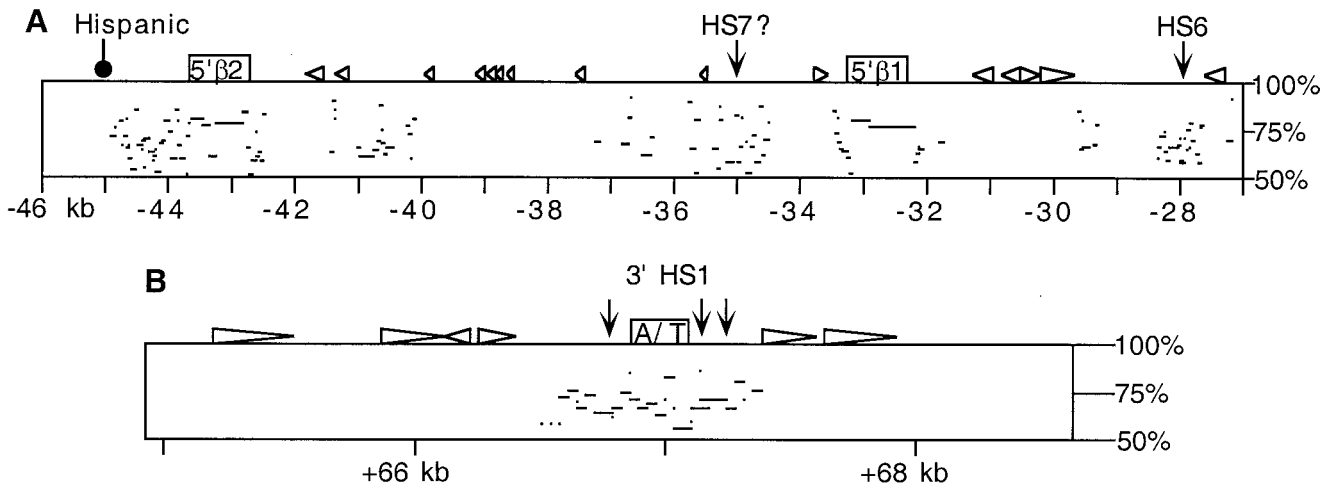


FIG. 2. PIP analyses of mouse and human sequences. (A) PIP of the region 5' of the β -globin LCR. The mouse sequence is represented at the top, and the percent identity of gap-free aligning segments from human is plotted beneath. The odorant receptor genes MOR5' β 1 and MOR5' β 2 are indicated as boxes, and repetitive elements shown as triangles. The positions of HSs and the 5' breakpoint of the human Hispanic deletion are indicated. Distance from the hypothetical Ey cap site in kb is shown on the bottom. For reference, the homologous human sequences that give rise to the alignments shown here correspond to HOR5' β 1 and HOR5' β 2 and surrounding regions as shown in Fig. 1. (B) PIP of the region of homology 3' of the β -globin genes. The mouse sequence is represented at the top, with comparison made to the human 3' HS1 sequence (24). The A/T-rich region is indicated by the shaded box.

β -globin locus, which involves an excision of 5' HS2–5 and an additional 20 kb upstream of the LCR including the newly identified HS 6 and 7, results in a complete loss of β -globin gene expression, a loss of generalized nuclease sensitivity, the absence of hypersensitivity at the remaining 5' HS1, and a transition from early to late replication in of the β -globin locus

in erythroid cells (4). The different phenotypes of the two deletions suggest that the additional conserved sequences included in the Hispanic deletion might contribute to the regulation of the β -globin locus.

The newly determined human sequences also reveal a lack of conservation in comparison to the rabbit β -globin LCR (22).

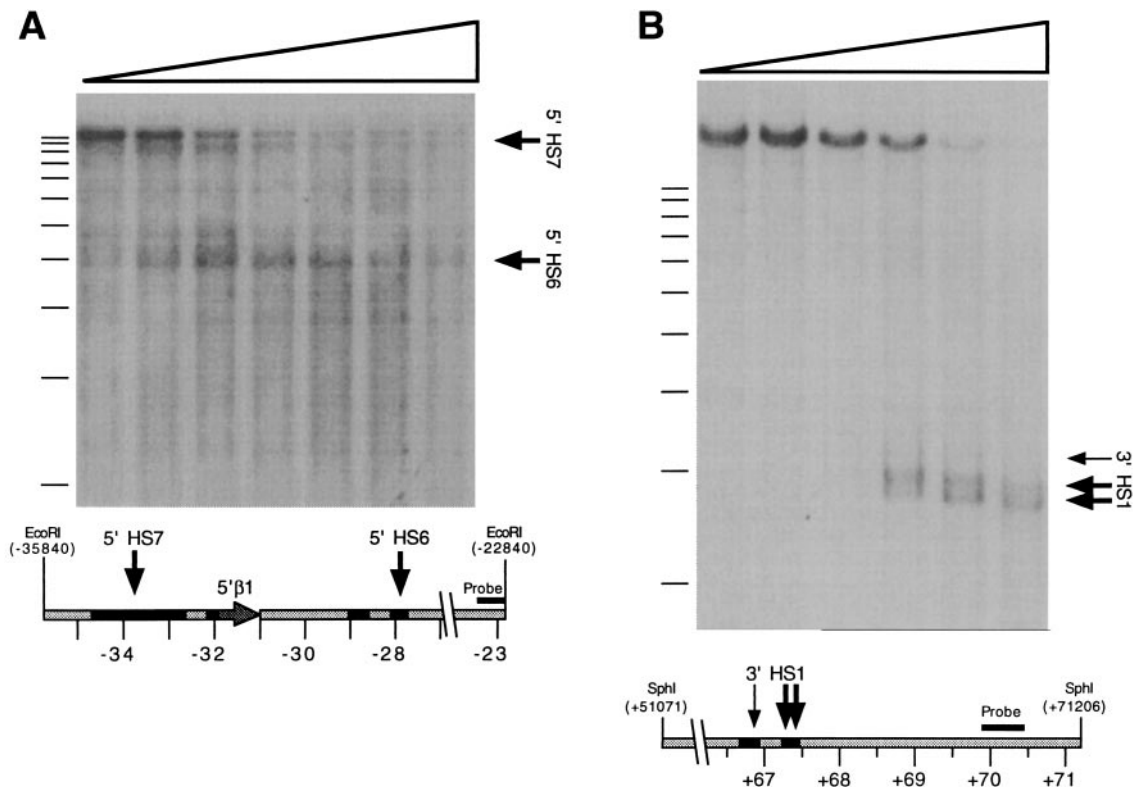


FIG. 3. DNaseI-HS mapping within the newly sequenced β -globin-flanking regions. (A) Mapping of human 5' HS6 and 5' HS7. A Southern blot of an *Eco*RI-digested DNase series prepared from human fetal liver was hybridized with a probe corresponding to an *Eco*RI/*Eco*0109I restriction fragment located at -23.4 to -22.8 kb. Arrows indicate DNaseI subbands evident on the autoradiogram and their corresponding positions on the human sequence represented below it. Filled regions on the diagram indicate significant ($>50\%$ identity) homology to mouse sequence, as shown in Fig. 2. (B) DNaseI-HSs located 3' of mouse β -globin genes. A Southern blot of an *Sph*I-digested DNase series prepared from phenylhydrazine-treated mouse spleen was hybridized with a probe corresponding to an *Acc*I/*Nsi*I restriction fragment located from $+69.9$ to $+70.5$ kb.

In rabbit, the regions that would correspond to 5' HS5 and 5' HS6 are absent, although extensive homology with regions both 5' and 3' of human HS5 and HS6 is maintained (data not shown). The deletion of this region in rabbit could indicate that 5' HS5 and 5' HS6 are nonessential for β -globin gene expression, which would be consistent with the observation that a deletion of this region in mouse has little measurable phenotype (18). It is possible, however, that rabbits have acquired an additional alteration within or near the locus that compensates for the absence of HS5 and HS6.

HSs 3' of the β -Globin Locus. The extensive sequence and structural conservation between human and mouse 5' of the previously defined LCRs prompted us to examine the possibility of additional conservation 3' of the β -globin locus as well. Previous studies have identified a nuclease-HS located ≈ 20 kb 3' of the adult β -globin gene in humans, termed 3' HS1 (23). Analysis of the mouse sequence 3' of the β -globin locus reveals a region located 21 kb downstream of β -minor that is homologous to human 3' HS1 (Fig. 2B), consisting of two conserved sequence blocks separated by a shared AT-rich region. Fine structure mapping of the human 3' HS1 previously demonstrated three distinct HSs in the region (24), all of which map to the sequences conserved between mouse and human.

We therefore asked whether these sites are also retained in mouse erythroid chromatin. DNaseI digestion of nuclei from mouse erythroid tissue reveals three HSs corresponding to the regions of homology with human (Fig. 3B). Although the 5'-most site is weak in these assays, the data indicate a further conservation of sequence and chromatin structure downstream of the β -globin gene. The function of these HSs is unknown. The region revealed no enhancer activity in transient transfection assays (24), although its potential activity in other assays has not been fully examined. Still, the conservation of these sites between mouse and human suggests that they may have a common regulatory role.

A Functional Olfactory Receptor Gene Cluster Surrounding the β -Globin Locus. Sequencing of the regions 5' and 3' of the β -globin locus in mouse and 5' of the β -globin locus in human reveals the presence of several ORFs corresponding to olfactory receptor genes (ORGs) (Fig. 1). We find five ORGs 5' of the β -globin LCR in mouse and three in human. We also find five ORGs 3' of the β -globin genes in mouse, whereas two are known to exist in human from analysis of the breakpoints of naturally occurring deletions known as HPFH-1 and HPFH-6 (25). Among these ORFs, the human $\text{HORS}'\beta 2$ and HPFH-6 sequences are pseudogenes on the basis of several missense and nonsense mutations that prevent the translation of a functional receptor molecule. All of the other genes we have identified in mouse and human encode full-length receptors, indicating that the mammalian β -globin loci are embedded within a cluster of genes encoding olfactory receptors.

Olfactory receptors are characterized as seven-transmembrane G protein-coupled receptor molecules expressed in sensory neurons of nasal epithelium (14), although expression in other tissues has been documented. Odorant receptors are encoded by a large family of genes that in the mouse consists of $\approx 1,000$ members (14, 15) that usually reside in large linked arrays containing up to 30 genes (16). In humans, these arrays have been found on nearly all of the autosomes (26). Olfactory receptor clusters have been mapped to both human chromosome 11 and mouse chromosome 7, on which the β -globin genes reside (26–28). None of the genes we have found in proximity to the β -globin locus, however, corresponds to the receptors identified in these studies.

Sequence analysis of the β -globin-proximal ORGs reveals that they share greater homology to each other than to any other known full-length receptor genes, with the exception of an ORG from rat termed RA1c (29). We have performed a phylogenetic analysis of the β -globin-proximal ORGs from mouse and human along with 82 published full-length ORG

sequences (Fig. 4). The resulting phylogenetic tree indicates that the newly identified receptor genes, along with those found near the HPFH breakpoints and the rat RA1c receptor, define a highly divergent group of olfactory receptor families with a closer relationship to receptors from fish than to other mammalian or avian receptors. It will be interesting to determine whether the RA1c receptor gene is located near the rat β -globin locus.

The presence of orthologous clusters of odorant receptor genes proximal to the β -globin locus in both mouse and human raises the question of whether this arrangement is conserved among other vertebrates. We find an odorant receptor gene located 7.1 kb downstream of the 3' end of the chicken ϵ -globin gene, which is the most downstream gene in the chicken β -globin locus (data not shown). This receptor, termed COR3 β , belongs to the same group of olfactory receptors composing the mammalian β -globin-proximal ORGs (Fig. 4). The presence of this gene adjacent to the chicken β -globin locus indicates that the genomic organization of this olfactory receptor cluster and the β -globin genes precedes the divergence of both birds and mammals from reptiles.

In olfactory sensory neurons, olfactory receptors are localized to a specialization of the dendrite where they interact with odors in the environment and transduce odor binding into alterations in membrane potential (30). Within the olfactory epithelium, four spatially defined domains have been identified in a dorsal/ventral dimension, such that a given odorant receptor is expressed in only one of four domains (31–32). Within a given domain, however, neurons expressing a given receptor appear to be randomly dispersed. Moreover, a given olfactory sensory neuron expresses only one odorant receptor gene (17). A small subset of ORGs are expressed in tissues

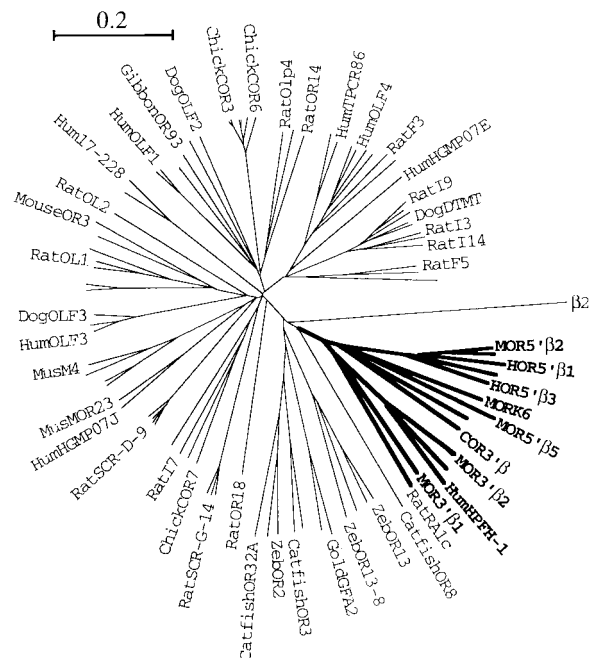


Fig. 4. Phylogenetic tree of ORGs. An unrooted phylogenetic tree was generated by using 82 polypeptide sequences corresponding to ORGs from human, mouse, rat, dog, gibbon, chicken, catfish, zebrafish, and goldfish, along with 14 of the ORGs found proximal to the β -globin loci of human, mouse, and chicken. The genes known to be proximal to the β -globin loci are indicated by bold type. The tree was confirmed by 1,000 rounds of "bootstrapping." Only intact, full-length sequences were used in the generation of the tree. Not all of the genes are labeled (see *Materials and Methods*). $\beta 2$ represents the β -2-adrenergic receptor and is shown here to indicate the likely position of the root (outgroup). The bar corresponds to the graphical distance equivalent to 20% divergence.

other than the nasal epithelium, including testis (15, 33), fetal heart (34), and notochord (35). The observation that the β -globin cluster is flanked by ORGs prompted us to determine whether these genes are expressed in olfactory neurons or erythroid cells.

We have examined the expression of newly identified mouse olfactory receptors in olfactory epithelium and fetal liver by *in situ* hybridization with RNA antisense probes. Probes corresponding to the 5' mouse genes, MOR5' β 1 and MOR5' β 2, as well as the 3' genes, MOR3' β 1 and MOR3' β 2, reveal a pattern of punctate staining of a small subpopulation of neurons within the olfactory epithelium (Fig. 5). All four probes hybridize to cells within the most dorsal domain of the nasal epithelium, corresponding to zone 1 (31). Within zone 1, each of the probes identifies 1–2% of the olfactory neurons. The frequency and punctate pattern of hybridization is identical to that observed with OR probes from genes at other chromosomal loci, and control hybridization with sense probes reveals no specific staining to olfactory epithelium. The MOR3' β 1 probe hybridized to a single band on Southern blot analysis, whereas the MOR5' β 1 probe hybridized to two bands (the 5' β 1 and 5' β 2 genes) under conditions of high stringency; because this stringency is lower than that of the *in situ* hybridization, the hybridization we observe results from expression of odorant receptor genes flanking the β -globin cluster, and not from elsewhere in the genome.

Experiments to determine whether the odorant receptor genes flanking the β -globin locus are also expressed in erythroid cells reveal no hybridization to 12.5-day postcoitum fetal liver (data not shown). This suggests that whereas the β -globin genes are expressed in erythroid cells, the neighboring ORGs are restricted in their expression to the olfactory epithelium. We cannot yet rule out expression of these odorant receptor genes in other erythroid developmental compartments, however.

The pattern of olfactory receptor expression, such that individual neurons express only 1 of 1,000 receptor genes, poses an interesting problem in transcriptional regulation of these genes. At present, however, little data exists that might serve to define the mechanism by which this regulation is accomplished. One of us (R.A.) has suggested a model in which a hierarchy of controls is exerted on the family of odorant receptor genes to activate expression of a single gene in each neuron. LCRs present in each cluster of odorant receptor genes may represent one aspect of such hierarchical control. In this context, the expression of a single receptor in each olfactory neuron could result from an LCR-mediated selection of an individual gene in a cluster, subsequent to activation of that cluster.

Another possibility is suggested by the hypothesis that enhancers and LCRs act by mechanisms similar to those used by silencers (36). For example, normal silencing of the MAT locus in yeast depends on a silencer element, which serves as a nucleation site for a repressive chromatin structure that then spreads for considerable distances. In a similar fashion, an LCR may serve as a nucleation site for active chromatin structure, including the formation of a nuclease-sensitive structure across a broad domain and also a more subtle alteration leading indirectly to gene activation. As applied to the problem of olfactory receptor gene expression, an LCR might serve to potentiate a cluster of genes, whereas selection of the gene to be expressed would be accomplished by another mechanism, perhaps involving promoter-proximal sequences.

The observation that the β -globin LCR resides adjacent to clusters of ORGs, and that this arrangement is evolutionarily conserved, raises questions about possible control of olfactory receptor gene expression by this region. Such a dual regulatory role for a mammalian LCR has not been documented previously, and at present there is no evidence for activity of this LCR in nonerythroid cells. It has been shown, however, that

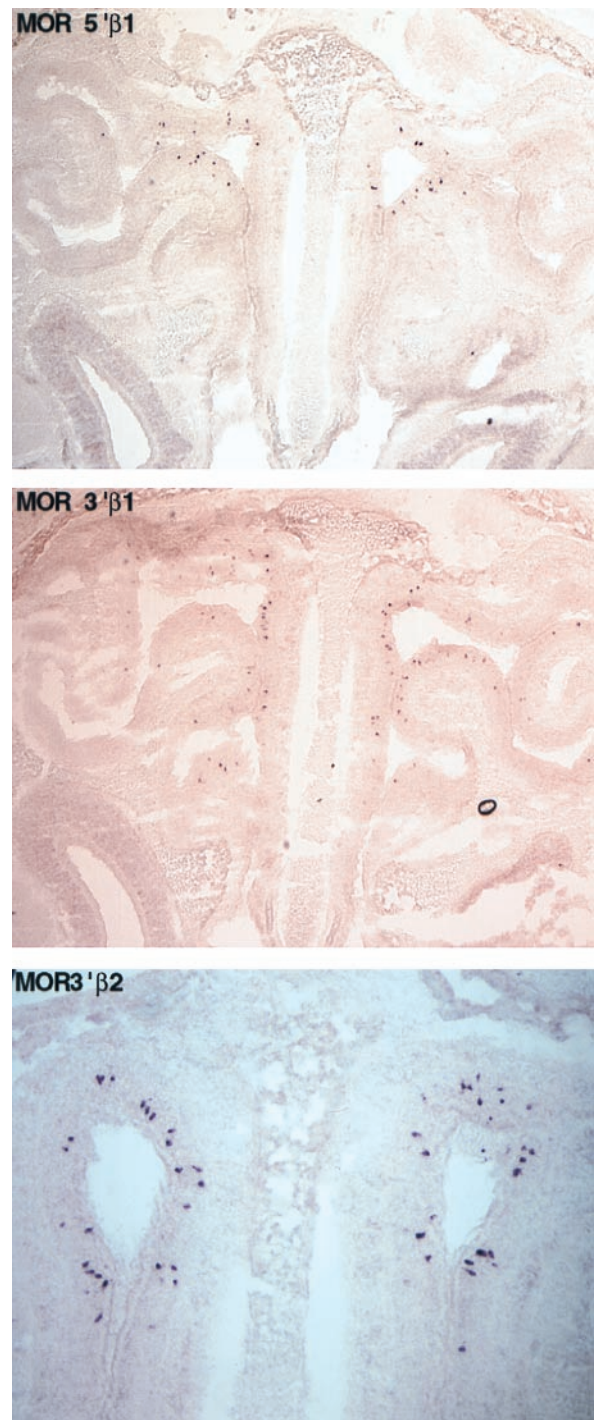


FIG. 5. *In situ* hybridization of mouse olfactory epithelium by using β -globin-proximal odorant receptor genes as probes. Coronal sections through the mouse nasal cavity were hybridized with digoxigenin-labeled antisense RNA probes prepared from the odorant receptor genes indicated (also see Fig. 1). The sections shown are oriented with the dorsal region toward the top and ventral toward the bottom; the nasal septum is visible down the center of each panel.

a portion of the TCR α/δ LCR is capable of driving reporter gene expression in a wide variety of tissues in transgenic mice (37). This observation is interesting because of the presence of the Dad-1 gene, which exhibits a broad pattern of expression in mice, near the endogenous TCR α/δ LCR. In addition, a pattern of HS distinct from that seen in T lymphocytes is observed over the TCR α/δ LCR region in several other mouse tissues (38). These results suggest that regulatory elements for

two sets of genes with different expression patterns are shared or overlap.

Whether or not the β -globin LCR is active for both sets of genes, another possibility raised by the juxtaposition of the β -globin locus with specific olfactory receptor gene arrays is that β -globin regulatory signals might be functionally separated from ORGs (in erythroid cells) and olfactory receptor gene regulatory signals from the β -globin genes (in sensory neurons). Such a separation might be accomplished by enhancer-blocking elements, such as those found at the 5' end of the chicken β -globin locus (39), and in the human T cell receptor locus (40). A direct test of dual function or of enhancer-blocking activity within the β -globin LCR will await the determination of expression of odorant receptor genes in mutant mice containing deletions within the β -globin LCR.

We thank R. Hardison and W. Miller for performing PIP analysis, E. Vanin, T. Townes and T. Ley for providing phage and P1 clones, and members of the Groudine lab for critical reading of the manuscript. This work was supported by National Institutes of Health Grant Nos. DK52854 and DK44746 (M.G.). M.B. is supported by the Helen Hay Whitney Foundation, J.H.v.D. by European Molecular Biology Organization Fellowship, and N.S. by the Japan Society for the Promotion of Science. M.A.B. is a Howard Hughes Medical Institute Physician Postdoctoral Fellow.

- Hardison, R., Slightom, J. L., Gumucio, D. L., Goodman, M., Stojanovic, N. & Miller, W. (1997) *Gene* **205**, 73–94.
- Grosveld, F., Dillon, N. & Higgs, D. (1993) *Baillieres Clin. Haematol.* **6**, 31–55.
- Martin, D. I. K., Fiering, S. & Groudine, M. (1996) *Curr. Opin. Genet. Dev.* **6**, 488–495.
- Forrester, W. C., Epner, E., Driscoll, M. C., Enver, T., Brice, M., Papayannopoulou, T. & Groudine, M. (1990) *Genes Dev.* **4**, 1637–1649.
- Epner, E., Reik, A., Cimborá, D., Telling, A., Bender, M. A., Fiering, S., Enver, T., Martin, D. I. K., Kennedy, M., Keller, G. & Groudine, M. (1998) *Mol. Cell* **2**, 447–455.
- Reik, A., Telling, A., Zitnik, G., Cimborá, D., Epner, E. & Groudine, M. (1998) *Mol. Cell Biol.* **18**, 5992–6000.
- Bonifer, C., Yannoutsos, N., Kruger, G., Grosveld, F. & Sippel, A. E. (1994) *Nucleic Acids Res.* **22**, 4202–4210.
- Aronow, B. J., Silbiger, R. N., Dusing, M. R., Stock, J. L., Yager, K. L., Potter, S. S., Hutton, J. J. & Wiginton, D. A. (1992) *Mol. Cell Biol.* **12**, 4170–4185.
- Madisen, L. & Groudine, M. (1994) *Genes Dev.* **8**, 2212–2226.
- Diaz, P., Cado, D. & Winoto, A. (1994) *Immunity* **1**, 207–217.
- Gribnau, J., de Boer, E., Trimborn, T., Wijgerde, M., Milot, E., Grosveld, F. & Fraser, P. (1998) *EMBO J.* **17**, 6020–6027.
- Wang, Y., Macke, J. P., Merbs, S. L., Zack, D. J., Klaunberg, B., Bennett, J., Gearhart, J. & Nathans, J. (1992) *Neuron* **9**, 429–440.
- Starck, J., Sarkar, R., Romana, M., Bhargava, A., Scarpa, A. L., Tanaka, M., Chamberlain, J. W., Weissman, S. M. & Forget, B. G. (1994) *Blood* **84**, 1656–1665.
- Buck, L. & Axel, R. (1991) *Cell* **65**, 175–187.
- Parmentier, M., Libert, F., Schurmans, S., Schiffmann, S., Lefort, A., Eggerickx, D., Ledent, C., Mollereau, C., Gerard, C., Perret, J., *et al.* (1992) *Nature (London)* **355**, 453–455.
- Ben-Arie, N., Lancet, D., Taylor, C., Khen, M., Walker, N., Ledbetter, D. H., Carrozzo, R., Patel, K., Sheer, D., Lehrach, H. & North, M. A. (1993) *Hum. Mol. Genet.* **3**, 229–235.
- Mombaerts, P., Wang, F., Dulac, C., Vassar, R., Chao, S. K., Nemes, A., Mendelsohn, M., Edmondson, J. & Axel, R. (1996) *Cold Spring Harbor Symp. Quant. Biol.* **61**, 135–145.
- Bender, M. A., Reik, A., Close, J., Telling, A., Epner, E., Fiering, S., Hardison, R. & Groudine, M. (1998) *Blood* **92**, 4394–4403.
- Vanin, E. F., Henthorn, P. S., Kiousis, D., Grosveld, F. & Smithies, O. (1983) *Cell* **35**, 701–709.
- Hardison, R. C., Oeltjen, J. & Miller, W. (1997) *Genome Res.* **7**, 959–966.
- Schaeren-Wiemers, N. & Gerfin-Moser, A. (1993) *Histochemistry* **100**, 431–440.
- Slightom, J., Bock, J., Tagle, D., Gumucio, D., Goodman, M., Stojanovic, N., Jackson, J., Miller, W. & Hardison, R. (1997) *Genomics* **39**, 90–94.
- Tuan, D., Solomon, W., Li, Q. & London, I. M. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 6384–6388.
- Fleener, D. F. & Kaufman, R. E. (1993) *Blood* **81**, 2781–2790.
- Forget, B. G. (1998) *Ann. N.Y. Acad. Sci.* **850**, 38–44.
- Rouquier, S., Taviaux, S., Trask, B. J., Brand-Arpon, V., van den Engh, G., Demaille, J. & Giorgi, D. (1998) *Nat. Genet.* **18**, 243–250.
- Buettner, J. A., Glusman, G., Ben-Arie, N., Ramos, P., Lancet, D. & Evans, G. A. (1998) *Genomics* **53**, 56–68.
- Sullivan, S. L., Adamson, M. C., Ressler, K. J., Kozak, C. A. & Buck, L. B. (1996) *Proc. Natl. Acad. Sci. USA* **93**, 884–888.
- Raming, K., Konzelmann, S. & Breer, H. (1998) *Recept. Channels* **6**, 141–151.
- Schild, D. & Restrepo, D. (1998) *Physiol. Rev.* **78**, 429–466.
- Ressler, K. J., Sullivan, S. L. & Buck, L. B. (1993) *Cell* **73**, 597–609.
- Vassar, R., Ngai, J. & Axel, R. (1993) *Cell* **74**, 309–318.
- Vanderhaeghen, P., Schurmans, S., Vassart, G. & Parmentier, M. (1997) *Genomics* **39**, 239–246.
- Blache, P., Gros, L., Salazar, G. & Bataille, D. (1998) *Biochem. Biophys. Res. Comm.* **242**, 669–672.
- Nef, S. & Nef, P. (1997) *Proc. Natl. Acad. Sci. USA* **94**, 4766–4771.
- Kamakaka, R. T. (1997) *Trends Biochem. Sci.* **22**, 124–128.
- Ortiz, B. D., Cado, D., Chen, V., Diaz, P. W. & Winoto, A. (1997) *EMBO J.* **16**, 5037–5045.
- Hong, N. A., Cado, D., Mitchell, J., Ortiz, B. D., Hsieh, S. N. & Winoto, A. (1997) *Mol. Cell Biol.* **17**, 2151–2157.
- Chung, J. H., Bell, A. C. & Felsenfeld, G. (1997) *Proc. Natl. Acad. Sci. USA* **94**, 575–580.
- Zhong, X. P. & Krangel, M. S. (1997) *Proc. Natl. Acad. Sci. USA* **94**, 5219–5224.