Molecular and Cellular Neuroscience xxx (2010) xxx–xxx

Contents lists available at ScienceDirect

Molecular and Cellular Neuroscience

journal homepage: www.elsevier.com/locate/ymcne

Homeodomain binding motifs modulate the probability of odorant receptor gene choice in transgenic mice

Anne Vassalli [a,1], Paul Feinstein [a,2], Peter Mombaerts [a,b,*]

[a] The Rockefeller University, 1230 York Avenue, New York, NY 10065, USA
[b] Max Planck Institute of Biophysics, Max-von-Laue-Strasse 3, 60438 Frankfurt, Germany

ARTICLE INFO

Article history:
Received 8 September 2010
Revised 11 October 2010
Accepted 1 November 2010
Available online xxxx

Keywords:
Odorant receptor
Olfactory receptor
Gene expression
Gene regulation

ABSTRACT

Odorant receptor (OR) genes constitute with 1200 members the largest gene family in the mouse genome. A mature olfactory sensory neuron (OSN) is thought to express just one OR gene, and from one allele. The cell bodies of OSNs that express a given OR gene display a mosaic pattern within a particular region of the main olfactory epithelium. The mechanisms and *cis*-acting DNA elements that regulate the expression of one OR gene per OSN — OR gene choice — remain poorly understood. Here, we describe a reporter assay to identify minimal promoters for OR genes in transgenic mice, which are produced by the conventional method of pronuclear injection of DNA. The promoter transgenes are devoid of an OR coding sequence, and instead drive expression of the axonal marker tau-β-galactosidase. For four mouse OR genes (*M71*, *M72*, *MOR23*, and *P3*) and one human OR gene (*hM72*), a mosaic, OSN-specific pattern of reporter expression can be obtained in transgenic mice with contiguous DNA segments of only ~300 bp that are centered around the transcription start site (TSS). The ~150 bp region upstream of the TSS contains three conserved sequence motifs, including homeodomain (HD) binding sites. Such HD binding sites are also present in the H and P elements, DNA sequences that are known to strongly influence OR gene expression. When a 19mer encompassing a HD binding site from the P element is multimerized nine times and added upstream of a *MOR23* minigene that contains the *MOR23* coding region, we observe a dramatic increase in the number of transgene-expressing founders and lines and in the number of labeled OSNs. By contrast, a nine times multimerized 19mer with a mutant HD binding site does not have these effects. We hypothesize that HD binding sites in the H and P elements and in OR promoters modulate the probability of OR gene choice.

Introduction

The mouse olfactory system relies on ~1200 distinct types of ORs to detect chemical information in the external environment (Buck and Axel, 1991). Odorous ligands bind to ORs on cilia of OSNs within the main olfactory epithelium (MOE) that lines the nasal cavity. Individual OSNs are thought to express only a single OR gene (Malnic et al., 1999), and express one allele of a given OR gene (Chess et al., 1994; Strotmann et al., 2000). Due to the size of the OR repertoire and to monogenic expression, the population of OSNs that express a given OR is relatively small, with the numbers varying over two orders of magnitude (Mombaerts, 2004). A given OR gene is expressed in a mosaic or punctate pattern in the mouse, within a characteristic region (typically referred to as zone) of the MOE (Ressler et al. 1993; Miyamichi et al., 2005). Axons of OSNs that express a given OR gene coalesce into one or a few glomeruli in each of the two halves of the two olfactory bulb in the mouse (Ressler et al., 1994; Mombaerts et al., 1996). OR gene expression is thus characterized by at least five features: monogenic expression on a per OSN basis, monoallelic expression on a per OSN basis, mosaic or punctate expression in the MOE, expression within a characteristic region of the MOE, and coalescence of OSN axons into glomeruli. These expression features are commonly referred to as OR gene choice.

As a first approach to understand the mechanisms and *cis*-acting DNA elements of OR gene choice, we and others have used transgenic mouse technology, in order to identify DNA segments of OR loci that reproduce some or all of these five features when inserted ectopically in the mouse genome (Bozza et al., 2009; Oka et al., 2006; Qasba and Reed, 1998; Rothman et al., 2005; Serizawa et al., 2000, 2003; Vassalli et al., 2002; Zhang et al., 2007). Our ~9 kilobase (kb) transgenes for *M71* and *MOR23* (Rothman et al., 2005; Vassalli et al., 2002) and the 10.5 kb transgene for *MOR262-12* (Zhang et al., 2007) are the smallest transgenes that provide the best reproduction of the five features of the endogenous OR genes. Most importantly, axons of transgene-

* Corresponding author. Max Planck Institute of Biophysics, Max-von-Laue-Strasse 3, 60438 Frankfurt, Germany. Fax: +49 69 6300 4002.
E-mail address: peter.mombaerts@biophys.mpg.de (P. Mombaerts).
[1] Current address: Center for Integrative Genomics, Université de Lausanne, 1015 Lausanne, Switzerland.
[2] Current address: Department of Biological Sciences, Hunter College, New York, NY 10065, USA.

1044-7431/$ – see front matter © 2010 Elsevier Inc. All rights reserved.
doi:10.1016/j.mcn.2010.11.001

expressing OSNs co-converge with OSNs expressing the corresponding endogenous OR locus to same glomeruli. These small transgenes, which contain the OR coding sequence, are typically referred to as minigenes.

The *M71* and *MOR23* minigenes (Rothman et al., 2005; Vassalli et al., 2002) consist of four segments from the OR locus: a nontranscribed region upstream of the TSS; one or two 5′ nontranslated exons; the OR coding sequence (which has no introns); and a 3′ nontranslated region. Deletion of both introns of *MOR23* still permits transgene expression: Tg3′Δ with 2.2 kb of genomic DNA (Vassalli et al., 2002). *MOR23* minigene expression requires a 405 bp region upstream of the TSS, which consists of a 257 bp fragment of a LINE-1 repetitive DNA element followed by 148 bp of unique sequence (Vassalli et al., 2002). For the *M71* minigene, a 161 bp region upstream of the TSS is required for transgene expression: Tg(161) with 7.1 kb of genomic DNA (Rothman et al., 2005). The *MOR262-12* minigene contains 358 bp upstream of the TSS (Zhang et al., 2007).

Conserved sequence motifs reside within these promoter regions (Rothman et al., 2005; Vassalli et al., 2002; Zhang et al., 2007). Two of these motifs lie in close proximity of each other: a putative homeodomain (HD) binding site and an OLF/EBP (O/E)-binding site. By site-directed mutagenesis of the *M71* minigene and of the corresponding sequences in the endogenous *M71* locus using gene targeting, we have demonstrated that these HD and O/E motifs are involved in transgene and endogenous OR gene expression, respectively (Rothman et al., 2005). We identified the LIM-homeodomain protein LHX2 using the HD binding site in the *M71* promoter as bait (Hirota and Mombaerts, 2004). We found that *Lhx2* is required for class II OR but not for class I OR gene expression and/or for maturation of the OSNs that express class II OR genes (Hirota and Mombaerts, 2004; Hirota et al., 2007).

Classical strategies of promoter mapping are performed in cell lines or *in vivo* with constructs in which short upstream genomic DNA segments are fused to a sequence encoding a reporter such as GFP or lacZ. All but the first (Qasba and Reed, 1998) transgenic OR constructs include the OR coding sequence and some surrounding sequences, because of the well-established role of the OR in coalescence of axons into glomeruli (Mombaerts et al., 1996; Mombaerts, 2006). Co-convergence of axons of transgene-expressing OSNs to the same glomeruli as axons of OSNs expressing the corresponding endogenous OR is the best criterion to assess the fidelity of the expression patterns. However, there is some evidence that the OR coding sequence itself is somehow involved in the regulation of OR gene expression (Nguyen et al., 2007). To dissect further the sequences with promoter activity in transgenic mice, there is thus a need for a reduced reporter assay that is devoid of OR coding sequence.

Here, we have developed such a transgenic reporter assay. We have used it to define minimal OR promoters for mouse and human OR genes in transgenic mice, which are produced conventionally by pronuclear injection of DNA. These promoter transgenes drive expression of the axonal maker tau-β-galactosidase. The transgenic OR promoters are ~300 bp contiguous segments centered around the TSS as defined by 5′RACE, or around the predicted TSS for the human OR gene as defined by sequence homology. We show that mutations in the HD and O/E binding sites abolish transgene expression of a 337 bp minimal promoter for *M71*. A segment of 317 bp that has high sequence identity to the P3 minimal promoter but could not be linked to transcripts containing an OR coding sequence, confers unexpected expression in up to 10% of OSNs, and has been termed the P element (Bozza et al., 2009). The ~150 bp regions upstream of these TSSs contain conserved motifs of HD and O/E binding sites, and another motif just upstream of the TSS. A similar arrangement of motifs is found in a central region of the 2.1 kb H element (Serizawa et al., 2003; Nishizumi et al., 2007). We further concentrated on a 13 bp perfect match containing a HD binding site that is shared among the P3 minimal promoter, the H and the P elements. When a 19 bp sequence from the P element that encompasses this 13 bp sequence is multimerized nine times and inserted upstream of a *MOR23* minigene (which contains the *MOR23* coding sequence), most transgenic founders and lines express the transgene, and many lines display a dramatic increase in the number of labeled OSNs. By contrast, a nine times-multimerized 19 bp sequence with a mutant HD binding site does not have these effects on the number of transgene-expressing founders and lines, and on the number of labeled OSNs; it is essentially neutral with regard to transgene expression.

We hypothesize that these HD binding sites modulate the probability of OR gene choice: the frequency with which a given OR gene is chosen for expression by an OSN within a characteristic region of the MOE.

Results

Transgenic promoters of ~300 bp contiguous segments

In our novel transgenic assay for OR promoter activity, there is no heterologous DNA sequence to provide basal promoter activity, and no OR coding sequence. The segments from OR loci are contiguous and ~300 bp in length. These segments consist of ~150 bp upstream of the TSS followed by ~150 bp downstream; the downstream segment corresponds, generally, to the first nontranslated exon. These segments are tested in the context of a downstream *tauLacZ* reporter sequence, which is followed by a rabbit β-globin polyadenylation signal (*-LacZpA*). We term these transgenes Tg-OR(x,y)-LacZpA, with OR the locus, and x and y the length in bp of sequence that is upstream and downstream, respectively, of the TSS as defined by 5′RACE.

In minigenes (Oka et al., 2006; Rothman et al., 2005; Vassalli et al., 2002; Zhang et al., 2007), the presence of an intact OR coding sequence enables monogenic expression and coalescence of labeled axons into glomeruli (Figs. 1A, D). By contrast, in our reduced reporter assay, the promoter transgenes do not contain an OR coding sequence. We find that labeled axons project diffusely to a broad domain of the OB (Figs. 1B, E, 2), instead of coalescing into one or a few glomeruli per half-OB, as is the case with minigenes (Figs. 1A, D). This phenotype of diffuse axonal projections is commonly observed with other transgenes (small or large) that lack a functional OR coding sequence (Serizawa et al., 2003) and with gene-targeted mutations that delete or otherwise cripple the OR coding sequence (Bozza et al., 2009; Feinstein et al., 2004; Grosmaitre et al., 2009; Lewcock and Reed, 2004; Shykind et al., 2004).

We have here studied four mouse OR genes and one human OR gene (Figs. 1–3).

Mouse M71

We have previously shown that a minigene with 9.2 kb of *M71* genomic DNA confers tau-β-galactosidase expression to OSNs within a dorsal region of the MOE, and enables axonal coalescence into glomeruli that reside within the dorsal–caudal OB (Vassalli et al., 2002). The genomic sequences in this Tg-M71-LacZ minigene, which was abbreviated as M71-TgA in Vassalli et al., 2002, consist of 2.3 kb upstream of the TSS, followed by the nontranslated exon, the 2.0 kb intron, and the coding exon of *M71*, with 3.8 kb downstream of the *M71* stop codon. The *IRES-taulacZ* sequence was inserted between the *M71* stop codon and the 3.8 kb downstream segment (Fig. 1A).

We have now dissected the *M71* promoter further in our promoter transgene assay. We used a 337 bp segment of the *M71* locus that consists of the same 161 bp upstream of the TSS as in Tg (161) (Rothman et al., 2005), followed by the entire 5′ nontranslated exon (176 bp), constructing promoter transgene Tg-M71(161,176)-LacZpA (Fig. 1B). Of four transgenic founders, three display mosaic β-galactosidase expression in the dorsal MOE (Figs. 1B, 3), similar to the *M71* minigene (Fig. 1A). By breeding three additional founders, we established three independent mouse lines carrying this transgene.

Please cite this article as: Vassalli, A., et al., Homeodomain binding motifs modulate the probability of odorant receptor gene choice in transgenic mice, Mol. Cell. Neurosci. (2010), doi:10.1016/j.mcn.2010.11.001

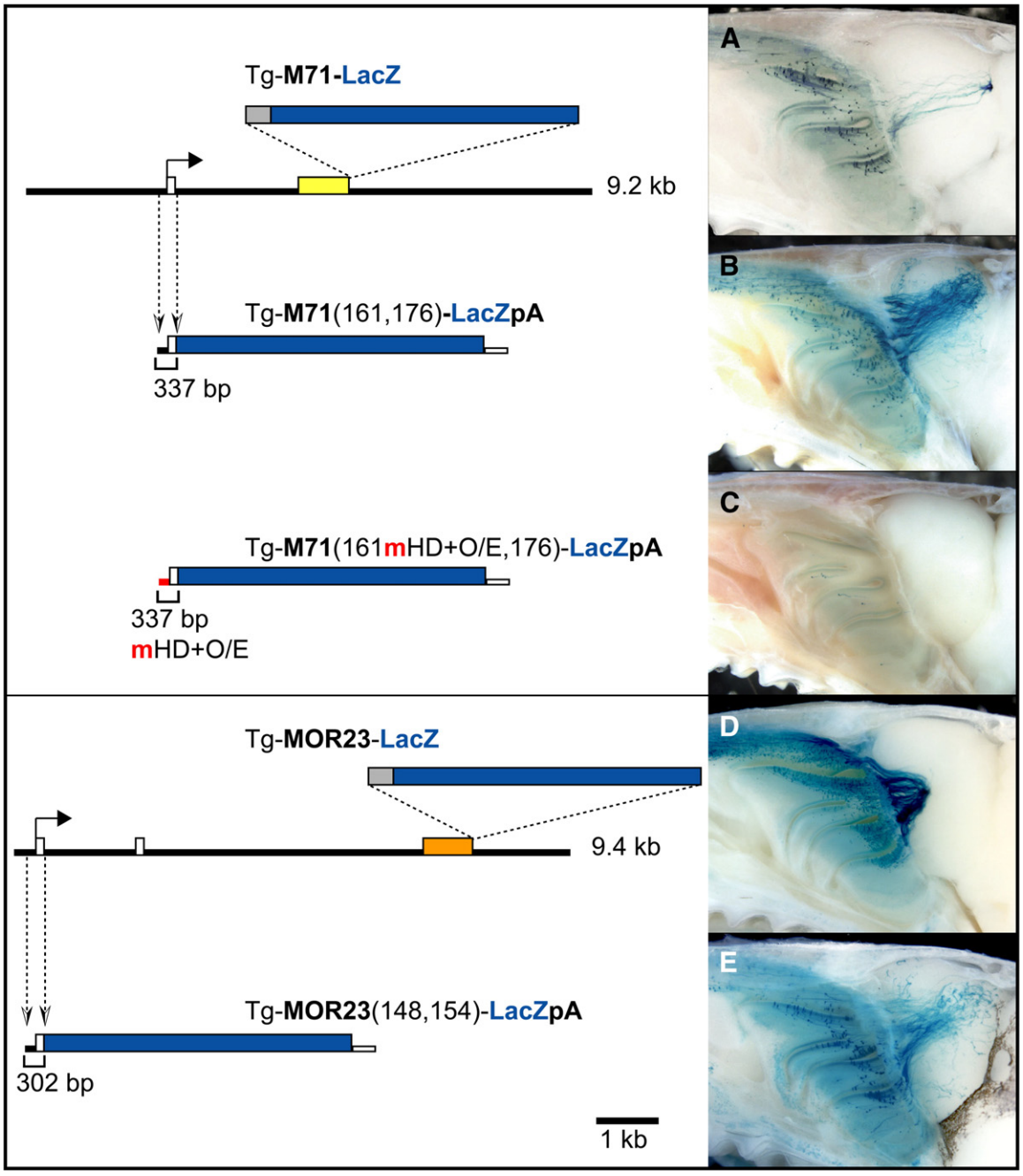


Fig. 1. Minimal *M71* and *MOR23* promoter transgenes derived from OR minigenes. Left, design of the transgenes. Right, X-gal stained wholemounts from a right half-head, view on the medial aspect; anterior is to the left, dorsal is to the top. (A) Minigene Tg-M71-LacZ (Vassalli et al., 2002). After the stop codon of the *M71* coding sequence (yellow box) is inserted an *IRES* (grey box) and the tauLacZ sequence (blue box). This transgene contains a total of 9.2 kb from the *M71* locus: 2.3 kb upstream of the TSS, the 5′ nontranslated exon (short white box), the 2.0 kb intron, the 930 bp *M71* coding sequence, and 3.8 kb downstream of the stop codon. Hooked arrow indicates TSS and direction of transcription. The wholemount is from a mouse of line M71-TgA described in Vassalli et al., 2002. X-gal labeled cells reside in the dorsal region of the MOE, and their axons coalesce into one glomerulus in the medial half-bulb. (B) Promoter transgene Tg-M71(161,176)-LacZpA. A 337 bp contiguous fragment from the *M71* locus centered around the TSS of *M71* is placed 5′ to the tauLacZ sequence (blue box), which is followed by the rabbit globin polyadenylation sequence (thin white box). The 337 bp fragment represents 161 bp (black box) upstream of the TSS and 176 bp (thick white box) of *M71* non-coding exon 1. This transgene does not contain an OR coding sequence. A transgenic founder shows X-gal labeled cells predominantly in the dorsal region of the MOE. In contrast to (A), labeled axons do not coalesce into glomeruli, but spread over a large domain in the dorsal aspect of the medial half-bulb. (C) Promoter transgene Tg-M71(161*m*HD + O/E,176)-LacZpA. Within the same 337 bp segment from the *M71* locus as in (B), two conserved motifs, the homeodomain (HD) and OLF/EBP (O/E) binding sites have been mutated from *TAATTG* to *TGGTTG* and from *ATCCCAGGAGAT* to *ATAGGAGGAGAT*, respectively. These mutations are within the 161 bp segment (thin red box) upstream of the TSS. A transgenic founder shows a very few X-gal labeled cells in the MOE. No labeled axons are observed on the olfactory bulb. (D) Minigene Tg-MOR23-LacZ (Vassalli et al., 2002). After the stop codon of the *MOR23* coding sequence (orange box) is inserted an *IRES* (grey box) and the taulacZ sequence (dark blue box). This transgene contains a total of 9.4 kb from the *MOR23* locus: 405 bp upstream of the TSS, two 5′ untranslated exons (short white boxes), the 930 bp *MOR23* coding sequence, and 1.7 kb downstream of the stop codon. Hooked arrow indicates TSS and direction of transcription. The wholemount is from a mouse of line TgSN-2 described in Vassalli et al., 2002. X-gal labeled cells are found in the dorsal region of the MOE, and their axons coalesce into one glomerulus in the medial half-bulb. (E) Promoter transgene Tg-MOR23(148,154)-LacZpA. A 302 bp contiguous fragment of the *MOR23* locus centered around the TSS of *MOR23* is placed 5′ to the tauLacZ sequence (blue box), which is followed by the rabbit globin polyadenylation sequence (thin white box). The 302 bp fragment represents 148 bp TSS (black box) upstream of the TSS and 154 bp of nontranslated exon 1 (thick white box). This transgene does not contain an OR coding sequence. The wholemount is from a mouse of line I. X-gal labeled cells reside predominantly in the dorsal region of the MOE. In contrast to (D), labeled axons do not coalesce into glomeruli, but project diffusely to a large subdomain centrally in the medial half-bulb.

Please cite this article as: Vassalli, A., et al., Homeodomain binding motifs modulate the probability of odorant receptor gene choice in transgenic mice, Mol. Cell. Neurosci. (2010), doi:10.1016/j.mcn.2010.11.001

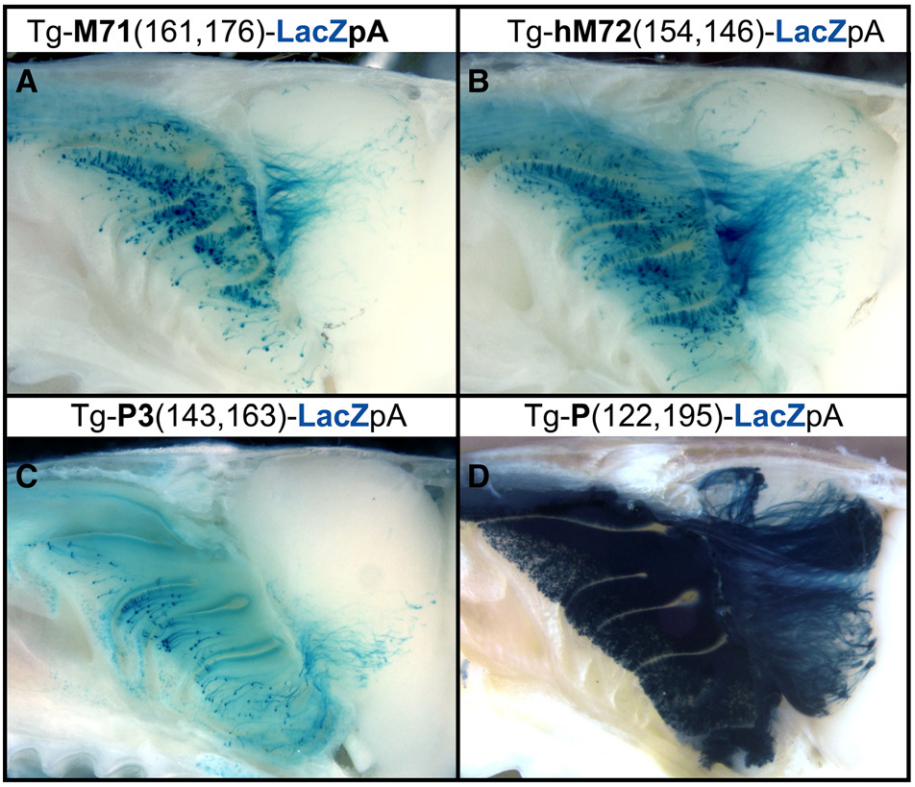


Fig. 2. Additional promoter transgenes. (A) Promoter transgene Tg-M71(161,176)-LacZpA. Wholemount from a mouse of line I. See also Fig. 1B. (B) Promoter transgene Tg-hM72 (154,146)-LacZpA. A 300 bp contiguous fragment from a human OR locus defined by homology to the *M71* and *M72* promoters, is placed 5′ to the tauLacZpA reporter. The 300 bp is presumed to represent 154 bp upstream of the hypothetical human TSS and 146 bp of a hypothetical nontranslated exon 1 (a putative splice donor site is observed following the 146 bp sequence). This transgene does not contain an OR coding sequence. Wholemount from a mouse of line IV. X-gal labeled cells reside predominantly in a ventral region of the MOE. Labeled axons project diffusely to a large subdomain of the medial half-OB. (C) Promoter transgene Tg-P3(143,163)-LacZpA. A 306 bp contiguous fragment from the *P3* locus centered around the TSS of P3 is placed 5′ to the tauLacZpA reporter. The 306 bp fragment represents 143 bp upstream of the putative TSS and 163 bp of *P3* nontranslated exon 1 (Lane et al., 2001). This transgene does not contain an OR coding sequence. Wholemount from a mouse of line IV. X-gal labeled cells reside predominantly in a ventral region of the MOE. Labeled axons project diffusely to a large subdomain in the ventral, medial half-OB. (D) Transgene Tg-P(122,195)-LacZpA. A 317 bp fragment (the P element) defined by homology to the minimal *P3* promoter is placed 5′ to the tauLacZpA reporter. This transgene does not contain an OR coding sequence. Wholemount from a mouse of P-LacZ-Tg line 8 (Bozza et al., 2009). X-gal labeled neurons reside throughout the entire MOE. Labeled axons project diffusely to the majority of the surface of the medial half-OB except for a dorsal, wedge-shaped domain.

One line does not express the transgene, and the two other lines have similar numbers of labeled OSNs as the gene-targeted strain M71-IRES-taulacZ (Fig. 3).

This 161 bp upstream sequence contains one HD and one O/E binding site (Rothman et al., 2005). We have previously shown that mutations of both the HD and O/E sites within a minigene that contains a slightly longer upstream sequence, Tg(491, Hom + O/E), result in nearly no transgene expression (Rothman et al., 2005). To test the relevance of these sites in the expression of promoter transgenes, the same five nucleotide substitutions as in minigene Tg (491, Hom + O/E) were introduced in order to convert the HD binding site *TAATTG* into *TGGTTG* and the O/E binding site *ATCCCAGGAGAT* into *ATAGGAGGAGAT*, constructing promoter transgene Tg-M71(161mHD + O/E, 176)-LacZpA (Fig. 1C). These mutations are known to interfere with the binding of HD and O/E proteins, respectively, to these DNA sequences. We find that these five substitutions also result in no transgene expression in 2/6 founders and low transgene expression in the other four founders (Fig. 3). No mouse strains were established. Thus, our transgenic reporter assay, which does not include the OR coding sequence and intronic sequences, confirms the relevance of the HD and O/E binding sites in the 161 bp region upstream of the TSS.

To determine whether the TSSs in the M71 promoter transgene are the same as in the endogenous *M71* gene, we performed 5′ RACE on RNA from olfactory mucosa of transgenic mice. Two independent Tg-M71(161, 176)-LacZpA lines (lines I and IV) reveal TSSs at the exact same nucleotide (data not shown) as the endogenous *M71* gene (Rothman et al., 2005). The promoter transgenes are thus not expressed from alternate promoters in the genome.

Mouse MOR23

We have previously described a *MOR23* minigene of 9.4 kb genomic DNA, with 405 bp upstream sequence and two 5′ nontranslated exons (Vassalli et al., 2002). An example of a mouse carrying this minigene Tg-MOR23-LacZ (line TgSN-2 in Vassalli et al., 2002) is shown in Fig. 1D. By fusing the two nontranslated exons directly with the *MOR23* coding sequence, we have shown that the introns are not necessary for transgene expression (TgΔΔ and Tg3'Δ Vassalli et al., 2002).

We have now further dissected the *MOR23* promoter in our promoter transgene assay. We used a 302 bp segment of the *MOR23* locus that consists of 148 bp upstream of the TSS (which represents all unique sequences of the 405 bp upstream sequence of Tg-MOR23-LacZ) and the entire 5′ nontranslated exon 1 of *MOR23* (154 bp), constructing promoter transgene Tg-MOR23(148,154)-LacZpA (Fig. 1E). Of four transgenic lines, two express β-galactosidase activity in a mosaic OSN-specific pattern (Fig. 1E). At four weeks of age, the distribution in the MOE is very similar to that of TgSN-2 mice, but there are a few X-gal stained cells that are more ventral (Fig. 1D). At postnatal day 8, the expression domain in both Tg-MOR23(148,154)-LacZpA lines is broader and extends ventrally (data not shown). We have also observed labeled cells in the ventral MOE of young mice of the gene-targeted MOR23-IRES-tauGFP strain, suggesting that these ectopic cells exist naturally but do not persist, presumably during a process of refinement (Vassalli et al., 2002).

Please cite this article as: Vassalli, A., et al., Homeodomain binding motifs modulate the probability of odorant receptor gene choice in transgenic mice, Mol. Cell. Neurosci. (2010), doi:10.1016/j.mcn.2010.11.001

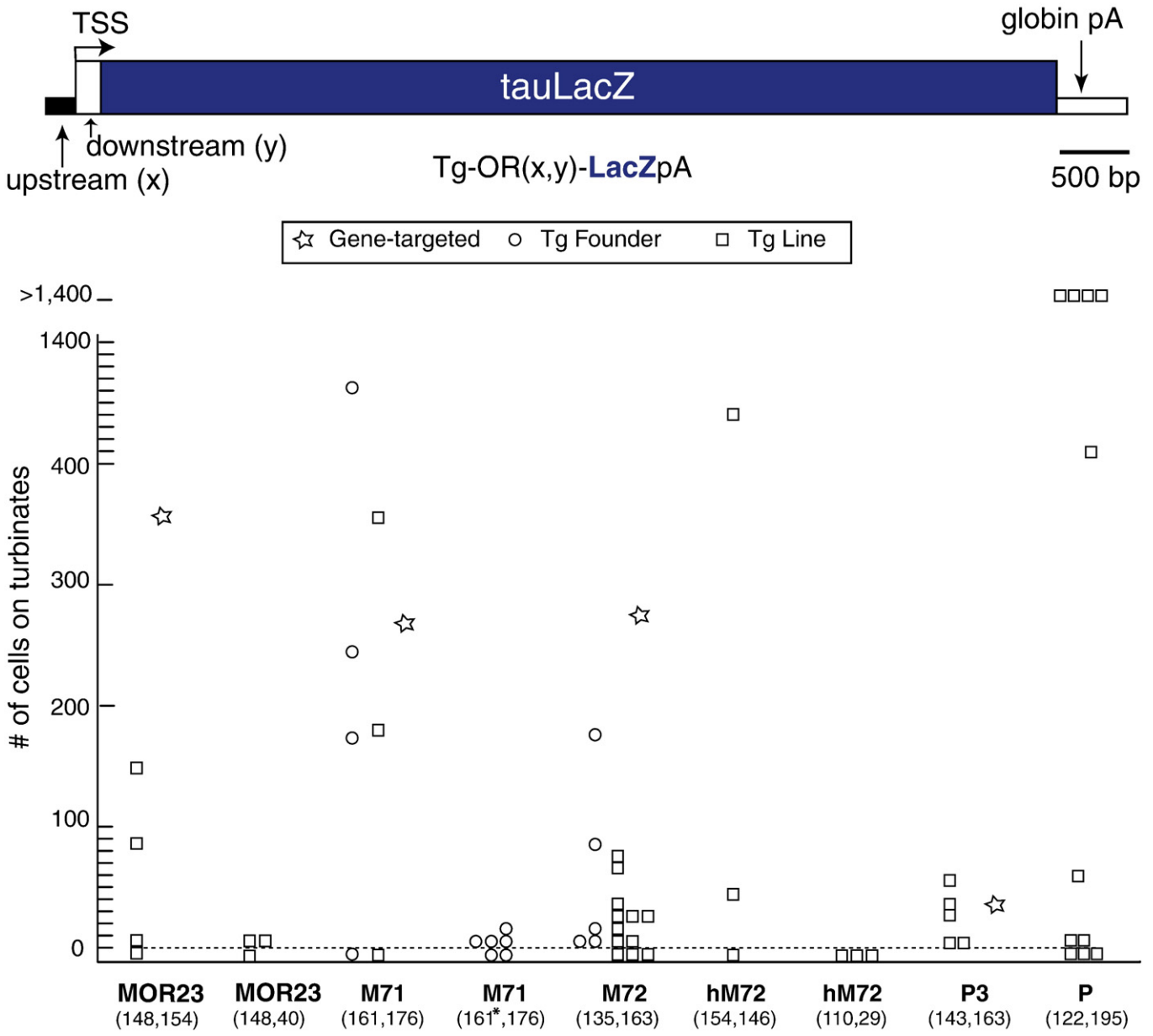


Fig. 3. Cell counts for the various promoter transgenes. Eight minimal OR promoters and the P element were tested for expression in our transgenic assay using the tauLacZ reporter and the rabbit globin polyadenylation sequence. None of these transgenes contains an OR coding sequence. Nomenclature of promoter fragments is OR(X,Y) with OR representing the name of the OR gene, X representing the sequence upstream of the TSS and Y the sequence downstream of the TSS. The mHD + O/E mutation in the M71 promoter is indicated with an asterisk after 161. Counts of X-gal labeled cells on the medial aspect of olfactory turbinates in wholemounts (in ranges of 10 s) are from transgenic founders (circles) and from lines (squares). Cell counts from lines are from at least three mice, which are hemi- or homozygous for the transgene. For reference, labeled cells were counted for at least three heterozygous mice (stars) of the corresponding lines with gene-targeted IRES-taulacZ insertions: MOR23, M71, M72, and P3.

We truncated the *MOR23* promoter from 154 to 40 bp downstream of the TSS, thereby deleting one of the two conserved O/E sites in exon 1, constructing promoter transgene Tg-MOR23(148,40)-LacZpA (Fig. 3). Among three transgenic lines, one does not express and two show a very few labeled neurons in the MOE, with an extended, ventralized distribution (data not shown). Although the number of lines is limited, the effect of this truncation suggests a role for the second O/E binding site in the first *MOR23* nontranslated exon.

Mouse M72

The *M72* OR gene is very similar to the *M71* OR gene: it is also expressed in the dorsal MOE, it has a single 5′ nontranslated exon, 96% bp identity in the OR coding sequence, and high sequence identity in the promoter region (Vassalli et al., 2002). *M71* and *M72* share three strikingly conserved motifs in the promoter region: two HD binding sites, one O/E binding site, and a third motif just upstream of the TSS (Vassalli et al., 2002; Fig. 5). No M72 minigenes have been reported.

We asked whether the *M71*-homologous region upstream of the *M72* TSS (135 bp) plus the first nontranslated exon of *M72* (163 bp) would be sufficient for transgene expression, constructing promoter transgene Tg-M72(135,163)-LacZpA. We obtained an unusually high efficiency of transgene integration, and generated 17 founders. We find that 5/5 founders and 9/12 mouse lines show mosaic β-galactosidase expression in the MOE (Fig. 3). Expression is ventralized relative to the normal pattern however (data not shown). Thus, our novel assay demonstrates that a ~300 bp region centered around the TSS generates *M72*-like promoter activity in a promoter transgene, without having characterized previously *M72* minigenes.

Please cite this article as: Vassalli, A., et al., Homeodomain binding motifs modulate the probability of odorant receptor gene choice in transgenic mice, Mol. Cell. Neurosci. (2010), doi:10.1016/j.mcn.2010.11.001

Human hM72

The human genome harbors on Chromosome 11 an OR gene that is orthologous to both the mouse *M71* and *M72* OR genes: *OR8A1*. We term this gene *hM72*, with 'h' for 'human'. The putative upstream region of *hM72* also contains the conserved HD and O/E binding sites and a predicted TSS (Fig. 5). In view of our success with the *M71* and *M72* promoter transgenes, we tested the function of the predicted *hM72* promoter in transgenic mice, constructing promoter transgene Tg-hM72(154,146)-LacZpA. Of three lines, two reveal mosaic expression in the MOE. The most robustly expressing line shows β-galactosidase-positive OSNs distributed across the entire MOE except for the most ventral aspect (Fig. 2B). We then constructed promoter transgene Tg-hM72(110,29)-LacZpA, in which the 5′ end is truncated by 44 bp (leaving the three conserved motifs intact) and the 3′ end of the first nontranslated exon is truncated by 117 bp. None of three transgenic lines show expression (Fig. 3). To our knowledge, hM72(154,146)-LacZpA defines the first promoter region for a human OR gene experimentally in mice.

Mouse P3

The mouse *MOR23*, *M71* and *M72* OR genes are expressed in the dorsal MOE. To extend our analysis of OR promoters to an OR gene that is expressed in the ventral MOE, we chose the OR gene *P3*, for which we have reported the gene-targeted strains P3-IRES-tauGFP and P3-IRES-taulacZ (Feinstein and Mombaerts, 2004). The complete exon–intron structure of *P3* is known (Lane et al., 2001). We tested a segment of 306 bp that consists of 143 bp upstream of the TSS and a portion of the first non-coding exon (163 bp): promoter transgene Tg-P3(143,163)-LacZpA. The 143 bp upstream region contains a variant HD motif and an O/E site in close proximity (Fig. 5). All five transgenic lines show mosaic β-galactosidase expression within the ventral MOE, similar in region and in numbers to P3-IRES-tauLacZ (Figs. 2C, 3).

Thus, short and proximal promoters are not limited to OR genes with expression in the dorsal MOE.

The P element

Unlike the *P3* OR gene, the exon–intron structure of the *P4* OR gene has not been reported. We performed 5′ RACE analysis for *P4* and identified a potential TSS site. Because the upstream region reveals no conserved motifs and no homology with the 306 bp *P3* promoter sequence, no promoter transgene was constructed for *P4*.

Interestingly, alignment of the 306 bp *P3* promoter sequence with the genomic region between the *P3* and *P4* coding sequences reveals a stretch of striking homology (Lane et al., 2001; Bozza et al., 2009; Fig. 5). This homology region is located 22.1 kb upstream of the putative *P4* TSS. We reasoned that this homology region may represent an additional or cryptic TSS for the *P4* OR gene, although we were unable to link it by RT-PCR with the *P4* coding sequence. Close comparison between this homology region and the 306 bp *P3* promoter sequence reveals a 317 bp equivalent segment, with a set of common HD and O/E binding sites. We have termed this 317 bp sequence the 'P element' (Bozza et al., 2009). Promoter transgene Tg-P(122,195)-LacZpA (referred to as P-Tg-LacZ in Bozza et al., 2009) was generated according to the same design principles as for the promoter transgenes described above.

Among 11 lines, eight show mosaic expression of β-galactosidase in the MOE (Figs. 2, 3). Remarkably, of these eight lines, in four the numbers of labeled cells are dramatically higher than in all the other promoter transgenes described above — as high as 10% of all cells in the MOE. This unusually high cell number can be seen both with a single-copy integration transgene (line 8, Fig. 2D) and with a multicopy integration transgene (line 11, data not shown). To date there is no explanation for this high frequency of expression of the P element across OSNs. We have demonstrated (Bozza et al., 2009) that promoter transgene Tg-P(122,195)-LacZpA can serve as a marker for the population of OSNs that express class II (but not class I) OR genes.

MOE expression patterns and axonal projections

MOR23 and *M71* minigenes reproduce closely the expression patterns in the dorsal MOE (Vassalli et al., 2002; Rothman et al., 2005). The novel *MOR23* and *M71* promoter transgenes generate β-galactosidase expression in patterns that are similar to the corresponding minigenes, but some additional expression in the ventral MOE is observed. When Tg-MOR23(148,154)-LacZpA mice are crossed to gene-targeted MOR23-IRES-tauGFP mice, we observe overlapping patterns of labeled cells in the MOE; curiously, labeled cells often come in pairs (Figs. 4A, B). The expression pattern of Tg-P3(143,163)-LacZpA is also shifted slightly ventrally relative to P3-IRES-tauLacZ mice (Fig. 2C). In a cross to gene-targeted P3-IRES-tauGFP mice, the patterns of labeled cells in the MOE overlap (Fig. 4D). Thus, the promoter transgenes produce patterns that are very similar but not identical in all their details to the endogenous counterparts.

Contrary to the minigenes, the promoter transgenes do not express an OR protein from the transgene. Their axons are thus not expected to coalesce into one or a few glomeruli per half-OB. In wholemounts, we observe that labeled axons project diffusely to a broad subdomain of the OB in all lines (Figs. 1B, E, 2). In histological sections, axons enter or pass through glomeruli. Some labeled axons of Tg-MOR23(148,154)-LacZpA mice enter or pass through the endogenous MOR23 glomeruli; axons that synapse in these glomeruli, would reflect the occasional co-expression of the promoter transgene with the endogenous MOR23 gene (Figs. 4C, C'). Some labeled axons of Tg-P3(143,163)-LacZpA mice enter or pass through glomeruli that are adjacent to the endogenous P3 glomeruli (Fig. 4E). In striking contrast, labeled axons of Tg-P(122,195)-LacZpA mice project, in varying extents, to numerous glomeruli across the OB except for the most dorsal part (Fig. 4F).

In conclusion, the diffuseness of the OSN axonal projections in the promoter transgenes is consistent with patterns reported for transgenic or gene-targeted mice in which an OR coding sequence is deleted and replaced with a marker (Bozza et al., 2009; Feinstein et al., 2004; Grosmaitre et al., 2009; Lewcock and Reed, 2004; Serizawa et al., 2003; Shykind et al., 2004). The fraction of the promoter transgene-expressing OSNs that coexpress an OR from an endogenous locus remains to be determined.

Sequence homology between minimal OR promoters and the H and P elements

Alignment of the ~150 bp sequence upstream of the TSS of *M71*, *M72*, and *hM72* reveals three highly conserved motifs (Fig. 5A, top): a HD binding site ($TAATN_5N_6$), an O/E binding site (*YYYCARRRR*), and a third conserved sequence (*AAACAG/CA*; Vassalli et al., 2002) near the TSS. The latter sequence is identical to an H1-box variant upstream of the Histone H1 gene (Eilers et al., 1994). A second conserved HD binding site is at the 5′ end of the promoter regions, but in the antiparallel orientation (Fig. 5A, top). These motifs are also found in a similar relative arrangement when the 143 bp sequence upstream of the TSS of mouse P3 and the corresponding 122 bp sequence of the P element in mouse are compared (Fig. 5A, middle).

The P3 minimal promoter and the P element share sequence motifs, but produce vastly different numbers of transgene-expressing OSNs. The high number of OSNs expressing the P transgene is reminiscent of the high numbers of *MOR28*-expressing OSNs when the H element (2125 bp) is placed upstream of a *MOR28* minigene (Serizawa et al., 2003). We reasoned that there may be sequences common to the H element and our 317 bp P element. Indeed, a scan of the H element sequence reveals several striking similarities in the

Please cite this article as: Vassalli, A., et al., Homeodomain binding motifs modulate the probability of odorant receptor gene choice in transgenic mice, Mol. Cell. Neurosci. (2010), doi:10.1016/j.mcn.2010.11.001

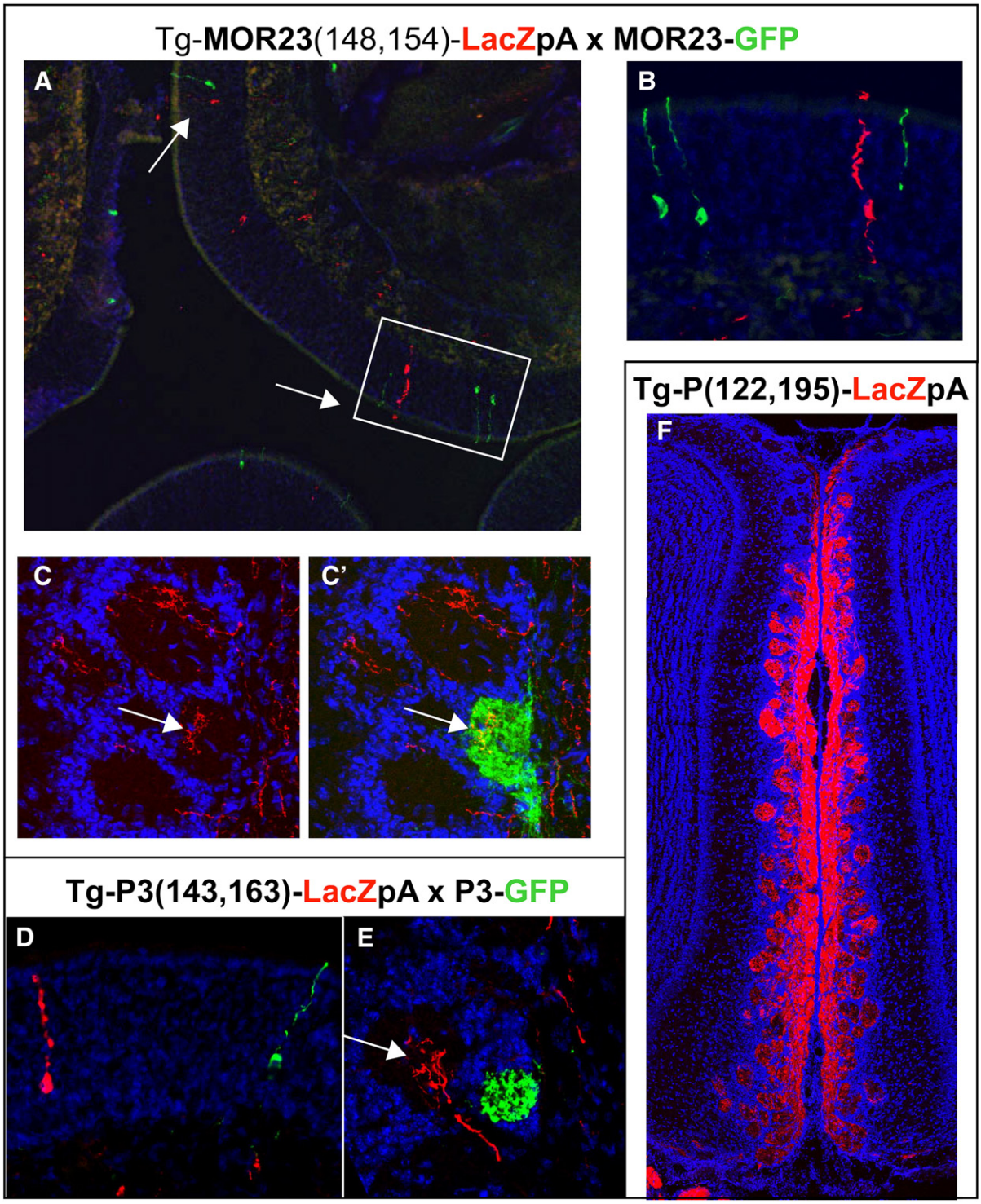


Fig. 4. Promoter transgene expression in the MOE and axonal projections. (A) Section through the dorsal aspect of the MOE from a Tg-MOR23(148,154)-LacZpA mouse (line I) crossed with MOR23-IRES-tauGFP. OSNs that express the promoter transgene are red-fluorescent (anti-β-galactosidase), and OSNs that express the endogenous *MOR23* locus are green-fluorescent (GFP). Nuclei are labeled blue. Red- and green-fluorescent OSNs are often in close proximity to each other (arrows). (B) Higher magnification of the box in (A). (C) Section through the OB of a Tg-MOR23(148,154)-LacZpA mouse (line I) crossed with MOR23-IRES-tauGFP. Glomeruli are outlined in blue by the nuclei of periglomerular cells. Red-fluorescent axons, which are from OSNs expressing the transgene, enter or pass through numerous glomeruli, including an endogenous MOR23 glomerulus (arrow). (C') Same section as (C). Green-fluorescent axons, which are from OSNs expressing the endogenous *MOR23* locus, coalesce into a glomerulus. Arrow indicates red-fluorescent axons that either enter or pass through the MOR23 glomerulus. (D) Section through the ventral aspect of the MOE from a Tg-P3(143,163)-LacZpA mouse (line IV) crossed with P3-IRES-tauGFP. OSNs that express the transgene are red-fluorescent (anti-β-galactosidase), and OSNs that express the endogenous *P3* locus are green-fluorescent (GFP). Nuclei are labeled blue. (E) Section through the olfactory bulb of a Tg-P3(143,163)-LacZpA mouse (line IV) crossed with P3-IRES-tauGFP. Glomeruli are outlined in blue by the nuclei of periglomerular cells. A red-fluorescent axon or axon bundle (anti-β-galactosidase) enters a glomerulus next to an endogenous P3 glomerulus, which is green-fluorescent. (F) Section through the medial portion of both bulbs from a Tg-P4(122, 195)-LacZpA mouse (line 8, Bozza et al., 2009). Glomeruli are outlined in blue by the nuclei of periglomerular cells. Red-fluorescent axons (anti-β-galactosidase) innervate many glomeruli, and to various extents.

Please cite this article as: Vassalli, A., et al., Homeodomain binding motifs modulate the probability of odorant receptor gene choice in transgenic mice, Mol. Cell. Neurosci. (2010), doi:10.1016/j.mcn.2010.11.001

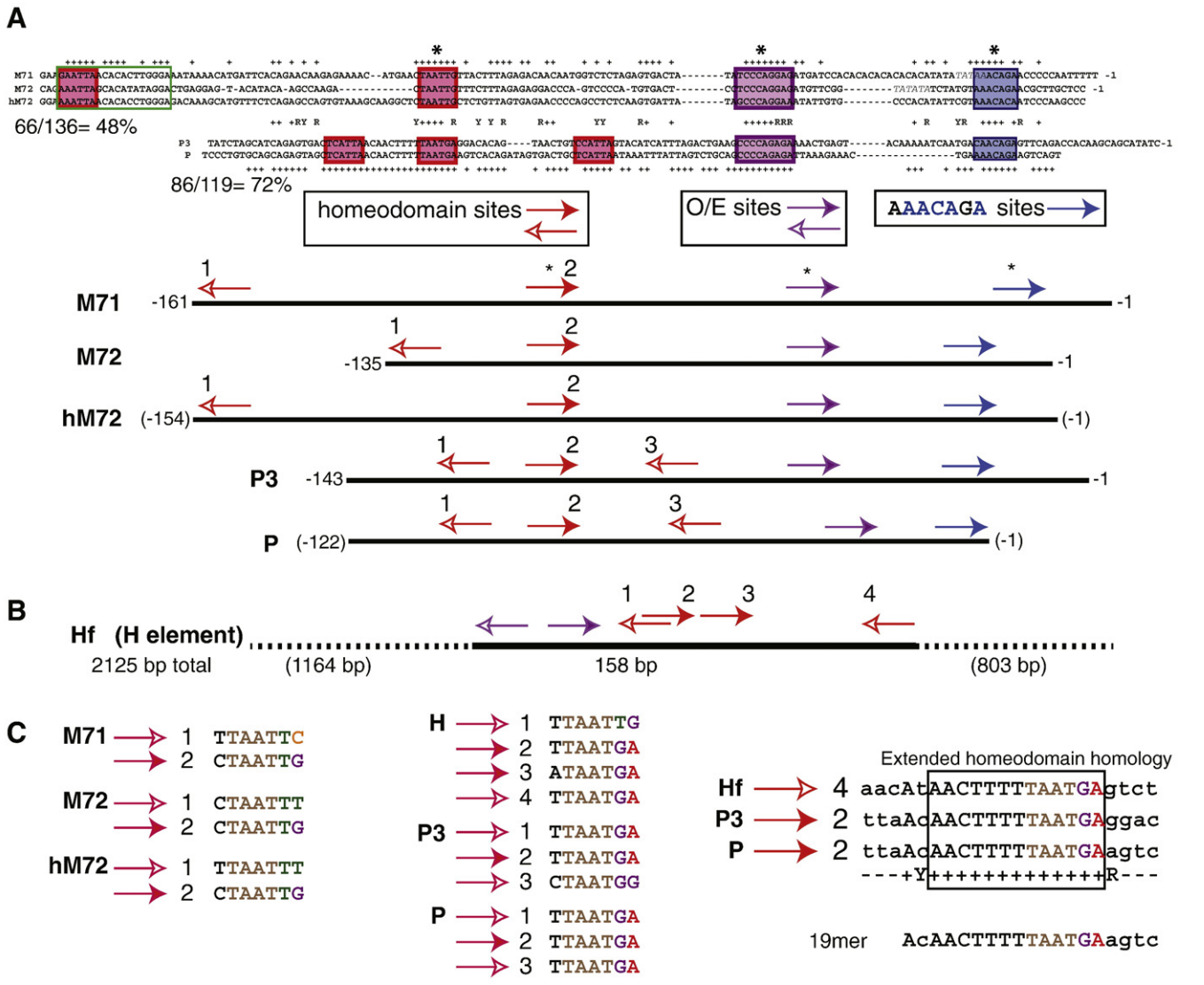


Fig. 5. Conserved sequence motifs in minimal OR promoters and the H and P elements. (A) (Top) Comparison of the sequences upstream of the TSS of mouse M71 (161 bp), mouse M72 (135 bp) and human *hM72* (154 bp) reveals four stretches of homology, which are indicated with colored, shaded boxes: two potential homeodomain (HD) binding sites (red boxes), one O/E binding site (purple box) and one site (blue box) of unknown function near the TSS. The green, non-shaded box depicts a stretch of additional sequence homology downstream of the first HD binding site (which is on the opposite strand) for *M71*, *M72* and *hM72*. The symbol + denotes identity among the three sequences. (Middle) Comparison of the sequences upstream of the TSS of *P3* (143 bp) and the corresponding site in the P element (122 bp) reveals five stretches of homology: three HD binding sites (red boxes), one O/E binding site (purple box) and the same site (blue box) of unknown function near the TSS. Sequence comparison among all five sequences is shown between the top and middle: it reveals the strongest homology in HD sites, O/E sites, and the *AAACAG/CA* site near the TSS. The symbol + denotes identity among the five sequences. *Y* is *T* or *C*, and *R* is *A* or *G*. (Bottom) Results from sequence analysis are schematized to show orientation of conserved sequences. HD sites are numbered. For reference, asterisks represent conserved sequences in *M71*, *M72* and *hM72* that we reported previously (Vassalli et al., 2002). The orientation of the motifs on either DNA strand is represented with arrows in different directions and with filled or unfilled arrowheads. (B) Sequence analysis of a central 158 bp region within the 2125 bp H element. Hf denotes a forward orientation with regard to the *MOR28* promoter in endogenous locus. The central 158 bp region is of interest, as it contains two O/E sites (purple arrows) and four HD sites (red arrows) in close proximity. These HD sites are identical to those in the core-H region described in Nishizumi et al., 2007. (C) Sequences of HD binding sites in the various promoter segments and the H and P elements. The H and P elements have three HD binding sites of the *TAATGA* type, and the *P3* promoter has two such sites. Core HD sequence *TAAT* is indicated in brown. A 19mer sequence from the P element encompassing the 13mer *AACTTTTTAATGA* that is shared with the H element and the P3 promoter, is multimerized nine times in the 9x19HD experiments.

number, orientation and types of HD binding sites along with neighboring O/E binding sites, particularly in a central 158 bp region (Figs. 5B, C).

We hypothesize that the precise sequence of the HD binding site may modulate the probability of gene expression. The HD binding sites in *M71*, *M72* and *hM72* all contain the *TAATT* pentamer sequence, but have variability at N_6. The P3 promoter and the P element contain the *TAATG* pentamer sequence, with P3 showing variability in one of its HD binding sites at N_6 (Fig. 5C). Comparison between the HD binding sites in the P element and in a central 158 bp region of the H element, which overlaps with the core-H region (Nishizumi et al., 2007), reveals three occurrences of *TAATGA* in common and an extended 13 bp perfect match (*AACTTTTTAATGA*) at one of these HD binding sites (Fig. 5C). This 13 bp sequence also occurs in the P3 promoter.

Thus, the *M71*, *M72*, *hM72*, and *P3* promoters and the P element share homologies in sequence and arrangement of conserved motifs. The H and P elements contain identical HD binding sites and a similar arrangement of motifs.

Effects of adding the H element in the MOR23 minigene

Many OR promoters contain HD binding sites upstream of their TSS suggesting a common mechanism for gene regulation (Michaloski et al., 2006; Hoppe et al., 2006). Site-directed mutagenesis of an HD

Please cite this article as: Vassalli, A., et al., Homeodomain binding motifs modulate the probability of odorant receptor gene choice in transgenic mice, Mol. Cell. Neurosci. (2010), doi:10.1016/j.mcn.2010.11.001

binding site in the *M71* promoter demonstrated its involvement in expression (Rothman et al., 2005). The exact conservation of the 13mer *AACTTTTTAATGA* encompassing a HD binding site in the P3 promoter and the P and H elements suggests that this site in particular is involved in OR gene expression.

An alternative strategy to study the role of HD binding sites, instead of deleting or mutating them, is to test whether the addition of an intact HD binding site(s) affects the numbers of transgene-expressing founders and lines, and the numbers of transgene-expressing OSNs. We chose to test such effects on a minigene rather than a promoter transgene, so that we could also evaluate axonal coalescence into glomeruli, which is the strictest criterion to compare expression patterns (Vassalli et al., 2002). In order to limit the effects from intronic sequences, we used as basis for our transgenic strategy the shortest OR minigene (TgΔ-MOR23) that recapitulates the endogenous *MOR23* expression pattern (Vassalli et al., 2002). This transgene contains 405 bp upstream of the *MOR23* TSS and an intact *MOR23* coding sequence, thereby enabling coalescence of axons into the same glomeruli as axons of OSNs that express the endogenous *MOR23* locus (Vassalli et al., 2002). The difference with TgΔ-MOR23 (Vassalli et al., 2002) is that we now inserted *IRES-tauGFP* instead of *IRES-tauLacZ* after the coding sequence (minigene TgΔ-MOR23-GFP). Fluorescent cells are observed in the dorsal MOE, and labeled axons converge to where MOR23 glomeruli are found normally (Fig. 6A).

In other studies, addition of the H element to a *MOR28* minigene resulted in a very high number of OSNs expressing the transgene (Serizawa et al., 2003, 2006). We tested whether a similar effect of the H element is also observed on our *MOR23* minigene. We inserted the full 2.1 kb H element upstream of TgΔ-MOR23-GFP, constructing transgene Hf-TgΔ-MOR23-GFP (Fig. 5B). ('f' denotes the forward orientation of the H element with regard to the *MOR28* promoter in the endogenous locus.) Most founders and lines show only modest increases in the numbers of labeled cells in the MOE, and GFP-tagged axons coalesce with MOR23-LacZ axons (Fig. 7 and data not shown). However, one founder and one mouse line exhibit a very large number of transgene-expressing OSNs in the dorsal region of the MOE (Figs. 6B, 7), with the majority of axons coalescing into a large glomerulus, which is however different from the endogenous MOR23 glomerulus (Fig. 6B).

Thus, in the particular configuration that we tested, the H element does not have the same dramatic and consistent effects on *MOR23* minigene expression as on *MOR28* (Serizawa et al., 2003). It is possible that the H element has a stronger effect on a ventrally expressed OR gene such as *MOR28* than a dorsally expressed OR gene such as *MOR23*. Alternatively, inhibitory sequences within the 2.1 kb H element may mask the effect of stimulatory sequences on *MOR23* minigene expression — a possibility that will be dealt with in the next and final series of constructs.

Effects on the MOR23 minigene of a multimerized 19mer containing an intact HD binding site

We hypothesized that the exceptionally strong effect of the H and P elements on gene expression could come from the triplet of *TAATGA* HD binding sites (Fig. 5C) that are in common between the H element (Fuss et al., 2007; Nishizumi et al., 2007; Serizawa et al., 2003) and the P element (Bozza et al., 2009). In addition, *TAATGA* is a preferential binding site for LHX2 (Berger et al., 2008). A direct test of the effect of HD binding sites is to multimerize and insert them upstream of the TgΔ-MOR23-GFP minigene. We chose the HD binding sequence *TAATGA* within the extended 13 bp *AACTTTTTAATGA*, because this sequence occurs as a perfect match in the P3 promoter and in both in the H and P elements. We hypothesized that *TAATGA* may work better as a multimer (at least a triplet), but we were concerned that concatenating the 13 bp may cause steric hindrance for HD protein binding. Thus, we added additional nucleotides *ac* and *agtc* respectively 5′ and 3′ to the HD site, giving rise to the 19mer sequence, *acAACTTTTTAATGAagtc* (Fig. 5C). This 19mer is from the P element. Our cloning strategy enabled us to concatenate this 19mer nine times (see Experimental methods), and we refer to it as 9xHD.

Many of the 15 founders and mouse lines with the 9xHD-TgΔ-MOR23-GFP transgene display a massive increase in the number of labeled MOE cells in a dorsal region of the MOE (Figs. 6C, D, 7). Out of a total of 24 founders and lines for 9xHD-TgΔ-MOR23-GFP or for this transgene coinjected with a differentially marked transgene (see below), all but one show expression of GFP (Fig. 7, summary; Fisher's exact test $p = 0.004$ compared to without 9xHD). The 9xHD sequence does not generally alter the zonal specificity of the MOR23 minigene: the distribution in the dorsal MOE is preserved in all 9xHD-TgΔ-MOR23-GFP and 9xmHD-TgΔ-MOR23-GFP founders and lines. Some lines show an additional narrow ventral domain of expression (data not shown).

To explore whether a single OR transgene copy is expressed in the 9xHD transgenes in spite of the very high number of transgene-expressing OSNs, we employed a second fluorescent marker IRES-taumRFP1 or IRES-taumCherry, constructing transgenes 9xHD-TgΔ-MOR23-RFP and 9xHD-TgΔ-MOR23-Cherry respectively (Fig. 6D). Coinjection of the GFP-marked minigene with the corresponding RFP- or Cherry-marked minigenes also results in founders and lines with high numbers of labeled MOE cells (Figs. 6D, 7). Importantly, nearly all labeled cells exhibit red or green fluorescence on a mutually exclusive basis (Fig. 6D and inset; 9x19RFP8: $n = 345$ red only, $n = 284$ green only, $n = 3$ red + green), demonstrating that ~99% of transgene-expressing OSNs express only one of the two types of transgenic constructs (Serizawa et al., 2000). By crossing lines of the 9xHD-TgΔ-MOR23 minigenes that have high numbers of labeled cells (lines 9x19-III, IV, VI; 9x19RFP8, 9; and 9x19Cherry13) to the corresponding gene-targeted line MOR23-IRES-taulacZ, we confirm that the two types of labeled axons co-coalesce into the same glomeruli (Figs. 6F–F'" and data not shown). Thus, these minigenes are expressed in a very similar way as the endogenous locus, but in many more OSNs.

Effects on the MOR23 minigene of a multimerized 19mer containing a mutant HD binding site

To rule out that this 19mer sequence merely functions as an insulator rendering the *MOR23* promoter less sensitive to inhibitory effects of flanking genomic sequences at the transgene integration sites, we introduced two *A*-to-*G* substitutions in the 19mer sequence (mHD for mutant HD) that disrupt the HD core binding sequence: *ACAACTTTTTGGTGAAGTC*. Importantly, founders and lines for the 9xmHD-TgΔ-MOR23-GFP minigene do not show the increased numbers of labeled MOE cells that founders and lines with 9xHD transgenes show (Figs. 6E and 7; Fisher's exact test $p = 0.028$). None of the 24 founders and lines show high numbers of labeled cells (>300 cells on turbinates), and ten exhibit no expression (Fig. 7; Fisher's exact test for non-expressors $p = 0.017$ compared to 9xHD). Axons from the 9xmHD-TgΔ-MOR23-GFP-labeled cells also co-coalesce with MOR23-LacZ labeled axons (Figs. 6G, G').

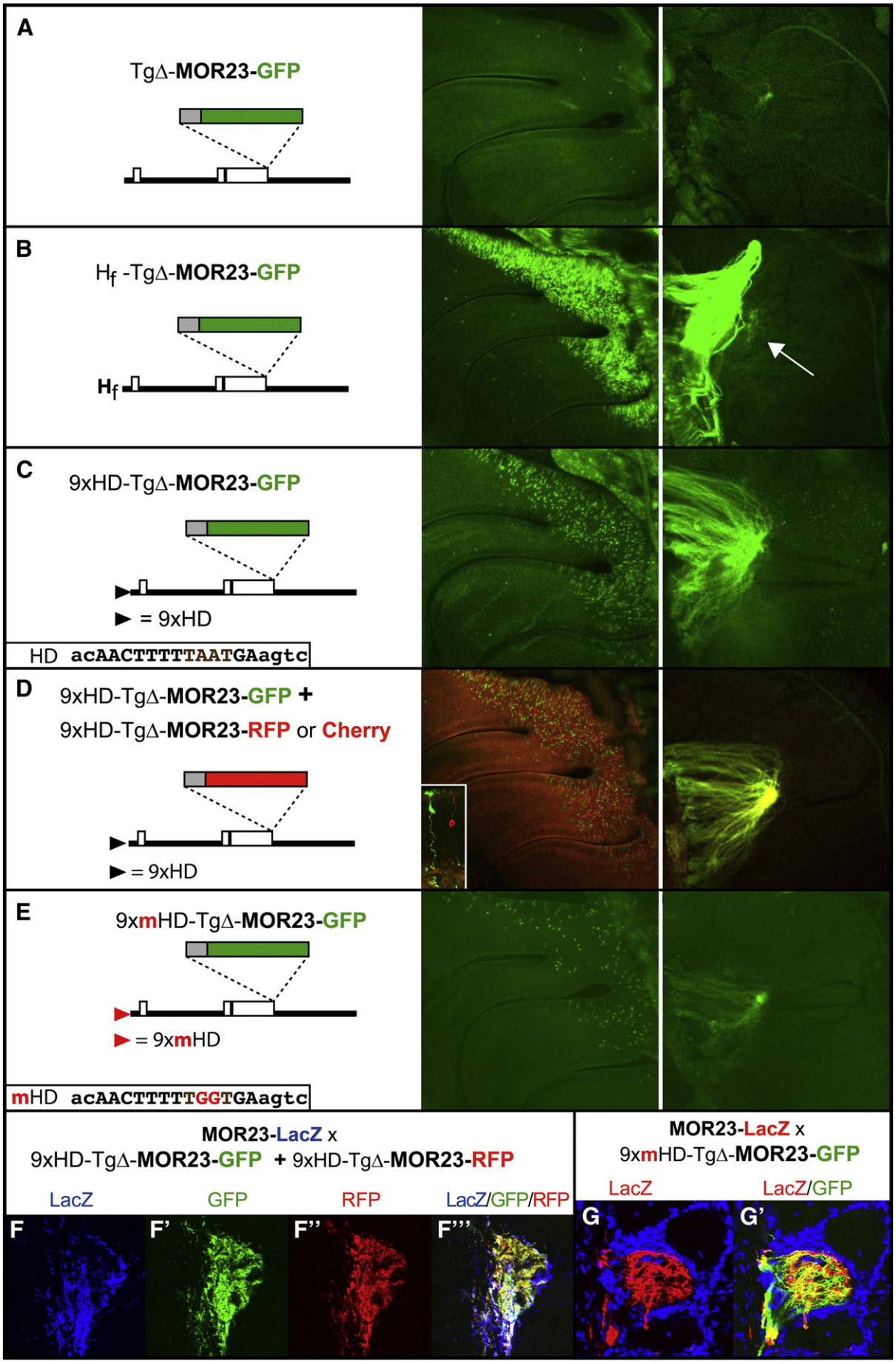
Thus, the drastic effect of 9xHD on expression of a MOR23 minigene is only observed if the HD binding site is intact, strongly implicating this site in OR gene choice.

Discussion

Promoter transgenes behave as ΔOR alleles

A series of studies have reported that OSNs that express either a transgene (Serizawa et al., 2003) or a gene-targeted allele (Bozza et al., 2009; Feinstein et al., 2004; Grosmaitre et al., 2009; Lewcock and Reed, 2004; Shykind et al., 2004) that carries a deletion or crippling mutation in the coding sequence of an OR, often coexpress another endogenous OR locus with an intact and functional OR coding sequence. The design of these knockout or deletion transgenes and

Please cite this article as: Vassalli, A., et al., Homeodomain binding motifs modulate the probability of odorant receptor gene choice in transgenic mice, Mol. Cell. Neurosci. (2010), doi:10.1016/j.mcn.2010.11.001



Please cite this article as: Vassalli, A., et al., Homeodomain binding motifs modulate the probability of odorant receptor gene choice in transgenic mice, Mol. Cell. Neurosci. (2010), doi:10.1016/j.mcn.2010.11.001

alleles is referred to as ΔOR. The coexpressed endogenous OR locus is presumably chosen at random — before, around the same time, or after ΔOR expression. The resulting phenotype is one of diffuse projection of OSN axons to a broad subdomain of the olfactory bulb. Each of these axons presumably innervates the glomerulus that is appropriate for the OR that is coexpressed by the OSN from an endogenous locus.

A popular explanation for the co-expression of a functional OR locus in ΔOR-expressing OSNs is the absence of some sort of negative feedback. An OSN would be able to sense whether it expresses a functional OR protein from the ΔOR locus or transgene. In the absence of a functional OR protein, the OSN would then sometimes proceed to a second attempt of activating an OR promoter elsewhere in the genome (Fuss and Ray, 2009; Imai and Sakano, 2009). During normal differentiation in wild-type mice, an OSN would sense a signal that emanates from the expression of a functional OR protein, and not proceed to activate another OR locus. This negative feedback is thought to contribute to the expression of one OR gene per OSN. There are no candidates for proteins or mechanisms underlying this negative feedback.

The phenotype of diffuse axonal projections for the promoter transgenes described in this study is very similar to the previously reported ΔOR phenotypes. Labeled axons do not coalesce into one or few discrete glomeruli per bulb, instead they project diffusely to a broad subdomain of the bulb. An extreme case is seen with the Tg-P(122,195)-LacZpA (also known as P-LacZ-Tg) mice, in which the P element drives expression of tau-β-galactosidase in ~10% of OSNs, which project their axons to a large majority of glomeruli in the bulb except for a dorsal domain (Bozza et al., 2009). With these P-LacZ-Tg mice, we have shown that on average ~9% of OSNs containing mRNA for a given class II OR coexpress the P-LacZ-Tg, and ~0.09% (100× less frequently) for a given class I OR (Bozza et al., 2009). Coexpression of endogenous OR loci with our promoter transgenes remains to be characterized.

Conserved motifs in OR promoter sequences

Our OR minigenes have helped to identify highly conserved motifs near the TSS of several mouse OR genes (Vassalli et al., 2002; Rothman et al., 2005). Other studies, which were not based on experimental evidence of expression, also came up with the same conserved motifs (Hoppe et al., 2006; Michaloski et al., 2006). Many of these motifs appear to be HD and O/E binding sites. But these studies could not rule out a role for other sequences in the minigenes, which contain between 2.2 and 9 kb of genomic DNA. Furthermore, there is some evidence for a role of the OR coding sequence itself in the regulation of OR gene expression (Nguyen et al., 2007, 2010).

Here, we have shown that mosaic, OSN-specific expression can be conveyed by very small sequences from OR loci in promoter-*taulacZ* transgenes that are devoid of an OR coding sequence: a mere ~300 bp is sufficient. Sequence comparisons between the mouse *M71*, mouse *M72* and *hM72* minimal promoters reveals 49% identity (66/136 bp) in the sequence upstream of the TSS. The homology between the *P3* minimal promoter and the P element in the region upstream of the TSS is 72% identity (86/119 bp). When these five sequences are aligned, only three small stretches of homology are found within the 71–88 bp upstream of the TSS: one HD binding site, one O/E binding site, and another motif *AAACAG/CA* just upstream of the TSS, near where RNA polymerase II would be expected to bind.

Two lines of evidence from our promoter transgenes further support a role of the conserved HD and O/E binding sites in OR gene expression. First, deletion of one O/E binding site of the *MOR23* promoter, in promoter transgene Tg-MOR23(148,40)-LacZpA, diminishes transgene expression drastically. Second, five nucleotide substitutions in both the HD and O/E sites in the *M71* promoter, in promoter transgene Tg-M71(mHD + O/E 161, 176)-LacZpA, preclude transgene expression, revealing only a few and faintly labeled OSNs. The identical mHD + O/E mutations introduced in the *M71* minigene with a similar upstream region (Rothman et al., 2005) give rise to a few and faintly labeled OSNs as well.

We hypothesize that an assembly of factors at the HD and O/E binding sites represents a crucial aspect of OR gene choice, in the activation and/or the maintenance of transcription. The absence of class II OR gene expression in mice mutant for the HD protein Lhx2 (Hirota et al., 2007) argues for its involvement in class II OR gene choice and/or maturation of OSNs that express class II OR genes. O/E proteins are four transcription factors of the repeated helix–loop–helix (rHLH) type and are highly expressed in the MOE. Physical coupling of bHLH and LIM-homeodomain transcription factors on promoters has been proposed to control the differentiation of motor neurons (Lee and Pfaff, 2003), and could be used also for OR genes in OSNs.

Transgenic expression patterns

The promoter transgenes do not always reproduce accurately the spatial expression patterns in the MOE. However, neither is the expression of other OR transgenes in the literature perfect in all cases, even with very large transgenes such as those based on yeast or bacterial artificial chromosomes (Nakatani et al., 2003; Qasba and Reed, 1998; Rothman et al., 2005; Serizawa et al., 2000; Vassalli et al., 2002; Zhang et al., 2007). Specifically, the *M72* and *P3* promoter transgenes tend to be expressed in a more ventral region of the MOE compared to their endogenous counterparts. The *M72* promoter sequence is shorter than *M71* and *hM72* by 13/61 bp upstream of a conserved HD binding site.

Fig. 6. Increase in the number of OSNs expressing a *MOR23* minigene by insertion of a 9xHD multimer upstream. (A) TgΔ-MOR23-GFP. This *MOR23* minigene (left) contains the *MOR23* coding sequence, and is similar in design to the TgΔ minigene described in Vassalli et al., 2002. Intron 2 is removed, and *IRES* (grey box) followed by *tauGFP* (green box) inserted into a *PacI* site after the stop codon of the *MOR23* coding sequence. Wholemount fluorescence from a transgenic founder (23DG-13) reveals dorsal expression in the MOE (left image), and coalescence of axons into a glomerulus with a position in the olfactory bulb that is similar to that of the endogenous MOR23 glomerulus (right image). (B) Hf-TgΔ-MOR23-GFP. One copy of the H element is inserted upstream of the TgΔ-MOR23-GFP minigene that is described in (A). This transgene contains the *MOR23* coding sequence. Wholemount fluorescence from a transgenic founder (Hf-27) reveals greatly increased numbers of labeled cells in the MOE (left image). Most of the labeled axons coalesce in a very large glomerulus, which is however not the endogenous MOR23 glomerulus, but more dorsal (right image). A minority of the labeled axons project to where the endogenous MOR23 glomerulus resides (arrow). (C) 9xHD-TgΔ-MOR23-GFP. A 19mer that is referred to as HD (sequence in inset) and encompasses the conserved 13mer *AACTTTTTAATGA*, which contains a HD binding site, is concatenated 9 times (9xHD, represented with black triangle) and inserted upstream of the TgΔ-MOR23-GFP minigene that is described in (A). This transgene contains the *MOR23* coding sequence. Wholemount fluorescence from a transgenic founder (9x19-11) reveals greatly increased numbers of labeled cells in the dorsal MOE (left image). In the bulb (right), labeled axons project to the endogenous MOR23 glomerulus (as confirmed in crosses to MOR23-IRES-tauLacZ with three lines of this construct: 9x19-III, IV, VI, data not shown). (D) 9xHD-TgΔ-MOR23-GFP minigene coinjected with either 9xHD-TgΔ-MOR23-RFP or 9xHD-TgΔ-MOR23-Cherry. All three transgenes contain the *MOR23* coding sequence. Wholemount fluorescence from a mouse of line RFP8 reveals red- or green-fluorescent cells in the dorsal MOE (left image). In the bulb (right), red and green axons co-coalesce into the endogenous MOR23 glomerulus (as confirmed in crosses to MOR23-IRES-tauLacZ, see also (F'–F")). Inset shows green- or red-fluorescent cells; double-labeled cells are rarely observed. (E) 9xmHD-TgΔ-MOR23-GFP. A mutant form of the 19mer termed mHD (inset) is concatenated 9 times (9xmHD) and inserted upstream of the TgΔ-MOR23-GFP minigene that is described in (A). The two *A*-to-*G* nucleotide substitutions in mHD affect the conserved *TAAT* core sequence that is common to all HD sites. This transgene contains the *MOR23* coding sequence. Wholemount fluorescence from a mouse of line 11 reveals normal numbers of labeled cells in the dorsal MOE (left image). In the bulb, labeled axons coalesce into the endogenous MOR23 glomerulus (as confirmed in crosses to MOR23-IRES-tauLacZ, see also (G) and (G')). (F,F',F",F"') Section through the bulb from a 9xHD-TgΔ-MOR23-GFP+9x HD-TgΔ-MOR23-RFP mouse of line RFP8 crossed to MOR23-IRES-tauLacZ (MOR23-LacZ). The endogenous MOR23 glomerulus (blue, antibody to β-galactosidase, F) is also innervated by GFP-labeled axons (green, F') and RFP-labeled axons (red, F"). The triple overlay in F"' shows convergence of the three types of axons to the same glomerulus. (G, G') Section through the bulb from a 9xmHD-TgΔ-MOR23-GFP mouse of line 11 crossed to MOR23-IRES-tauLacZ. The endogenous MOR23 glomerulus is labeled red with an antibody to β-galactosidase (G) and is coinnervated by GFP-labeled axons (G'). Glomeruli are surrounded by periglomerular cells (blue).

Please cite this article as: Vassalli, A., et al., Homeodomain binding motifs modulate the probability of odorant receptor gene choice in transgenic mice, Mol. Cell. Neurosci. (2010), doi:10.1016/j.mcn.2010.11.001

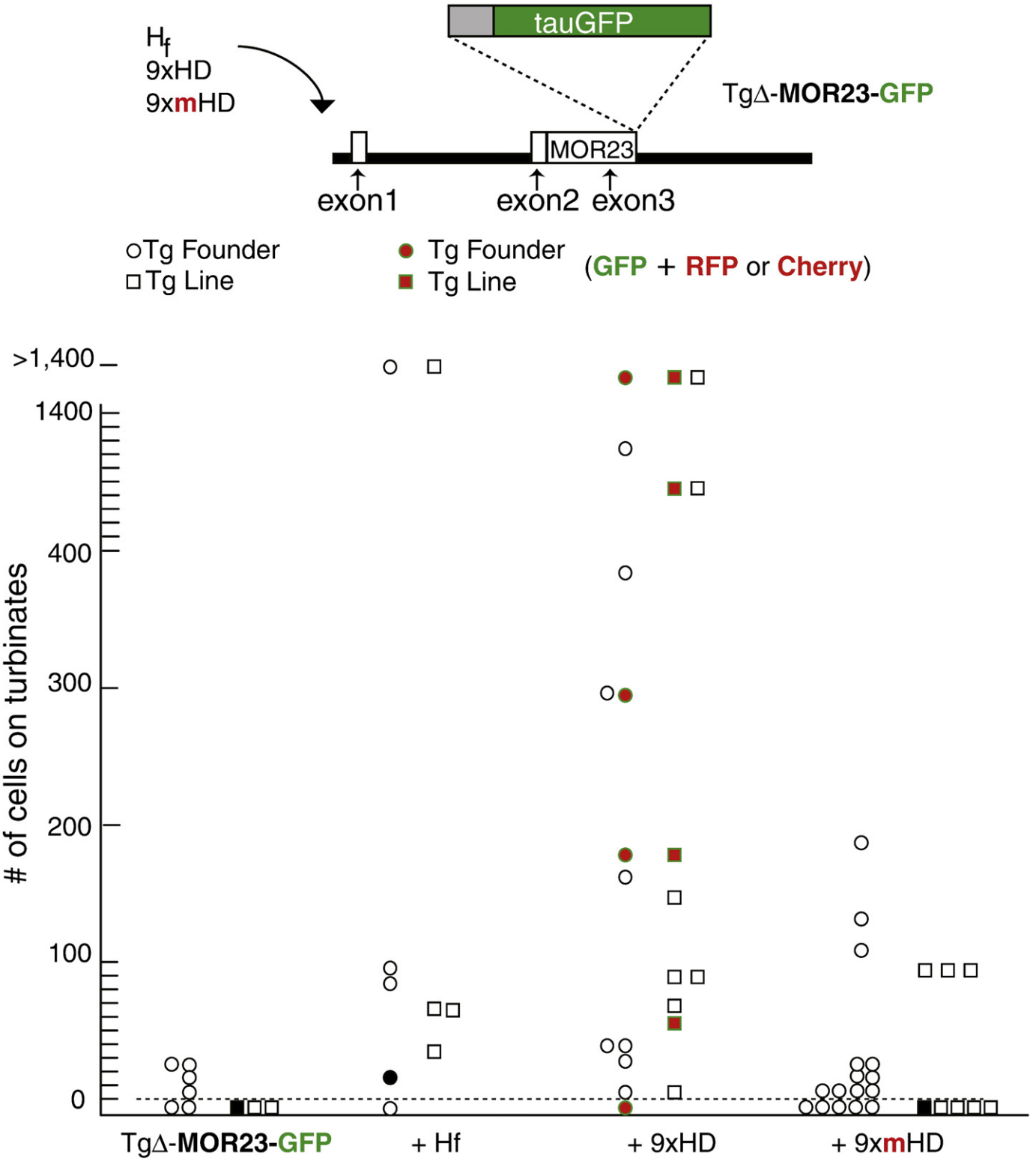


Fig. 7. Cell counts for the various TgΔ-MOR23-GFP minigenes. (Top) The TgΔ-MOR23-GFP minigene consists of a 405 bp sequence upstream of the first non-coding exon, then nontranslated exons 1 and 2, followed by exon 3 (containing the *MOR23* coding sequence), and finally 1.7 kb of 3′ sequence. The intron between exon 1 and exon 2 has been removed. *IRES* (grey box) followed by *tauGFP* (green box) is inserted into a *PacI* site after the stop codon of the *MOR23* coding sequence (Vassalli et al., 2002). Variations of this minigene come in three forms: + Hf, + 9xHD or + 9xmHD. Counts of labeled cells on olfactory turbinates (in groups of 10 s) are for transgenic founders (circles) or lines (squares). Cell counts of mice derived from coinjections of transgenes containing GFP and RFP, or GFP and Cherry, are also included for transgenic founders (red-filled circles with a green outline) and lines (red-filled squares with a green outline). Black-filled squares or circles represent lines or founders with a low expression level per cell and in cells throughout the MOE, a likely consequence of genomic position effects due to the integration site. One 9x19HD founder had mosaic pan-neuronal expression combined with a dorsal pattern of expression, which was not included in the figure because of difficulty in counting.

Please cite this article as: Vassalli, A., et al., Homeodomain binding motifs modulate the probability of odorant receptor gene choice in transgenic mice, Mol. Cell. Neurosci. (2010), doi:10.1016/j.mcn.2010.11.001

The full control region may have additional sequences with positive or negative regulatory regions. The *P3* promoter transgene contains 27 bp fewer upstream of the HD binding site compared to the *M71* promoter family, and may also miss some regulatory sequences. These ventralized patterns are not only observed with minigenes and promoter transgenes, but also with gene-targeted mutations in the *M71* promoter (Rothman et al., 2005). Our interpretation is that OR gene choice may also be regulated by suppression or negative regulation in certain regions of the MOE (Rothman et al., 2005); when these sequences are missing or incomplete, ectopic expression can occur.

The 337 bp genomic sequence that we have termed the P element (Bozza et al., 2009) was selected based on homology with the experimentally validated *P3* promoter sequence. We have no evidence that the P element acts a promoter for an OR gene in the cluster, but it can act as a promoter for *tauLacZ* in a transgenic context. Unexpectedly this 337 bp sequence behaves very differently from the other eight short promoters tested in this series of transgenic constructs: up to 10% of OSNs express P-LacZ-Tg (Bozza et al., 2009). The occurrence of HD and O/E binding sites in the P element is consistent with their role in OR gene expression.

Further experiments

These promoter segments of ~300 bp can be reduced further and dissected in our assay of promoter-*tauLacZ* transgenes. We retained 5′ nontranslated exon sequences in order to maintain stability of the taulacZ mRNAs. For 55% of ~200 ORs analyzed, the endogenous transcripts contain open reading frames of at least ten codons upstream of the presumptive AUG start codon within 5′ nontranslated exons (Michaloski et al., 2006). Whether these short, upstream open reading frames have a regulatory function, can be tested by deletion or by addition of heterologous or artificial 5′ nontranslated exons. The most critical sequence could be the region encompassing the HD and O/E binding sites. These two types of binding sites are the only conserved motifs identified among ~200 putative OR promoters (Michaloski et al., 2006). The third motif *AAACAG/CA* located just upstream of the TSS is not conserved when many ORs are examined.

H and P elements

The 2.1 kb H element was discovered by its high sequence homology between human and mouse genomic sequence in the *MOR28* gene cluster (Serizawa et al., 2000, 2003). Its presence is required for expression of three of seven adjacent OR genes including *MOR28* (Fuss et al., 2007), and a core region has been identified (Nishizumi et al., 2007). When the H element is placed 5 kb upstream of the TSS in an otherwise non-functional *MOR28* minigene, transgene expression is not only restored, but the numbers of expressing OSNs are extremely high (Serizawa et al., 2003). It appears that the relocation of the H element ~50 kb closer to the *MOR28* TSS increases the probability of expression. Similarly, our P-LacZ-Tg promoter transgene is expressed in ~10% of OSNs. We suggest that the H and P elements have evolved high-affinity binding sites, enabling them to operate on OR promoters from their remote locations. Evolution of strong, long-distance regulatory elements within and nearby OR gene clusters would favor the maintenance of OR function.

A comparison of the sequence of the mouse H and P elements reveals a set of three HD binding sites in close proximity to each other, and an associated O/E site. The three *TAATGA* HD binding sites in the P element are within a 45 bp stretch of DNA, or 75 bp if the O/E site is included. The *TAATGA* HD binding sites of the H element are within 92 bp of each other, and within 125 bp of the first O/E site. A 13 bp perfect match between the two elements includes one of the HD binding sites. This sequence alone cannot explain the strength of the activity of the P element in P-LacZ-Tg, because it is also conserved in the P3 promoter, which is expressed in much fewer cells (Feinstein and Mombaerts, 2004). Interestingly, the P3 promoter contains only one additional *TAATGA* HD site, but the P and H elements contain two such additional sites.

We tested the notion of tightly linked HD binding sites cooperating to achieve high numbers of expressing OSNs by multimerizing a 19mer from the P element. The MOR23 coding sequence was included in these transgenic constructs in order to evaluate patterns of axonal coalescence. The numbers of expressing transgenic lines and of expressing OSNs are increased dramatically with the 9xHD constructs. Despite the elevated numbers of transgene-expressing OSNs, the transgene appears to expressed from a single copy in most cells. The involvement of a functional HD binding site in this effect is strongly suggested by the absence of such a boost with the mutant 9xmHD constructs. Axons of OSNs expressing MOR23 from the 9xHD transgene co-converge into the same glomeruli as axons of OSNs expressing MOR23 from the endogenous locus, suggesting that the level of MOR23 expression per OSN is normal.

We hypothesize that this multimerized 9xHD sequence increases the probability of transgenic OR gene expression by enabling strong cooperative interactions among HD binding proteins. Similarly, the triplets of *TAATGA* HD binding sites in the H and P elements together would constitute an exceptionally high-affinity binding site for HD binding proteins. This assembly of HD binding proteins and perhaps other factors would then serve as a binary on-off switch for adjacent OR genes in the cluster: the 'choice' of one OR gene for high-level transcription. Likewise, we hypothesize that HD binding sites within OR promoters themselves modulate the probability of OR gene choice. Mechanistically, HD binding motifs could modulate the probability of OR gene choice by attracting polymerase assembly factors, enhancing transcription initiation, or stabilizing transcription elongation.

Conclusion

The total number of OSNs that express a given OR in a mouse varies over two orders of magnitude (Mombaerts, 2004). We speculate that this number may be critically determined by the number and the strength of the HD binding sites in regulatory elements such as H and P elements, by the number and strength of the HD binding sites in OR promoters, and by the distance between these regulatory elements and OR promoters.

Experimental methods

Promoter transgenes

The *Pac*I–*Eco*RI–IRES–*Nco*I–tauLacZ–*Xba*I fragment (Mombaerts et al., 1996) was ligated with a 556 nt *Xba*I-rabbit globin pA–*Eco*RV–*Pac*I–*Xba*I fragment, thereby generating the *Pac*I cassette IRES–tauLacZpA. Next, the *Eco*RI–IRES–*Nco*I fragment was replaced with OR (X,Y) promoter segments of the type *Eco*RI–*Eco*RV–OR (X,Y)–*Kozak*–*Nco*I that were generated by PCR, resulting in transgenes flanked by *Pac*I restriction sites: OR (X,Y)–tauLacZpA. (X,Y) indicate the bp position upstream and downstream, respectively, of the TSS as defined by 5′ RACE. The *Nco*I site contains the start codon *ATG* for methionine. The nucleotides *GCCA* were inserted upstream of the *Nco*I site (*CCATGG*) to mimic a Kozak sequence, insuring more efficient translation of the β-galactosidase protein. The mouse promoter fragments are from the 129 genetic background. The hM72 (154,146) fragment is from BAC 210L3 (Genome Systems). All promoter fragments were sequenced. These plasmids are publicly available from Addgene (Cambridge, MA, USA).

MOR23 minigenes

The TgΔ-MOR23 minigene is a *Sac*I–*Nhe*I fragment, but with intron 2 fused with intron 3 (Vassalli et al., 2002). A *Pac*I site is located 3

Please cite this article as: Vassalli, A., et al., Homeodomain binding motifs modulate the probability of odorant receptor gene choice in transgenic mice, Mol. Cell. Neurosci. (2010), doi:10.1016/j.mcn.2010.11.001

nucleotides after the *MOR23* stop codon, where the *Pac*I cassette IRES-tauGFP (-GFP) was inserted generating TgΔ-MOR23-GFP. The H element was amplified by PCR from 129 mouse genomic DNA origin including 5′ and 3′ *Sac*I restriction sites and cloned into the *Sac*I site of the TgΔ-MOR23-GFP minigene generating Hf-TgΔ-MOR23-GFP. The identified conserved sequences in the H element are identical in the C57BL/6J strain.

The transgenes 9xHD-TgΔ-MOR23-GFP and 9xmHD-TgΔ-MOR23-GFP were constructed as follows. A pair of unphosphorylated oligos containing three times the HD-containing sequence *ACAACTTTTTAATGAAGTC* or three times the mutant sequence *ACAACTTTTTGGTGAAGTC* from 129 mouse genomic DNA origin were annealed and ligated into the *Sac*I digested TgΔ-MOR23-GFP minigene. The mutation (AA to GG) abolishes the A_2A_3 core of the $T_1A_2A_3T_4$ HD recognition site; other pairs of *AA* nucleotides are retained in the oligos. Clones were screened for a triplet insertion in the same orientation (same as HD site 2 in the P element, Fig. 5C) and sequenced. Oligos:

3xHD-1:
CACAACTTTTTAATGAAGTCACAACTTTTTAATGAAGTCACAACTTTTTAATGAAGTCGAGCT
3xHD-2:
CGACTTCATTAAAAAGTTGTGACTTCATTAAAAAGTTGTGACTTCATTAAAAAGTTGTGAGCT
3x*m*HD-1:
CACAACTTTTTGGTGAAGTCACAACTTTTTGGTGAAGTCACAACTTTTTGGTGAAGTCGAGCT
3x*m*HD-2:
CGACTTCACCAAAAAGTTGTGACTTCACCAAAAAGTTGTGACTTCACCAAAAAGTTGTGAGCT

*Pac*I cassettes containing IRES-taumRFP1 or IRES-taumCherry were ligated into the 9xHD-TgΔ-MOR23 minigene.

Transgenic mice

Plasmid inserts were released for pronuclear microinjection into F2 (C57BL/6J × CBA) zygotes, with *Pac*I for the promoter transgenes and with *Pme*I–*Nhe*I for the MOR23 minigenes. Founders included in the analysis are distinct from founders bred with C57BL/6J mice to transmit the transgene through the germline and establish lines. For each transgenic construct that gives expression, one representative line has been cryopreserved as sperm at The Jackson Laboratory (Bar Harbor, Maine, USA), and is publicly available.

Sequences of promoter transgenes

For unambiguous identification of the genomic segments from OR loci that are included in the promoter transgenes, below are the sequences of the variable, relevant parts. Under the name of the transgene is the Addgene Number ID for the plasmid. Sequences that are underlined represent a *Pac*I restriction site; italics, *EcoR*I restriction site; underlined italics, *EcoR*V restriction site; forward slashes, boundary between restriction sites and promoter sequence; capitals underlined, Kozak sequence; capitals italics, *NcoI* restriction site encompassing *ATG* start codon. The length of the sequence between the forward slashes corresponds to the sum of the numbers x and y. Under each sequence is listed our line number, the JR number assigned by The Jackson Laboratory, and the official name of the line in a mixed C57BL/6 × CBA background that is publicly available from The Jackson Laboratory.

Tg-M71(161,176)-LacZpA
Addgene 15605
ttaattaa*gaattc*gatatc/gaagaattaacacacttgggaaataaaacatgattcacagaacaagagaaaacatgaactaattgttactttagagacaacaatggtctctagagtgactatatcccaggagatgatccacacacacacacatatatataaacagaaccccaatttttagtcatattatatcacctttccaaatgacaatctccttttccggactagactgggctctttaagtaaattttgtggtgtatacacatctgcatacttctagaatttactgaatattgaagccacagacattgcctgaagtttcatttgagacgctgtttcatttccttcagcacag/GCCA*CCATGG*
line IV, JR#7972, B6;CBA-Tg(Olfr151-tauLacZ)4Mom/MomJ

Tg-M71(161mHD + O/E,176)-LacZpA
Addgene 15606
ttaattaa*gaattc*gatatc/gaagaattaacacacttgggaaataaaacatgattcacagaacaagagaaaacatgaactggttgttactttagagacaacaatggtctctagagtgactatataggaggagatgatccacacacacacacatatatataaacagaaccccaatttttagtcatattatatcacctttccaaatgacaatctccttttccggactagactgggctctttaagtaaattttgtggtgtatacacatctgcatacttctagaatttactgaatattgaagccacagacattgcctgaagtttcatttgagacgctgtttcatttccttcagcacag/GCCA*CCATGG*

Tg-MOR23(148,154)-LacZpA
Addgene 15603
ttaattaa*gaattc*gatatc/aactgatgcacaacatcttcagaaacttctgaaggaagggaagaaagaaacacctgggaagcagcctctggggaatcccaaaatgtcacccccctgagataagtaaaaaccccctgtttttaataaaatcccctggtatctatttgtgacacgacttttctggttccccctgagataccaagggagctgtgaagaaaggagagccaggtgggtgtggtagacagtgttgtttccttggagatttctcttgatctctgaaaccacagaggcttcttctcctattagcaccactttcagagcctgg/GCCA*CCATGG*
line I, JR#7973, B6;CBA-Tg(Olfr16-tauLacZ)1Mom/MomJ

Tg-MOR23(148,40)-LacZpA
Addgene 15604
ttaattaa*gaattc*gatatc/aactgatgcacaacatcttcagaaacttctgaaggaagggaagaaagaaacacctgggaagcagcctctggggaatcccaaaatgtcacccccctgagataagtaaaaaccccctgtttttaataaaatcccctggtatctatttgtgacacgacttttctggttccccctgagataccaagggagctgtg/GCCA*CCATGG*

Tg-M72(135,163)-LacZpA
Addgene 15607
ttaattaa*gaattc*gatatc/cagaaattagcacatataggactgaggagtacatacaagccaagactaattgtttctttagagaggacccagtccccatgtgactcctcccaggagatgttcggtatatatctatgtaaacagaacgcttgctccactatcatgatgattccttttttctaacagtgccaggcccttgcagtaaatttagcactagatataaatccatttatattatctgttttcagcatctgctaattgttgagaccaaaggcatcccttaaagtttccagaggctcatcacttcccttccacaaag/GCCA*CCATGG*
line V4-7, JR#7974, B6;CBA-Tg(Olfr160-tauLacZ)V4-7Mom/MomJ

Tg-hM72(154,146)-LacZpA
Addgene 15608
ttaattaa*gaattc*gatatc/aggaaaattaacacacctgggagacaaagcatgtttctcagagccagtgtaaagcaaggctctaattgctctgttagtgagaaccccagcctctcaagtgattatagcccaggaaatattgtgcccacatattcgtaaacacaatcccaagcccccagtcttagtgatccatttttaatgaatttgagtaaatgtaggactcagtataaatcagtatggttcctgcattttctgattgttgaaaccaaggcagccccttaaagtttccagcagactgtgcttttatttttcctcatcacag/ GCCA*CCATGG*
line I, JR#7975, B6;CBA-Tg(OR8A1-tauLacZ)1Mom/MomJ

Tg-hM72(110,29)-LacZpA
Addgene 15609
ttaattaa*gaattc*gatatc/cagtgtaaagcaaggctctaattgctctgttagtgagaaccccagcctctcaagtgattatagcccaggaaatattgtgcccacatattcgtaaacacaatcccaagcccccagtcttagtgatccatttttaatgaa/GCCA*CCATGG*

Tg-P3(143,163)-LacZpA
Addgene 15610

Please cite this article as: Vassalli, A., et al., Homeodomain binding motifs modulate the probability of odorant receptor gene choice in transgenic mice, Mol. Cell. Neurosci. (2010), doi:10.1016/j.mcn.2010.11.001

ttaattaa*gaattcgatatc*/tatctagcatcagagtgactcattaacaacttttaatgaggacacagtaactgtccattagtacatcatttagactgaagccccagagaaaactgagtacaaaaatcaatgacaacagagttcagaccacaagcagcatatcggagtgaagagagtggactttctcagcagaagccaccaggcttttactgttgagtctccagggtctttgcaggtccccagtggctatgattctggatataaaaatagagggaccaggggctttgtgggaggaaaggacaactgagcagagtgatctgagacctg/GCCA*CCATGG*
line IV, JR#7976, B6;CBA-Tg(Olfr713-tauLacZ)4Mom/MomJ

Tg-P(122,195)-LacZpA (referred to as P-LacZ-Tg in Bozza et al., 2009)
Addgene 15611
ttaattaa*gaattcgatatc*/tccctgtgcagcagagtagctcattaacaacttttaatgaagtcacagatagtgactgctcattaataaatttattagtctgcagccccagagattaaagaaactgaaaacagaagtcagtgacaacagtgtttagccagagtagctgcatcagagtgaaaagaGggaacttctcttggcagactctgaaacttttactgttggacctctagggtccttgcaggtcccagaaactgtaattctgaaaatagaaacctgggggactaagaacttgatgggatgaaggaaaaacagagccaaatgatcccaggcttg/GCCA*CCATGG*
line 8, JR#6742, Tg(P-taulacZ)8Mom/MomJ
line 11, JR#6743, Tg(P-taulacZ)11Mom/MomJ
line 13, JR#6793, Tg(P-taulacZ)13Mom/MomJ

IDs of MOR23 minigenes and transgenic lines

TgΔ-MOR23-GFP
Addgene 15612

Hf-TgΔ-MOR23-GFP
Addgene 15613
line Hf#7, JR#7977, B6;CBA-Tg(Hf/Olfr16-tauGFP)7Mom/MomJ
line Hf#47, JR#7978, B6;CBA-Tg(Hf/Olfr16-tauGFP)47Mom/MomJ

9xHD-TgΔ-MOR23-IRES-GFP
Addgene 15614
line III, JR#7979, B6;CBA-Tg(H/Olfr16-tauGFP)3Mom/MomJ
line IV, JR#7980, B6;CBA-Tg(H/Olfr16-tauGFP)4Mom/MomJ
line VI, JR#7981, B6;CBA-Tg(H/Olfr16-tauGFP)6Mom/MomJ

9xHD-TgΔ-MOR23-IRES-taumRFP plus p9xHD-TgΔ-MOR23-IRES-GFP
Addgene 15615 for 9xHD-TgΔ-MOR23-IRES-taumRFP
line RFP8, JR#7982, B6;CBA-Tg(H/Olfr16-taumRFP,H/Olfr16-tauGFP)8Mom/MomJ
line RFP9, JR#7983, B6;CBA-Tg(H/Olfr16-taumRFP,H/Olfr16-tauGFP)9Mom/MomJ

9xHD-TgΔ-MOR23-IRES-taumCherry plus p9xHD-TgΔ-MOR23-IRES-GFP
Addgene 15616 for 9xHD-TgΔ-MOR23-IRES-taumCherry
line Cherry#11, JR#7984, B6;CBA-Tg(H/Olfr16-taumCherry,H/Olfr16-tauGFP)11Mom/MomJ
line Cherry#13, JR#7985, B6;CBA-Tg(H/Olfr16-taumCherry,H/Olfr16-tauGFP)13Mom/MomJ

9xmHD-TgΔMOR23-IRES-tauGFP
Addgene 15167
line 11, JR#7986, B6;CBA-Tg(H*/Olfr16-tauGFP)11Mom/MomJ
line 25, JR#7987, B6;CBA-Tg(H*/Olfr16-tauGFP)25Mom/MomJ

5′RACE

RNA was prepared from mice between postnatal day 15–28 using the Qiagen RNeasy Kit. Reverse transcription was carried out with the SMART RACE cDNA amplification Kit (Clontech/Takara cat. no. 63914). Oligos used for 5′RACE of the *P4* gene are as follows: 1573 (5′ *CAGAATGACTAGGTTGAAGCCGATCTCCAT* 3′), CAT135 (5′ *CCCAGCAGAGTAACCAGGTAGATGGTGAG* 3′), and CAT136 (5′ *GCTTGTGTCCTGGGCAATCAGAGTCCTC* 3′). Transgene 5′RACE analysis was with an internal LacZ oligo.

Analysis of LacZ expression

Blue precipitate for wholemount analysis was obtained with X-gal (Mombaerts et al., 1996). Mice for cryosectioning were drop-fixed for ~3 h in PLP (2% formaldehyde, 13.5 mg/ml lysine, 2.1 mg/ml sodium periodate, and 0.1 M sodium phosphate buffer at pH 7.4). Immunohistochemistry was with rabbit anti-β-galatosidase antibody (Cappel) at 1:1000 dilution, followed by Red-X-conjugated or Cy5-conjugated goat affinity purified antibodies to rabbit IgG (Jackson ImmunoResearch) at 1:200 dilution. Fluorescent sections were analyzed with a Zeiss LSM 510 confocal microscope.

Acknowledgments

We thank Krishna Amuluru for technical support; the Transgenic Facility at The Rockefeller University for generating transgenic mice; Tom Bozza for sharing of data prior to publication, and some initial cloning steps for the IRES-taumRFP1 and IRES-taumCherry cassettes; Mona Khan for organization of the cryopreservation and plasmid submission and for critical reading of the manuscript; and Bolek Zapiec for help with the figures. Grant support came from the National Institute on Deafness and Other Communication Disorders (A.V., P.F., and P.M.) and from the Max Planck Society (P.M.)

References

Berger, M.F., Badis, G., Gehrke, A.R., Talukder, S., Philippakis, A.A., Peña-Castillo, L., Alleyne, T.M., Mnaimneh, S., Botvinnik, O.B., Chan, E.T., et al., 2008. Variation in homeodomain DNA binding revealed by high-resolution analysis of sequence preferences. Cell 133, 1266–1276.
Bozza, T., Vassalli, A., Fuss, S., Zhang, J.J., Weiland, B., Pacifico, B., Feinstein, P., Mombaerts, P., 2009. Mapping of class I and class II odorant receptors to glomerular domains by two distinct types of olfactory sensory neurons in the mouse. Neuron 61, 220–233.
Buck, L., Axel, R., 1991. A novel multigene family may encode odorant receptors: a molecular basis for odor recognition. Cell 65, 175–187.
Chess, A., Simon, I., Cedar, H., Axel, R., 1994. Allelic inactivation regulates olfactory receptor gene expression. Cell 78, 823–834.
Eilers, A., Bouterfa, H., Triebe, S., Doenecke, D., 1994. Role of a distal promoter element in the S-phase control of the human H1.2 histone gene transcription. Eur. J. Biochem. 223, 567–574.
Feinstein, P., Bozza, T., Rodriguez, I., Vassalli, A., Mombaerts, P., 2004. Axon guidance of mouse olfactory sensory neurons by odorant receptors and the β2 adrenergic receptor. Cell 117, 833–846.
Feinstein, P., Mombaerts, P., 2004. A contextual model for axonal sorting into glomeruli in the mouse olfactory system. Cell 117, 817–831.
Fuss, S.H., Omura, M., Mombaerts, P., 2007. Local and cis effects of the H element on expression of odorant receptor genes in mouse. Cell 130, 373–384.
Fuss, S.H., Ray, A., 2009. Mechanisms of odorant receptor gene choice in *Drosophila* and vertebrates. Mol. Cell. Neurosci. 41, 101–112.
Grosmaitre, X., Fuss, S.H., Lee, A.C., Adipietro, K.A., Matsunami, H., Mombaerts, P., Ma, M., 2009. SR1, a mouse odorant receptor with an unusually broad response profile. J. Neurosci. 29, 14545–14552.
Hirota, J., Mombaerts, P., 2004. The LIM-homeodomain protein Lhx2 is required for complete development of mouse olfactory sensory neurons. Proc. Natl Acad. Sci. USA 101, 8751–8755.
Hirota, J., Omura, M., Mombaerts, P., 2007. Differential impact of Lhx2 deficiency on expression of class I and class II odorant receptor genes in mouse. Mol. Cell. Neurosci. 34, 679–688.
Imai, T., Sakano, H., 2009. Odorant receptor gene choice and axonal projection in the mouse olfactory system. Results Probl. Cell Differ. 47, 57–75.
Hoppe, R., Breer, H., Strotmann, J., 2006. Promoter motifs of olfactory receptor genes expressed in distinct topographic patterns. Genomics 87, 711–723.
Lane, R.P., Cutforth, T., Young, J., Athanasio, M., Friedman, C., Rowen, L., Evans, G., Axel, R., Hood, L., Trask, B.J., 2001. Genomic analysis of orthologous mouse and human olfactory receptor loci. Proc. Natl Acad. Sci. USA 98, 7390–7395.
Lee, S.K., Pfaff, S.L., 2003. Synchronization of neurogenesis and motor neuron specification by direct coupling of bHLH and homeodomain transcription factors. Neuron 38, 731–745.

Please cite this article as: Vassalli, A., et al., Homeodomain binding motifs modulate the probability of odorant receptor gene choice in transgenic mice, Mol. Cell. Neurosci. (2010), doi:10.1016/j.mcn.2010.11.001

Lewcock, J., Reed, R., 2004. A feedback mechanism regulates monoallelic odorant receptor expression. Proc. Natl Acad. Sci. USA 101, 1069–1074.
Malnic, B., Hirono, J., Sato, T., Buck, L.B., 2006. Combinatorial receptor codes for odors. Cell 96, 713–723.
Michaloski, J.S., Galante, P.A., Malnic, B., 2006. Identification of potential regulatory motifs in odorant receptor genes by analysis of promoter sequences. Genome Res. 16, 1091–1098.
Miyamichi, K., Serizawa, S., Kimura, H.M., Sakano, H., 2005. Continuous and overlapping expression domains of odorant receptor genes in the olfactory epithelium determine the dorsal/ventral positioning of glomeruli in the olfactory bulb. J. Neurosci. 25, 3586–3592.
Mombaerts, P., 2004. Odorant receptor gene choice in olfactory sensory neurons: the one receptor-one neuron hypothesis revisited. Curr. Opin. Neurobiol. 14, 31–36.
Mombaerts, P., 2006. Axonal wiring in the mouse olfactory system. Annu. Rev. Cell Dev. Biol. 22, 713–737.
Mombaerts, P., Wang, F., Dulac, C., Chao, S.K., Nemes, A., Mendelsohn, M., Edmondson, J., Axel, R., 1996. Visualizing an olfactory sensory map. Cell 87, 675–686.
Nakatani, H., Serizawa, S., Nakajima, M., Imai, T., Sakano, H., 2003. Developmental elimination of ectopic projection sites for the transgenic OR gene that has lost zone specificity in the olfactory epithelium. Eur. J. Neurosci. 18, 2425–2432.
Nguyen, M.Q., Zhou, Z., Marks, C.A., Ryba, N.J.P., Belluscio, L., 2007. Prominent roles for odorant receptor coding sequences in allelic exclusion. Cell 131, 1009–1017.
Nguyen, M.Q., Marks, C.A., Belluscio, L., Ryba, N.J.P., 2010. Early expression of odorant receptors distorts the olfactory circuitry. J. Neurosci. 30, 9271–9279.
Nishizumi, H., Kumasaka, K., Inoue, N., Nakashima, A., Sakano, H., 2007. Deletion of the core-*H* region in mice abolishes the expression of three proximal odorant receptor genes in *cis*. Proc. Natl Acad. Sci. USA 104, 20067–20072.
Oka, Y., Katada, S., Omura, M., Suwa, M., Yoshihara, Y., Touhara, K., 2006. Odorant receptor map in the mouse olfactory bulb: in vivo sensitivity and specificity of receptor-defined glomeruli. Neuron 52, 857–869.
Qasba, P., Reed, R.R., 1998. Tissue and zonal-specific expression of an olfactory receptor transgene. J. Neurosci. 18, 227–236.
Ressler, K.J., Sullivan, S.L., Buck, L.B., 1993. A zonal organization of odorant receptor gene expression in the olfactory epithelium. Cell 73, 597–609.
Ressler, K.J., Sullivan, S.L., Buck, L.B., 1994. Information coding in the olfactory system: evidence for a stereotyped and highly organized epitope map in the olfactory bulb. Cell 79, 1245–1255.
Rothman, A., Fenstein, P., Hirota, J., Mombaerts, P., 2005. The promoter of the mouse odorant receptor *M71*. Mol. Cell. Neurosci. 28, 535–546.
Serizawa, S., Ishii, T., Nakatani, H., Tsuboi, A., Nagawa, F., Asano, M., Sudo, K., Sakagami, J., Sakano, H., Ijiri, T., et al., 2000. Mutually exclusive expression of odorant receptor transgenes. Nat. Neurosci. 3, 687–693.
Serizawa, S., Miyamichi, K., Nakatani, H., Suzuki, M., Saito, M., Yoshihara, Y., Sakano, H., 2003. Negative feedback regulation ensures the one receptor-one olfactory neuron rule in mouse. Science 302, 2088–2094.
Serizawa, S., Miyamichi, K., Takeuchi, H., Yamagishi, Y., Suzuki, M., Sakano, H., 2006. A neuronal identity code for the odorant receptor-specific and activity-dependent axon sorting. Cell 127, 1057–1069.
Shykind, B.M., Rohani, S.C., O'Donnell, S., Nemes, A., Mendelsohn, M., Sun, Y., Axel, R., Barnea, G., 2004. Gene switching and the stability of odorant receptor gene choice. Cell 117, 801–815.
Strotmann, J., Conzelmann, S., Beck, A., Feinstein, P., Breer, H., Mombaerts, P., 2000. Local permutations in the glomerular array of the mouse olfactory bulb. J. Neurosci. 20, 6927–6938.
Vassalli, A., Rothman, A., Feinstein, P., Zapotocky, M., Mombaerts, P., 2002. Minigenes impart odorant receptor-specific axon guidance in the olfactory bulb. Neuron 35, 681–696.
Zhang, Y.-Q., Breer, H., Strotmann, 2007. Promoter elements governing the clustered expression pattern of odorant receptor genes. Mol. Cell. Neurosci. 36, 95–107.

Please cite this article as: Vassalli, A., et al., Homeodomain binding motifs modulate the probability of odorant receptor gene choice in transgenic mice, Mol. Cell. Neurosci. (2010), doi:10.1016/j.mcn.2010.11.001