DEVELOPMENTAL BIOLOGY **188,** 1–16 (1997)
ARTICLE NO. DB978575

Temporal and Spatial Control of Murine GATA-3 Transcription by Promoter-Proximal Regulatory Elements

Ken H. Lieuw, Guo-long Li, Yinghui Zhou, Frank Grosveld,* and James Douglas Engel[1]

*Department of Biochemistry, Molecular Biology and Cell Biology, Northwestern University, Evanston, Illinois 60208-3500; and *Department of Cell Biology and Genetics, Erasmus University, 3000 DR Rotterdam, Holland*

GATA-3 is expressed in a temporally dynamic manner and fulfills vital functions during vertebrate fetal development. Homozygous mGATA-3 mutant embryos die at midgestation, thus complicating the analysis of its contribution to the development of specific cell fates in the many tissues where it is expressed during embryogenesis. We show here that the elements controlling GATA-3 regulation can be precisely refined, using transgenic mice, to discrete *cis*-acting domains: within 6 kb surrounding the transcriptional initiation site, separate sequences were found to control the expression of mGATA-3 in early muscle masses, in a subset of PNS neurons, in the genital tubercle, and in the branchial arches. The branchial arch regulatory element is particularly robust and was refined to a discrete enhancer sequence lying between nt −2832 and −2462 from the transcription initiation site. The enhancer contains potential binding sites for many well-characterized transcription factors, suggesting that mGATA-3 transcriptional activity may be regulated by these proteins (or related family members) in the mesenchyme of the arches that contribute to formation of the jaw. These studies show that discrete regulatory elements required for the elaboration of complex developmental programs can be individually localized, suggesting that the developmentally transient expression of individual transcription factors collaboratively contributes to the temporal and spatial pattern of cellular differentiation leading to the formation of adult anatomy.

INTRODUCTION

The family of GATA transcription factors has been shown to play key roles in regulating eukaryotic development. The vertebrate GATA factor family shares a highly conserved C_4 zinc finger DNA-binding motif, while amino acid sequences outside of the zinc finger domain exhibit considerably greater divergence, and the expression profiles of the individual family members are both diverse and developmentally dynamic (Yamamoto *et al.*, 1990). GATA-1 is predominantly expressed in myeloerythroid lineage cells, while an alternatively initiated and differentially spliced form of the mRNA produces the same protein in the testis (Ito *et al.*, 1993; Tsai *et al.*, 1989). GATA-2 is more broadly expressed in hematopoietic progenitors than is GATA-1, but is also found in many nonhematopoietic cell lineages. GATA-4, -5, and -6 form a distinct subfamily that is differentially expressed in endodermally derived tissues and in the developing heart (Arceci *et al.*, 1993; Kelley *et al.*, 1993; Lavernere *et al.*, 1994; Maeda, 1994). GATA-3, first reported to be most prevalent in T lymphocytes and in the embryonic brain (Yamamoto *et al.*, 1990), has since been found to display a developmentally dynamic expression pattern in the embryo that only later becomes predominantly restricted to T cells in the adult (George *et al.*, 1994).

Before 8.5 days postcoitus (dpc) during mouse development, GATA-3 is found only in the ectoplacental cone, which later gives rise to the placenta. By 10.5 dpc, GATA-3 mRNA localizes to specific regions of the embryo proper, principally in the central nervous system (CNS) in the mesencephalon and spinal cord, in peripheral nervous system (PNS) structures such as cranial, sympathetic, and dorsal root ganglia, in the branchial arches, and in the embryonic liver, otic vesicles, and somites. As the embryo develops,

[1] To whom correspondence should be addressed. Fax: (847) 467-2152. E-mail: d-engel@nwu.edu.

0012-1606/97 $25.00

the expression profile of GATA-3 changes in these differentiating organs and tissues. By 14.5 dpc, GATA-3 is expressed in the outer layers of the mesencephalon as well as in more clearly defined regions of the CNS (e.g., the myelencephalon and pons and in the spinal cord), while specific ganglia in the PNS (e.g., trigeminal ganglia and the facial acoustic complex) persistently express GATA-3. Less abundant GATA-3 expression is also found in the embryonic liver, kidney, adrenal medulla, vomeronasal organ, and thymic rudiment at this stage of embryonic development. However, in the neonatal and adult mouse, GATA-3 is abundantly expressed only in T lymphocytes, with expression either greatly diminished or extinct in many sites that previously expressed the factor during embryogenesis. Therefore, it is possible that GATA-3 may be directing, albeit perhaps only transiently, important developmental programs during fetal development in many cell lineages.

The vital role played by the GATA family of transcription factors during embryogenesis has been most clearly demonstrated by gene targeting in embryonic stem cells, followed by the observation that the phenotypes displayed by GATA-1, -2, and -3 targeted mutations result in embryonic lethality, thus revealing an indispensable role for each individual factor (Pandolfi *et al.*, 1995; Pevny *et al.*, 1991; Tsai *et al.*, 1994). While phenotypic abnormalities arise in many tissues, all three germ-line mutants display profound defects in hematopoiesis. The GATA-3 germ-line mutants die between 10.5 and 11.5 dpc with obvious deformations in the CNS, in fetal liver hematopoiesis, and with a common abdominal hemorrhaging (Pandolfi *et al.*, 1995). One initially surprising phenotype observed in the null-mutant embryos was that the craniofacial region seemed to often be badly malformed, an unanticipated deformity since expression of GATA-3 in this region had not been previously documented. However, upon more refined analysis of further mRNA *in situ* data, we found that GATA-3 was indeed expressed in the branchial arch region during midgestation (Lieuw *et al.*, 1995).

In order to address whether or not, and if so how, GATA-3 might contribute to the formation or elaboration of specific cell lineages during vertebrate embryogenesis, we have begun to investigate the transcriptional regulation of this gene. We initially used transient transfection into GATA-3-expressing or nonexpressing cell lines to attempt to study its transcriptional regulation, but accrued only modest evidence for tissue-specific gene control (George *et al.*, 1994). However, when the identical mGATA-3 sequences were examined in transgenic mice, appropriate tissue-specific expression of the reporter gene was observed, and thus we have continued these studies using transgenic assays. We report here recent progress in this endeavor: the delineation and description of the characteristics of transcriptional regulatory sequences within the locus that spatially and temporally control a subset of the sites at which GATA-3 is expressed during embryogenesis.

When random sequence blocks in the vicinity of the mGATA-3 gene promoter were examined in transgenic mice, we found that regional regulatory domains could be individually defined: in early muscle cells [conferred by mGATA-3 element(s) lying between nt −308 and +1002], in a specific subset of PNS neurons (−2052 to −308), and in the genital tubercle (primordium of the adult external genitalia) and branchial arches (−4500 to −2052). We then refined the position of a 370-bp enhancer between nt −2832 and −2462 that specifically regulates mGATA-3 expression in the branchial arches. We conclude that this element is likely to be the one required for specifying GATA-3 contribution to formation of the jaw in normal embryos (visualized as grossly defective structures in GATA-3 homozygous germ-line mutants), and may thus provide key insight into the earliest phases of craniofacial development.

MATERIALS AND METHODS

DNase I Hypersensitive (HS) Site Mapping

The protocol of Forrester *et al.* (1990) was adapted for this analysis with only minor modifications. Approximately 3×10^8 cells were washed in phosphate-buffered saline (PBS) and resuspended in reticulocyte standard buffer (RSB) with 0.5% NP-40. The cells were Dounce homogenized at 0°C to isolate nuclei and examined for lysis by trypan blue staining. Lysed cells were diluted to 50 ml with cold RSB, collected by low speed centrifugation at 4°C, and finally resuspended in 3 ml of cold RSB. Ten equal aliquots of nuclei were then treated with DNase I, added in a range of 0.25 to 80 μg/ml (final concentration). After a 15-min incubation at 37°C, an equal volume of stop buffer [0.6 *M* NaCl, 20 m*M* Tris–HCl (pH 7.6), 10 m*M* EDTA, 1% SDS, 0.5 mg/ml proteinase K] was added to the reactions. After further incubation at 37 to 45°C for 4 hr to overnight to lyse and digest the nuclei, the DNA was isolated by phenol–chloroform extraction. For Southern blotting, 20 μg of DNA was digested with 100 units of the given restriction enzyme.

Transgenic Mice

Plasmid expression constructs (described below) were isolated from vector DNA by restriction enzyme digestion followed by preparative agarose gel electrophoresis. DNA to be injected was purified using Gene Clean (Bio 101) followed by an Elutip column (Schleicher and Schuell) or spinning thorough a Probind spin filter (Millipore) to remove particulate material. Recovered DNA was finally resuspended in microinjection buffer (10 m*M* Tris–HCl, pH 7.5, 0.1 m*M* EDTA) and injected at a concentration of 1–5 ng/ml. SPF mice were purchased from Charles River or Jackson Labs. Fertilized eggs from CD-1 mice were used for microinjection and transferred into the oviducts of pseudopregnant foster mothers. Embryos were isolated at 10.5 to 13.5 dpc for transient transgenic analysis or were allowed to progress to term to generate transgenic lines (Dillon and Grosveld, 1993; Hogan *et al.*, 1986). Transgenic mice were identified by dot blot, Southern blot, and/or the polymerase chain reaction using lacZ primers (5′.TACCACAGCGGATGGTTCGG.3′ and 5′.GTGGTGGTTATGCCGATCGC.3′) to generate a 352-bp band (Zimmerman *et al.*, 1994).

Whole-Mount lacZ Staining

Embryos were fixed in 1% formaldehyde, 0.2% gluteraldehyde, 2 m*M* $MgCl_2$, 5 m*M* EGTA, and 0.02% NP-40 in PBS for 15 min at ambient temperature. After three washes in PBS plus 0.02% NP-40 for 15 min each, embryos were stained overnight at ambient temperature in a solution containing 5 m*M* $K_3Fe(CN)_6$, 5 m*M* $K_4Fe(CN)_6$, 2 m*M* $MgCl_2$, 0.01% sodium deoxycholate, and 0.02% NP-40 containing 1 mg/ml X-gal (Whiting *et al.*, 1991). Stained embryos were then fixed again and embedded in paraffin for sectioning or were transferred into glycerol (by serial 10% increases to a final concentration of 50%) for photography.

Plasmid Constructions

−308placZ. The plasmid subclone R1.3HNC, derived from genomic clone λ7e, was used as the base plasmid, with the neomycin selection cassette replaced by the lacZ gene (George *et al.*, 1994; Pandolfi *et al.*, 1995). The lacZ gene was isolated by *Not*I digestion of pNASSβ (Clontech, Inc.), followed by blunting with Klenow polymerase. The neomycin resistance gene was removed from R1.3HNC by digesting with *Nco*I (at the mGATA-3 translation initiation codon; +1002 in exon 2) followed by S1 nuclease treatment to remove the endogenous GATA-3 translation initiation codon. The lacZ gene was then ligated into this plasmid, thereby generating −308placZ. R1.3HNC provides the poly(A) signal from the plasmid pMC1NeoA$^+$ (Stratagene).

−2052placZ. −308placZ was digested with *Xho*I (present in the polylinker at the 5′ end) and *Hin*dIII (at +523 bp, in mGATA-3 intron 1). A 2.6-kb *Xho*I to *Hin*dIII genomic DNA fragment of mGATA-3 (−2052 to +523; George *et al.*, 1994) was then ligated to this plasmid to generate −2052placZ.

−4500placZ and −3200placZ. −2052placZ was digested with *Bgl*II (at −1945) plus *Eco*RI (at −308) to generate a 1.6-kb fragment. A *Not*I (−4500) to *Bgl*II (−1945) fragment was isolated from a genomic subclone of λ7e (George *et al.*, 1994) and then ligated with both the *Bgl*II/*Eco*RI fragment and pGEMEX-2 that had been digested with *Not*I plus *Eco*RI. −3200placZ was generated by digesting −4500placZ with *Nco*I and *Bam*HI prior to microinjection.

(−3400/−775)tklacZ. The plasmid pTKβ (Clontech) was digested with *Sal*I to release the HSV thymidine kinase promoter directing lacZ transcription and recloned into pGEMEX2 (which provided more useful restriction enzyme sites) to generate ptkGEMEX. Putative mGATA-3 regulatory elements were then ligated to this plasmid for subsequent *in vivo* analysis. The *Sac*I to *Nsi*I fragment bordering the promoter of mGATA-3 was isolated from the genomic clone λ7e and ligated to ptkGEMEX to generate (−3400/−775)tklacZ.

(−3400/−2052)tklacZ, (−2052/−775)tklacZ, and (−4500/−2052)tklacZ. A *Sac*I to *Xho*I fragment was isolated from genomic clone λ7e and ligated into ptkGEMEX, thereby generating (−3400/−2052)tklacZ. Similarly, the other constructs were created using *Xho*I and *Nsi*I to generate (−2052/−775)tklacZ and *Not*I plus *Xho*I to generate (−4500/−2052)tklacZ.

(−2832/−2052)tklacZ, (−3400/−2462)tklacZ, and (−3400/−2900)tklacZ. A *Sac*I to *Xho*I fragment (−3400 to −2052) of clone λ7e was subcloned into pGEM7, which was then used for the generation of both 5′ and 3′ exonuclease III serial deletions using the Erase-a-Base kit from Promega, Inc. Three deletions (one 5′ and two 3′) were used to generate these three constructs by inserting them into ptkGEMEX.

(−2832/−2462)tklacZ. The 5′ exoIII deletion pGEM7 subclone containing mGATA-3 sequences from −2832 to −2052 in pGEM7 was digested with *Hin*dIII and *Bam*HI (at position −2462) and ligated to ptkGEMEX to generate (−2832/−2462)tklacZ.

(−2462/−2052)tklacZ. A *Bam*HI (−2462) to *Xho*I (−2052) fragment was isolated from the genomic subclone S1.7 (George *et al.*, 1994) and inserted into *Bam*HI/*Xho*I-digested ptkGEMEX.

−3400ΔlacZ. A GATA-3 genomic subclone (8-2; George *et al.*, 1994) containing sequences from −4500 to −308 in pBluescript (Stratagene) was transferred into plasmid vector YIP5, leaving unique *Nco*I and *Bam*HI sites at positions −3200 and −2462, respectively. This plasmid was digested with *Nco*I and *Bam*HI to delete the intervening 738 bp, filled using Klenow polymerase, and religated to create NRΔYIP. The remaining genomic DNA fragment from −3400 to −308 was then released by digestion with *Eco*RI and *Bam*HI and then ligated to the *Eco*RI/*Bam*HI fragment of −300placZ in pGEMEX-2 (Promega) to create −3400ΔlacZ. The insert was isolated from vector sequences by *Sph*I and *Sfi*I digestion prior to microinjection.

RESULTS

DNase I Hypersensitive Site Mapping Localizes Potential mGATA-3 Transcriptional Control Elements

An oft-employed strategy to identify candidate transcriptional regulatory domains for eukaryotic genes is HS site mapping, which identifies regionally altered chromatin configurations in nuclei that can be detected by differential sensitivity to exogenous DNase I (Groudine and Weintraub, 1981). While GATA-3 is expressed in many different cell lineages during embryogenesis, we chose to initially examine three murine cell lines in these assays: a T lymphocyte cell line (BW 5147.3), a neural crest-derived sympathoadrenal representative line (C1300) (Pons *et al.*, 1982), and a fibroblast cell line (3T3). Both BW 5147.3 and C1300 cells express abundant GATA-3 mRNA, whereas 3T3 cells express none (George *et al.*, 1994; Leonard *et al.*, 1993).

The most informative HS mapping data are shown in Fig. 1, where several DNase I HS sites were found in the proximity of the mGATA-3 gene transcription initiation site in the three cell lines. The locus was surveyed using multiple restriction enzyme digests and probes, but only three representative Southern blots are shown; the approximate positions of the restriction fragment endpoints and indirect end-labeled probes used in the three central panels are depicted in Fig. 1A. The 3T3 cells contain two prominent HS sites located near the transcriptional start site (at −1.0 kb and at approximately the mRNA cap site), and these same constitutive sites were found in all three cell lines; an additional HS site immediately inside the gene was detected in T cells and fibroblasts, but not in C1300 cells (Fig. 1B). In a second, more promoter-distal region of the locus, two additional sites were identified in both cell types that express the gene (located −3.1 and −2.6 kb 5′ to the transcriptional start site; Fig. 1C), but not in fibroblasts. At the most distant region 5′ to the start site in the experiments reported here, the T cell line displayed four prominent tissue-specific HS sites,

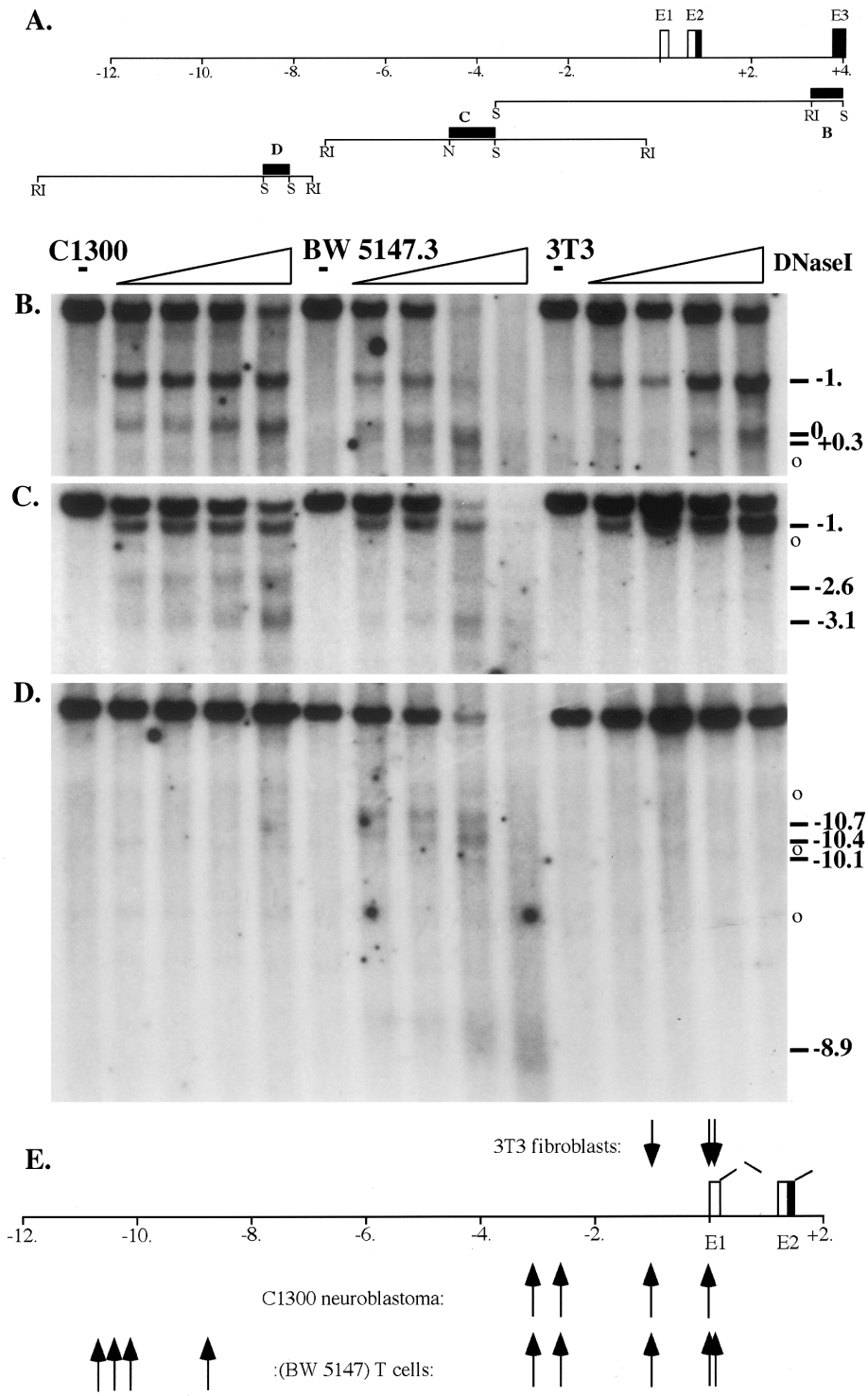


FIG. 1. DNase I hypersensitive site mapping of the mGATA-3 promoter. (A) An abbreviated map of the 5′ end of the mGATA-3 gene locus is depicted (George *et al.*, 1994). The approximate positions of restriction enzyme fragments and probes (B, C, and D) used in the Southern blotting experiments (shown below) are depicted beneath the map. RI, *Eco*RI, S, *Sac*I, N, NotI. (B) A Southern blot showing the positions of DNase I hypersensitive sites surrounding the transcription initiation site. Nuclei were isolated from C1300 neuroblastoma cells (first five lanes), from BW 5147 cells (middle five lanes), or from 3T3 fibroblasts (right five lanes) and then subjected to digestion with increasing amounts of DNase I (see Materials and Methods). Isolated DNA from the nuclei was then digested with *Sac*I and probed with the 3.1-kb *Sac*I to *Eco*RI fragment represented diagramatically in (A) as probe B. The parental *Sac*I band is 7.2 kb, and the sizes of

at approximate position −8.9 and three closely spaced sites at positions −10.1, −10.4, and −10.7 kb; none of these sites were present in C1300 or 3T3 cells (Fig. 1D). A search for other HS sites lying either within or 3′ to the locus failed to identify additional strong sites within the 42 kb described by the original overlapping λ recombinants (George *et al.*, 1994). Therefore, two HS sites that appear to be unique to GATA-3-expressing cells were identified at −2.6 and −3.1 kb, while four additional sites were identified far upstream which appear to be specific for T cells (Fig. 1E).

To further define the activity of the four T cell HS sites, we examined two other T lymphocyte samples in additional HS mapping experiments. The same array of T-cell-specific sites identified in the BW 5147.3 cell line (Fig. 1D) was also found in freshly isolated murine thymocytes, while none of the sites were present in another well-characterized murine T cell line, LSTRA (Glynn *et al.*, 1964; Voronova *et al.*, 1984), which we found did not express GATA-3 mRNA or protein (G.-l.L., K.-C. Lim, and J.D.E., data not shown). Thus the T-cell-specific hypersensitivity displayed at the −8.9, −10.1, −10.4, and −10.7 sites correlated perfectly with mGATA-3 transcription in the three different T cell types examined.

The mGATA-3 T-Cell-Specific HS Sites Do Not Confer T Cell Enhancer Activity

As a direct test to determine whether or not the T-lymphocyte-specific DNase I HS sites represented the mGATA-3 T cell enhancer activity, various constructs were prepared linking DNA fragments containing these sites to the heterologous thymidine kinase gene promoter, which was then used to direct expression of the lacZ gene in transgenic mice (i.e., an enhancer trap assay; Allen *et al.*, 1988; Zimmerman *et al.*, 1994).

First, a contiguous DNA fragment containing all four of the T-cell-specific HS sites (from −12 to −8 kb) was linked to the thymidine kinase promoter used to direct lacZ expression. From nine independent lines, no T cell expression was observed by either histochemical staining or reverse transcriptase polymerase chain reaction. Other constructs that contained either individual HS sites or different combinations of the four sites were also not expressed in T cells. Thus, several constructs, prepared to examine the postulated regulatory activity of the four T-cell-specific DNase I HS sites (Fig. 1E), were not capable of conferring lacZ expression in transgenic thymocytes (data not shown).

To determine whether or not the GATA-3 gene promoter might be required to mediate some postulated cooperative enhancer/promoter-specific interaction with these HS sites (e.g., Li and Noll, 1994), we then tested a construct containing all four sites linked *in cis* to a reporter containing the minimal mGATA-3 gene promoter (−308placZ; see below); however, no expression was detected in T cells derived from nine transgenic mice bearing this construct. Further analysis of a contiguous segment of genomic sequence including all nine of the tissue-specific and ubiquitous HS sites described in Fig. 1 still did not confer T cell expression. We conclude that sequences lying outside the boundaries of these sites must be required for T-cell-specific mGATA-3 transcription, perhaps in collaboration with the sites defined here.

Muscle and Craniofacial Ganglia Transcriptional Control Elements Lie between −2052 and +1002 of the mGATA-3 Gene

We previously reported a preliminary analysis of mGATA-3 transcription by examining three transient transgenic embryos bearing a construct containing sequences from −2052 through +1002 of the gene linked to lacZ (called −2052placZ) which showed a reporter transgene staining pattern that was coincident with the normal mGATA-3 mRNA expression profile in a variety of tissues (George *et al.*, 1994). We concluded that the elements controlling transcription in some tissues, and at some of the proper developmental stages, must lie within the boundaries described by this mGATA-3 transgene, but also that the element(s) conferring appropriate regulation in several other tissues where the gene is normally expressed (e.g., in T lymphocytes, in the developing kidney, and in the CNS) are absent. Because the three transient transgenic embryos were examined at different stages of embryonic development (9.5, 11.0, and 12.5 dpc), it was not possible to compare their expression patterns to rule out potentially confounding effects of transgene integration position or mosaicism. Thus, we expanded the analysis of this potentially rich transcriptional regulatory domain.

Fertilized eggs were microinjected and were either analyzed again as transient transgenic embryos or allowed to progress to term in order to generate stably transmitting lines. We analyzed 16 additional transient transgenic embryos as well as 15 established lines harboring the −2052placZ construct. We made the assumption that inconsistent patterns of lacZ expression were due to integration position effects, and therefore that only a consistent

---

the hypersensitive bands are shown on the right. An ''o'' indicates the positions of faint bands detected in the absence of DNase I treatment (''−'' lanes). (C) A Southern blot depicting HS sites detected in the middle part of the promoter. DNA isolated from DNase I-treated nuclei was digested with *Eco*RI and probed with a 1.1-kb *Not*I/*Sac*I radiolabeled fragment (A, probe C). 3T3 cells (which do not express GATA-3) exhibit hypersensitivity only at the ubiquitous −1.0-kb site. (D) A Southern blot examining the most distal part of the cloned locus. A 5.0-kb *Eco*RI genomic DNA digest was assayed using a 700-bp *Sac*I probe (A, probe D). (E) A summary of the positions of HS sites (indicated by the arrows) identified in the three cell lines.

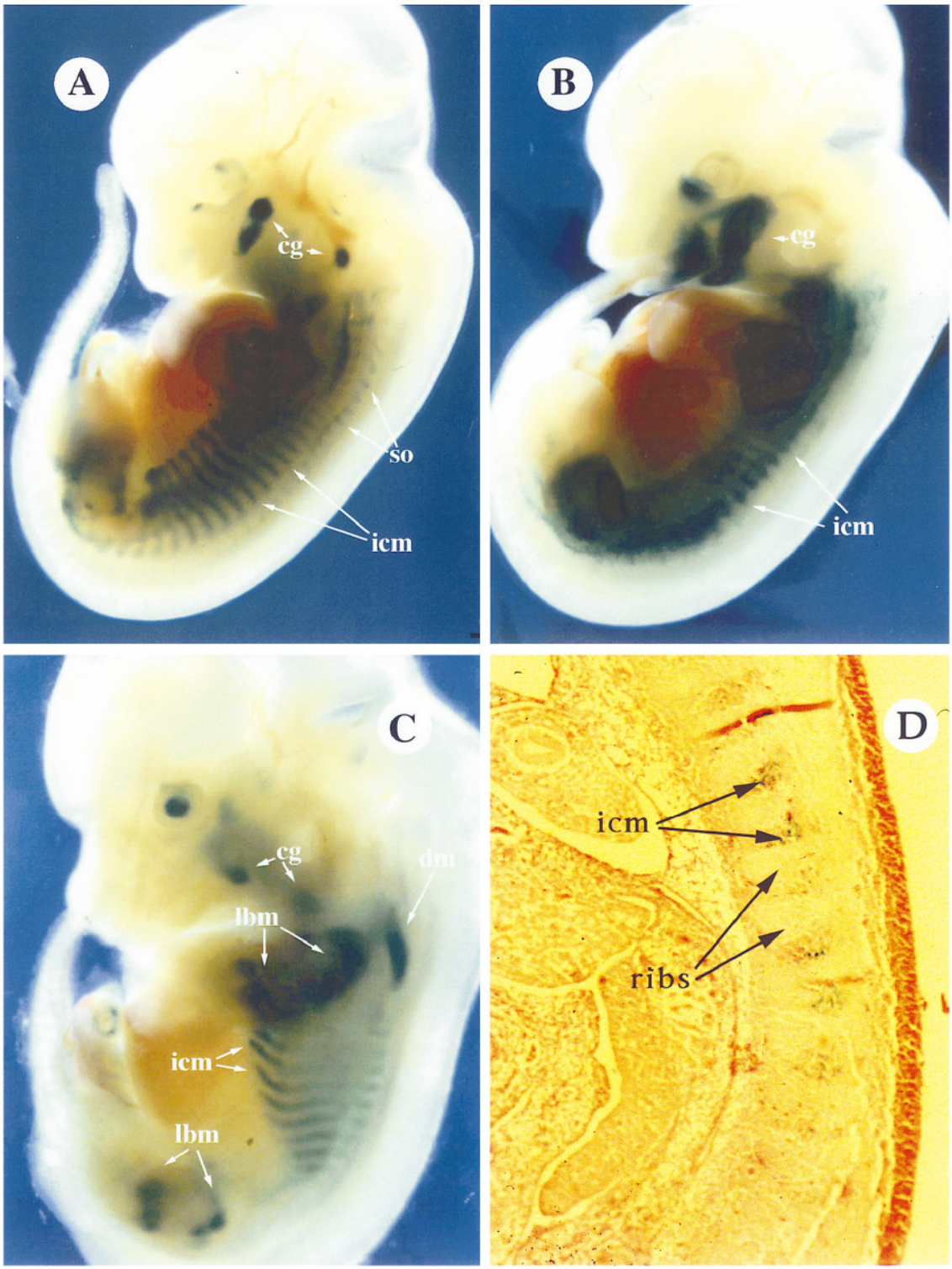


FIG. 2. −2052placZ expression in transgenic embryos. (A and B) Embryos analyzed as transient transgenics (Materials and Methods) while (C) is an example of staining patterns recovered from established lines. (D) A sagittal section of the whole-mount embryo shown in (B). All four embryos show lacZ expression in the premuscle masses of the limb buds (lbm) and intercostal muscles of the ribs (icm) as well as in cranial ganglia (cg) of the peripheral nervous system. The areas of light blue color detected in these photographs (in the heads of these embryos) are not true lacZ staining, but a result of the background color seen through transparent areas. (A) A lateral view of a

pattern displayed by a collection of transgenics harboring a particular construct was conferred by sequences within the transgene.

The −2052placZ reporter gene was usually expressed in the premuscle masses in both the intercostal area of the ribs and in the limb buds. There was also consistent expression in the PNS (in trigeminal ganglia and the facial acoustic complex) coincident with the mRNA *in situ* data (Fig. 2). Therefore, the region between −2052 and +1002 appears to harbor the spatial and temporal information used to direct expression of mGATA-3 to premuscle masses and to a subset of PNS neurons, thus confirming our earlier observations that the regulatory elements present in this transgene were sufficient to confer expression in a subset of tissues where GATA-3 is normally found. However, the observation that only 6 of 16 transients and 3 of 15 established lines display any lacZ expression whatsoever (final results are summarized in Fig. 6) suggests that this construct is very sensitive to transgene integration position effects and can be silenced by chromatin surrounding the integration sites.

We then examined a much smaller reporter transgene (−308placZ) to determine whether or not it might be possible to refine the position of sequences necessary for directing the expression of GATA-3 to premuscle and to cranial ganglia. Eleven transient transgenics and 9 lines were examined. As summarized in Fig. 6, −308placZ was also very sensitive to position effects. Only 3 of 11 transients and 5 of 9 established lines expressed the transgene, very often in ectopic sites. We therefore concluded that this ''minimal promoter'' construct may behave in a manner similar to an enhancer trap, which falls under the immediate regulatory influence of transcriptional control elements present at the (random) genomic integration sites (Allen *et al.*, 1988). The only tissue that consistently reflected normal GATA-3 expression was in the premuscle intercostal masses between the ribs and in the limb buds (Fig. 3). Thus, sequences between −2052 and −308, including only one of the constitutive HS sites identified in Fig. 1, appear to include sequences that may be necessary for regulation of mGATA-3 cranial ganglion expression, whereas the region from −308 to +1002, harboring one constitutive HS site, harbors weak, but reproducible, information conferring control over muscle-specific expression. Computer examination of the DNA sequences within the −308 to +1002 region of the GATA-3 gene revealed multiple CANNTG sequences, suggesting that myogenic E box transcription factors may be at least partly responsible for the muscle expression detected in lines bearing this transgene (e.g., Edmonson and Olson, 1993).

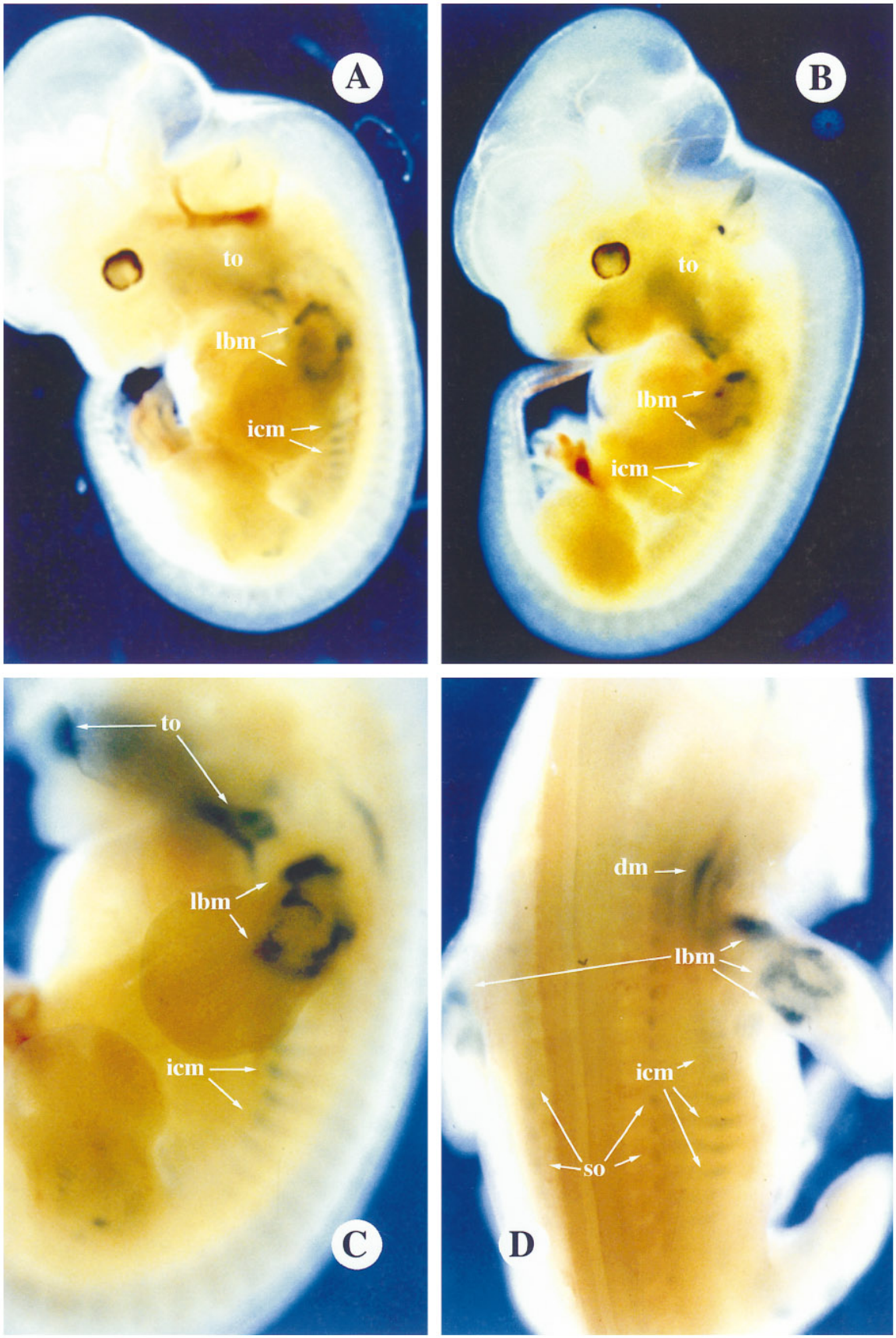
Sequences within −4500 to −2052 of the mGATA-3 Gene Regulate Genital Tubercle and Branchial Arch Expression

In order to identify other regions necessary for proper transcriptional regulation, transgenes were prepared that contained mGATA-3 sequences further removed from the promoter and then linked to the lacZ gene. The largest construct analyzed contained sequences from −4500 to +1002 (−4500placZ), and 5 transient transgenic embryos and 10 stable lines were examined. After staining with Xgal, the expression patterns were carefully compared to rule out sites where the transgene was expressed inconsistently.

As observed with the −2052placZ construct, −4500placZ was also capable of conferring expression in a subset of tissues where GATA-3 is usually found, although it too was expressed at some ectopic sites. The −4500 transgene displayed consistent expression in the premuscle masses between the ribs and in the limb buds, as well as in the cranial ganglia, as expected, since it includes all the sequences examined before (compare Fig. 2 with Fig. 4C), but additional expression was reproducibly detected in the branchial arches and in the genital tubercle (best seen in Figs. 4A, 4B, and 4D).

The branchial arches give rise to the jaws, tongue, and palate (among other sites) in the craniofacial region later during development, and was of particular interest since the mGATA-3 homozygous mutant embryos displayed gross morphological defects in this region (Pandolfi *et al.*, 1995). Expression of GATA-3 in the genital tubercle was initially surprising, since it had not been reported previously. However, since the tissue sections analyzed in previous studies did not include the extreme caudal end of the embryo, it seems possible that this expression was simply overlooked in the initial survey (George *et al.*, 1994). In addition to conferring GATA-3-specific expression at new anatomical sites, −4500placZ was considerably more resistant to transgene integration position effects, since 4 of 5 transients and 9 of 10 transgenic lines displayed some lacZ expression (Fig. 6). Thus, we drew the tentative conclusion that the region between −4500 and −2052 contains the spatial and temporal information for both branchial arch and genital tubercle regulation of mGATA-3.

12.0-dpc transgenic embryo showing expression in premuscle masses and cranial ganglia. Note that there is some expression detected in somites (so). (B) A comparable stage transient transgenic embryo from a different integration which shows expression in the same tissues as in (A) but seems to be slightly broader as well as more intense. (C) A lateral view of a 12.5-dpc embryo showing a representative expression pattern in the developing intercostal muscles slightly later in development. There appears to be ectopic expression in the eye of this transgenic line, even though GATA-3 has been reported to be present in lens fibers (Oosterwegel *et al.*, 1992). (D) GATA-3 transgene expression is found in the developing intercostal muscles rather than the developing ribs in −2052placZ transgenic animals. The same labeling pattern is detected in the embryonic intercostal muscle with −4500placZ and −300placZ.


A
to
lbm
icm
B
to
lbm
icm
to
lbm
icm
C
dm
lbm
icm
so
D

To determine whether or not the temporal expression exhibited by this construct reflected that of the endogenous mGATA-3 gene, the expression of the −4500 transgene from various developmental stages was also examined. The earliest site of expression was found to be at 8.0 dpc in the branchial arches and in the tail bud (which gives rise to the genital tubercle as well as other posterior structures; Fig. 4A). Previously, we did not detect abundant embryonic GATA-3 expression before 9.5 dpc by mRNA *in situ* analysis (George *et al.*, 1994); however, prior to 8.5 dpc, GATA-2 and GATA-3 mRNAs are both abundantly transcribed in the ectoplacental cone, which later develops into the placenta (Ng *et al.*, 1994), but by 9.5 dpc, GATA-3 also accumulates in the dermomyotome of the somites. While the −4500 transgene fails to be expressed in the somites at this stage, expression was detected later (by 11.0 dpc; data not shown) in the premuscle masses of the intercostal muscles of the ribs, another site of normal GATA-3 mRNA accumulation later in development (Fig. 4C). Thus, while some of the temporal information required for the earliest control of muscle-specific expression appears to be missing (compared to the strong somitic dermomyotome labeling detected by *in situ* hybridization) (George *et al.*, 1994), the correct patterning information for expression of mGATA-3 in the branchial arches and in the genital tubercle does appear to be present in the transgene described by the boundaries within −4500placZ. A smaller construct, −3200placZ, also gave the same expression pattern, thereby further localizing the branchial arch and genital tubercle element(s).

Refining the mGATA-3 Branchial Arch Transcriptional Control Element

We next focused on the activity of sequences containing three of the HS sites (at −1.0, −2.6, and −3.1) found between −3400 and −775 of the promoter by examining their expression using the enhancer trap (HSV tk gene promoter-directed) lacZ plasmid. After ligation to the reporter gene, consistent expression was detected in the branchial arches, the genital tubercle, and the limb buds, showing that elements capable of directing a heterologous promoter to expression in these appropriate tissues were present within that sequence (Fig. 5B).

Since the GATA-3 null mutants display a profound craniofacial defect (Pandolfi *et al.*, 1995), we focused on refining the regulatory element required for branchial arch-specific expression. As also shown in Fig. 5C, an element between −3400 and −2052 was found to be sufficient for directing lacZ expression to both the branchial arches and genital tubercle, whereas sequences from −2052 to −775, as anticipated (Fig. 6), directed no expression in any of the transgenic embryos. When sequences between −3400 and −2462 were tested, transgene expression was detected in only the branchial arches (Fig. 5E), thus further localizing this element (and deductively localizing the genital tubercle element as well). Since sequences from −2832 to −2052 were shown to be capable of directing expression to both the branchial arch and the genital tubercle (Fig. 5D), we predicted that the branchial arch element would lie between nucleotide positions −2832 and −2462, whereas the genital tubercle element would lie between positions −2462 and −2052.

Constructs linking these two separate sequence elements to the tk/lacZ enhancer trap plasmid were prepared, and as anticipated, transgene expression in the branchial arches was reproducibly detected in animals bearing the −2832 to −2462-directed reporter gene (Figs. 5G and 5H). However, none of 28 transgenic embryos that were examined bearing the −2462 to −2052-directed reporter gene displayed expression in the genital tubercle (Fig. 6, and data not shown). Thus, mGATA-3 expression in the genital tubercle appears to require additional information from sequences lying 5′ to position −2462 which might cooperate with sequences between −2462 and −2052. It is also possible that critical sequences for genital tubercle expression were bisected at position −2462 in generating this construct.

Interestingly, a consistent pattern was observed when sequences lying between −2462 and −2052 were examined (Figs. 5F and 5I): both the tip of the nose and a region of the face bordering the expression boundaries found with the (−2832/−2462)tklacZ construct were labeled in these animals. Since GATA-3 is not expressed in this region, the sequences tested, at least in combination with the thymidine kinase promoter, seem to be misinterpreting the spatial information encoded by the −2462 and −2052 sequence in the mGATA-3 gene promoter, resulting in aberrant expression. Thus, it seems likely that the two regulatory elements specifying the branchial arch and genital tubercle expression are both required to confer proper spatial information to both regulatory patterns. The fact that the entire sequence from −2832 to −2052 expresses the transgenes more frequently, as well as being more robust in the enhancer assay than either element alone, supports this hypothesis.

FIG. 3. Expression of −308placZ in transgenic embryos. (A and B) Lateral views of two different 12.0-dpc transgenic embryos showing expression in the premuscle masses of the limb buds and intercostal muscles. There is also consistent expression in the developing tongue (to) in both embryos. However, embryo (B) also displays ectopic expression in the jaw and eyes; although GATA-3 is normally expressed in these organs, the expression pattern detected in this particular animal is considered ectopic, since the same pattern was detected only in a minority of the −300placZ transgenics. (C and D) Lateral and dorsal views of a third 12.0-dpc embryo also showing expression of the transgene in the premuscle masses of the limb buds and intercostal muscles. Note that there is particularly strong staining in the developing tongue in (C) and that there is weak expression in the somites (so) in (D).

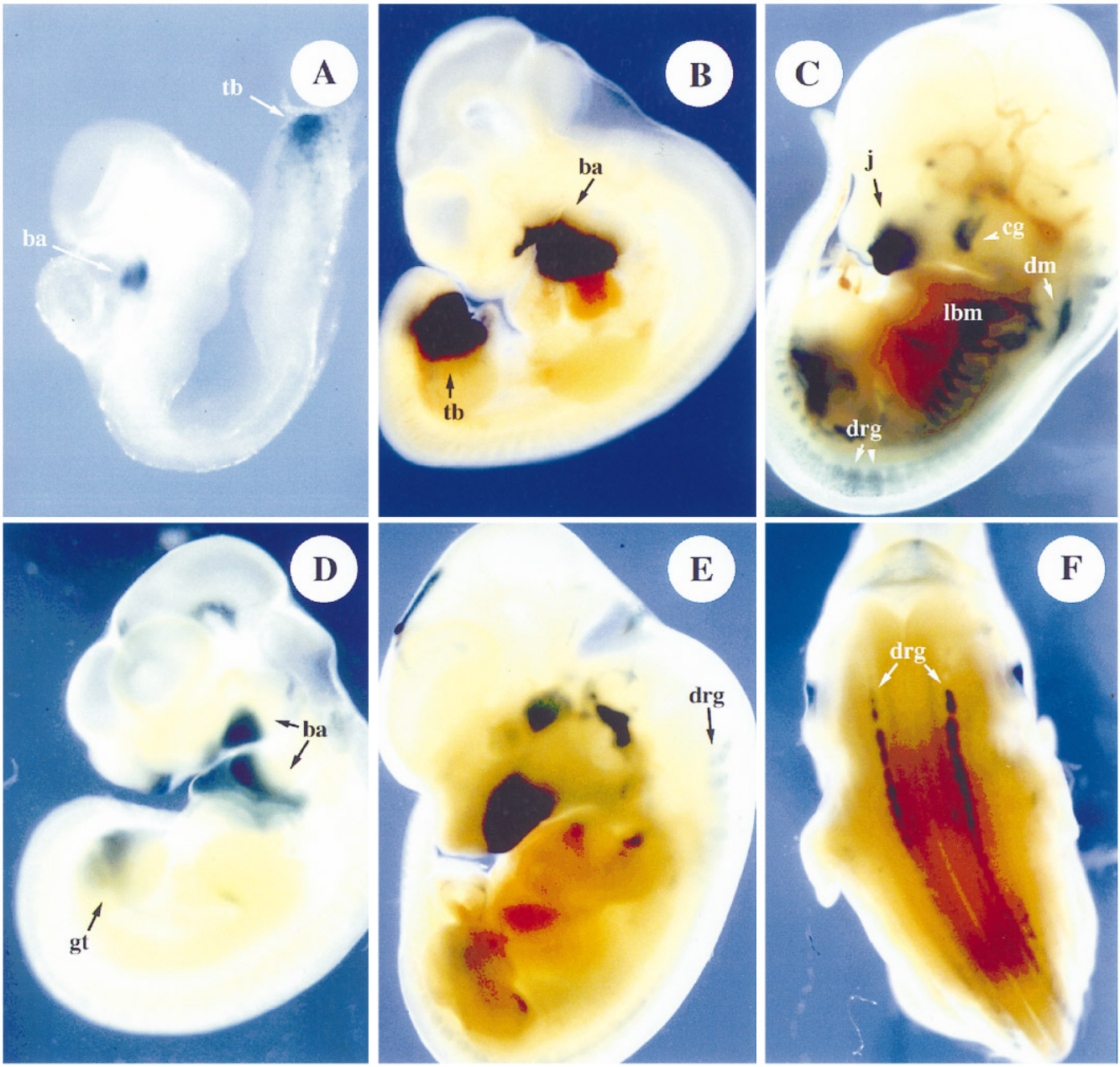


FIG. 4. Expression of −4500placZ in transgenic mice. (A, B, and C) Lateral views of three embryos analyzed at three different developmental stages from the same transgenic line to depict the temporal expression profile of this transgene. (A) Expression is first detected in the branchial arch (ba) and tail bud (tb) at around 8.0 dpc. (B) A 10.0-dpc embryo shows elaborated expression of the transgene in both the maxillary and the mandibular components of the first branchial arch as well as the tail bud, which later localizes to the genital tubercle (gt). Note that by 12.5 dpc (C), the transgene is expressed widely throughout the jaw (j), in dorsal root ganglia (drg), in cranial ganglia (cg), and in the premuscle of the limb buds and intercostal region. (D) A separate transgenic line gives rise to the same spatial transgene expression pattern at 10.5 dpc as that detected in (B). Note that this integration also results in an ectopic pattern in the midbrain. (E and F) A transient transgenic embryo showing a similar spatial pattern found in the two previous lines. In this embryo, there is weaker expression in somite-derived musculature structures, as well as expression again in, but a different region of, the midbrain.

None of the transgenics produced from the (−2832/−2462)tklacZ construct express strongly [in comparison to the (−2832/−2052)tklacZ construct] in the branchial arches (compare Figs. 5G and 5H to 5D), suggesting that additional positive regulatory elements affecting branchial arch expression may be present 3′ to position −2462. When a 738-bp region containing both the HS sites at −3.1 and −2.6 were deleted within the construct containing sequences from −3400 to +1002, both branchial arch and genital tubercle expression were affected, providing further support for such a mechanism (Fig. 5I). While there was residual branchial arch expression even after deletion of the −2.6 HS site, genital tubercle expression was ablated. Thus, the sequences lying between positions −2832 and −2462 act as an independent enhancer that alone is sufficient for directing a heterologous promoter to expression in the branchial arches, and additionally appears to be required for cooperation with sequences located between −2462 and −2052 for specifying genital tubercle expression. The summary results of all the transgenic assays are represented diagramatically in Fig. 6.

The sequence of this branchial arch enhancer is shown in Fig. 7, and as anticipated, it contains potential binding sites for multiple transcription factors. While the DNA sequence does not predict which of these factors might actually be present in mesenchymal primordia that develop into the jaw, the characterization of the sequence as displaying the dual properties of many transcription factor-binding sites being present and densely localized within the sequence serves as prima facia evidence that this segment of DNA could indeed serve well as a rather typical eukaryotic enhancer, which has been directly demonstrated in these functional assays (Fig. 5).

DISCUSSION

In attempting to gain greater insight into the complex embryonic regulation of GATA-3 expression, we describe here the regional definition of several *cis*-regulatory elements that appear to be able to recapitulate aspects of normal control over this gene during midgestation in transgenic mice. After the potential regulatory domains of the locus were first identified using DNase I HS site mapping, further functional analysis of these regions was undertaken using transgenic mice. While no function could be attributed to a group of T-cell-specific HS discovered far 5′ to the promoter (below), a potent regulatory element was shown to be present among sequences that were able to confer branchial arch- and genital tubercle-specific expression. We then localized the mGATA-3 branchial arch-specific enhancer activity to a 370-bp sequence that contains one of the HS sites (at −2.6 kb) identified in GATA-3-expressing cells. When a region containing two of the sites (at −2.6 and −3.1) was deleted within an otherwise intact promoter (bearing sequences from −3400 to +1012), expression in both the branchial arches and the genital tubercle was severely affected. By enhancer trap analysis, the majority of the branchial arch enhancer activity was localized to positions −2462 to −2834, but it seems that within the context of the endogenous promoter, there is residual sequence information still present in a construct missing the −2.6-kb HS site which allows residual branchial arch transcription. Perhaps these HS sites are required to efficiently express GATA-3 in the branchial arches, but when integrated into a permissive chromatin environment, sufficient sequence information remains to provide low level, but temporally and spatially correct activity.

Even though the *cis*-regulatory elements of the GATA-3 gene appear to be arranged in a modular organization, only the branchial arch element was defined here as one capable of independently directing a heterologous promoter to expression in a discrete cell population. We also conclude that the genital tubercle element cooperates with at least a subset of sequences found within this branchial arch element to provide spatially correct information to patterning of GATA-3 expression in the primordia of the adult genitalia. Other spatially specific regions described in the present analysis may require collaboration or cooperative interaction with the GATA-3 gene promoter in order to be revealed.

When sequences between −3400 and −775 were linked to the tk gene promoter, transgene expression was revealed in the limb buds as well as in the branchial arches and genital tubercle. However, the location of lacZ expression in the limb buds differed from the pattern found with the endogenous promoter, suggesting that correct spatial patterning requires information from the endogenous promoter (compare Fig. 3 to Fig. 5). It has been shown previously that enhancers and promoters can display promiscuous synergy and work in different combinations (Kermekchiev *et al.*, 1991). However, in analyzing spatially specific gene regulation, it is important to remember that individual elements can act to either activate or suppress transcription, and thus both activities may be required to combinatorially generate specific patterns of gene expression. Thus a combinatorial interaction of multiple elements is probably required to elaborate an entirely appropriate tissue-specific transcription to mGATA-3, and consequently it is not surprising that this analysis showed enhanced spatial specificity that more closely resembled the normal expression pattern when more distal mGATA-3 regulatory elements were examined in the context of its own promoter.

Because GATA-3 is expressed in multiple tissues, in both a temporally and a spatially dynamic manner, it is perhaps not surprising that we have been unable to fully recapitulate the normal expression pattern of the endogenous gene in functional analysis of only 6 kb of DNA surrounding the transcriptional initiation site. While there are now a number of examples showing that gene-proximal sequences are often sufficient to recapitulate the endogenous expression pattern, there are also a growing number of examples where the regulation of other genes appears to be more complex. For example, the myogenin regulatory pattern likely falls

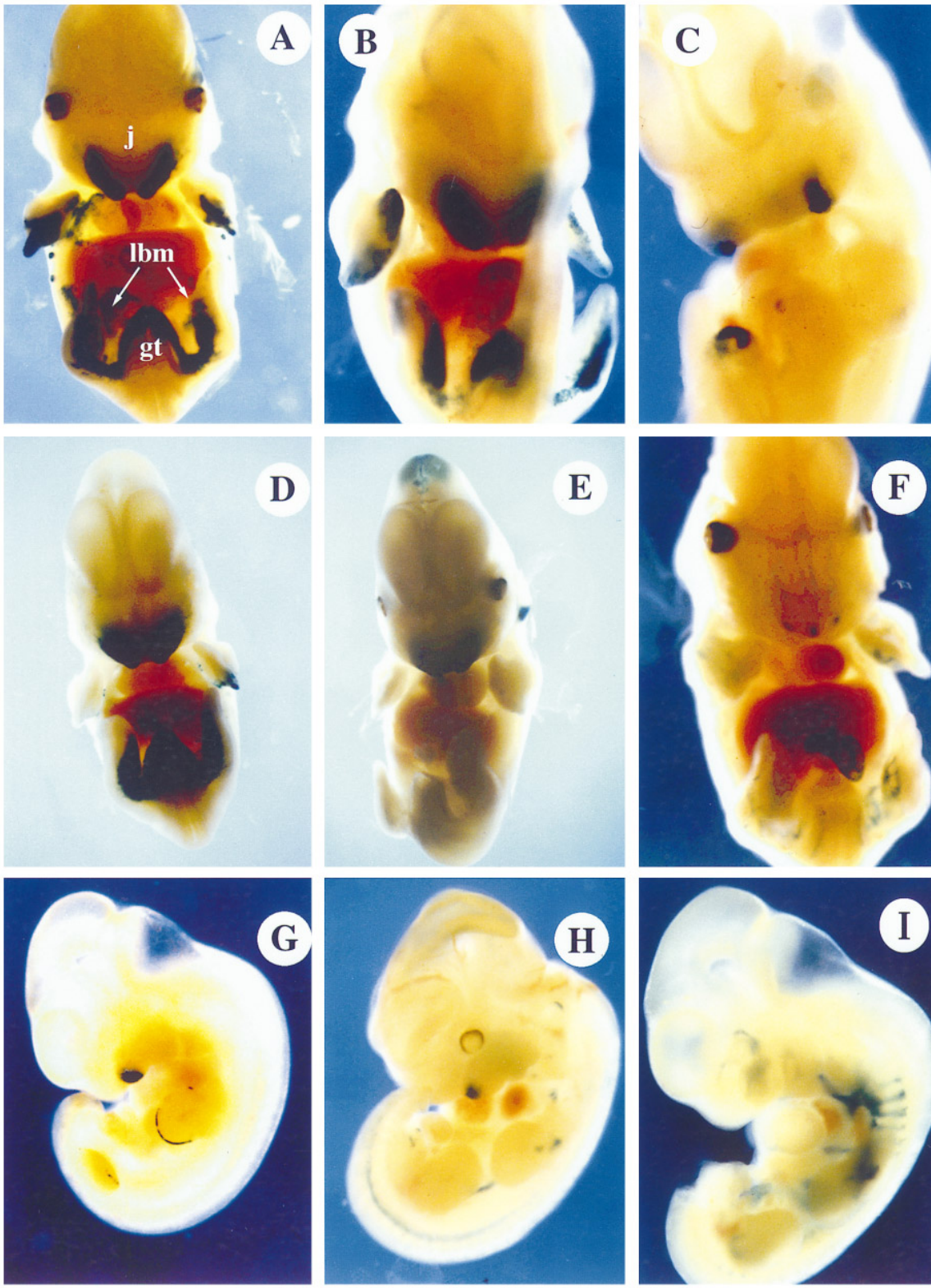


FIG. 5. Localization of the mGATA-3 branchial arch element. (A) A 13.0-dpc (−4500/−2052)tklacZ (Materials and Methods) transient transgenic embryo showing strong expression in the branchial arch-derived facial structures and the genital tubercle with continuous staining into the posterior limb buds. There is also expression in the nipples (visible on the lateral aspects of the thorax in this view),

into the former more simple type, in which its transcriptional control appears to be fully recapitulated by sequences lying between −1565 and +18 to the gene (Cheng *et al.*, 1993); similarly, only 2.6 kb of 5′ flanking sequence appears to be sufficient for proper regulation of the hox-2.6 gene (Whiting *et al.*, 1991). More complex regulatory interactions have been well documented in *Drosophila* (Arnosti *et al.*, 1996; Cai and Levine, 1995) and are now being revealed in a growing number of vertebrate studies as well. Contrary examples to the relatively simple myogenin and hox-2.6 patterns were detailed through the analysis of En-2 transcriptional regulation, where 9.5 kb of the genomic locus was required to recapitulate appropriate transcription, while in analyzing both nestin and CD4 transcriptional regulation, modular enhancers or silencers located in introns were shown to be required for their proper regulation (Sawada *et al.*, 1994; Zimmerman *et al.*, 1994). Even more complex regulation has been defined in the murine Steel locus, in which a DNA rearrangement located more than 100 kb 5′ to the coding region has been shown to cause deregulated transcription (Bedell *et al.*, 1995). From the present analysis, the GATA-3 locus seems to be subject to the latter, more complex type of gene regulation, and we conclude that sequences required to recapitulate the entire GATA-3 expression profile during midembryogenesis clearly lie outside of the boundaries defined by the sequences examined here.

GATA-3 Regulation in T Cells

The best documented site of expression, as well as most of the known cellular target genes, for GATA-3 in adult animals is in T lymphocytes (Hambor *et al.*, 1993; Joulin *et al.*, 1991; Ko *et al.*, 1991; Landry *et al.*, 1993; Marine and Winoto, 1991; Siegel *et al.*, 1995), but it may also be required for hematopoietic progenitor development since there appears to be a profound defect in definitive hematopoiesis in the germ-line mutant animals (Pandolfi *et al.*, 1995). Moreover, while there is overlapping expression of GATA-2 and GATA-3 in many embryonic tissues, GATA-3 appears to be the only member of this family expressed in T lymphocytes. Thus, we postulated that GATA-3 might play an important role in the genesis and/or elaboration of the T cell lineage, which has recently been documented (A. Karis and F. Grosveld, personal communication; Ting *et al.*, 1996).

In attempting to decipher the mechanism for transcriptional regulation of GATA-3 in T lymphocytes, DNase I HS mapping was performed, and four T-cell-specific HS sites were identified, located as far as 10.7 kb 5′ to the mGATA-3 transcription start site. In several assays, when these sites were tested for the ability to direct either the endogenous or a heterologous gene promoter to T-cell-specific expression, we were unable to verify that they play any functional role in GATA-3 transcriptional regulation (K.H.L. and G.-l.L., unpublished observations). It is nonetheless formally possible that these sites are important for GATA-3 expression in T cells and that the assay itself was limiting. It is also possible that other regulatory elements, lying outside the boundaries described in the present study, are needed in combination with these more gene-proximal HS sites to confer T-cell-specific expression. One possible explanation for why we were unable to detect T-cell-specific expression with the constructs tested is that the lacZ gene that we are using contains DNA sequences which prohibit its expression in T lymphocytes, since we observed none in all the transgenics analyzed here. Alternatively, the HS sites may be conferring activities to the gene that do not play a transcriptional regulatory role; these may have to do with structural alterations in the chromatin of T lymphocytes dealing with replication or chromosome dynamics that would not be detected in a transcription assay.

The Role of GATA-3 in Craniofacial Development

Since GATA-3 homozygous mutant embryos display profound craniofacial deformities and since GATA-3 is expressed early during development in the branchial arches, GATA-3 seems to play an important role in face morphogenesis. The creation of many knockout mice displaying a craniofacial phenotype has highlighted the complexity of craniofacial development and the fact that many genes play previously unexpected roles in this process (Richman and

---

which was consistently detected in the −4500placZ transgenic lines. (B) A 13.0-dpc (−3400/−775)tklacZ transient transgenic embryo with the same spatial expression pattern found in (A), but missing nipple expression. (C) A 12.0-dpc (−3400/−2052)tklacZ transgenic embryo with expression that is weaker than, but spatially very similar to, the branchial arch-derived jaw and genital tubercle expression. The tail was intentionally removed to allow better visualization of the genital tubercle in this animal. (D) A 12.5-dpc (−2835/−2052)tklacZ transgenic embryo showing that the branchial arch element and the genital tubercle element localize within this region. (E) A 12.5-dpc (−3400/−2462)tklacZ embryo, displaying only the branchial arch-derived jaw expression. There is also light ectopic expression in the midbrain of this embryo. (F) An example of the only consistent pattern detected in (−2462/−2052)tklacZ transgenic mice. There was no genital tubercle expression in any of the transgenic embryos, while aberrant expression was found consistently at the tip of the nasal prominence, suggesting that the (−2835/−2462 and −2462/−2052) regions are both required to confer proper spatial expression in both regions (compare to E). (G and H) Two examples of transgenic embryos (11.5 dpc) showing specific expression in the branchial arches in the (−2835/−2462)tklacZ embryos. The majority of the transgenics bearing this construct showed specific branchial arch expression, although the overall intensity of staining seems less robust than that detected with larger constructs (compare to A and B). (I) An 11.5-dpc, −3400ΔlacZ embryo showing residual expression in the branchial arch even after deletion of sequences between −3200 and −2462 (two of the HS sites, located at −3.1 and −2.6, are removed in this deletion).

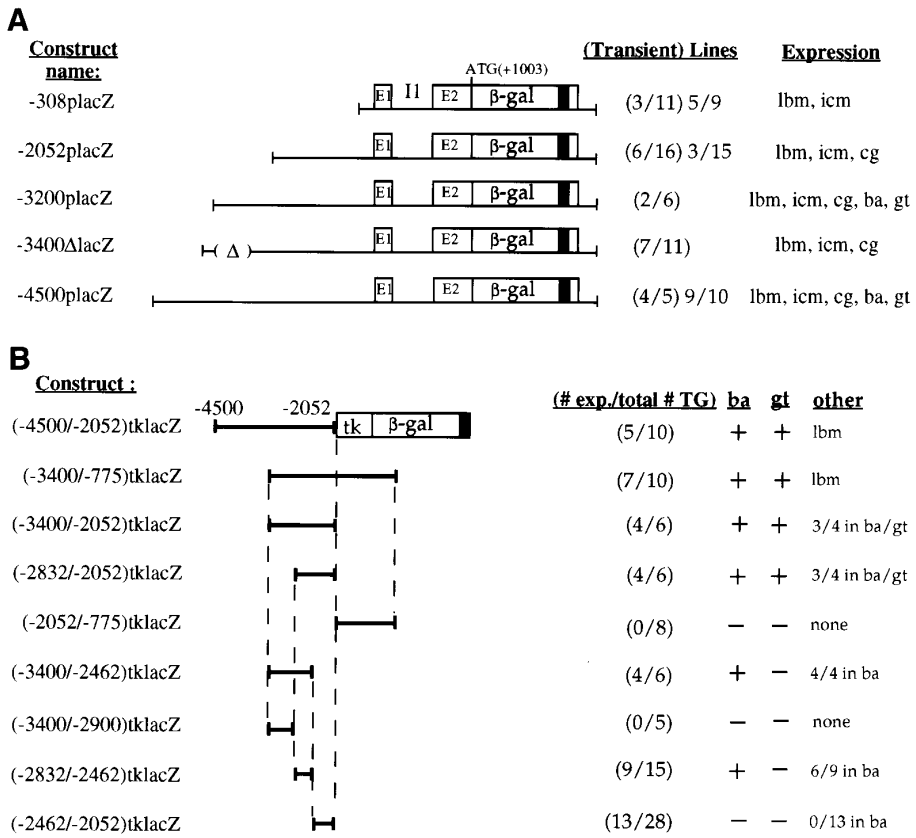


FIG. 6. Summary and expression characteristics of the mGATA-3/lacZ transgenes. Diagrams of the three GATA-3 promoter-directed transgenes are shown on the top three lines: −308, −2052, and −4500 describe the 5′ boundaries of the three constructs, and all were joined to the lacZ gene at mGATA-3 nt position +1002 (Materials and Methods; George *et al.,* 1994). The lightly stippled box represents the lacZ coding sequence (“*β*-gal”), the black box represents the poly(A) addition signal and mGATA-3 exons are depicted by open boxes. The number of transgenics that expressed the transgene are shown in the column “(Transient) Lines,” separated from the total number of transgenics analyzed by a slash. The numbers of transgenics analyzed as transient transgenic embryos are surrounded by parentheses, while the comparable numbers for transgenic lines are not. A summary of the consistent expression pattern is also included: lbm (limb bud muscle), icm (intercostal muscle), cg (cranial ganglia), ba (branchial arches), and gt (genital tubercle). A summary of the enhancer trap analysis is represented diagramatically on the bottom nine lines. The genomic DNA sequence analyzed is represented by the thick black line, and the approximate (for −4500 and −3400) or precise 5′ and 3′ (separated by a slash) mGATA-3 genomic locus endpoints of each fragment analyzed are shown in the construct name. All of the constructs were linked to the HSV thymidine kinase gene promoter which was used to direct lacZ expression. The number of expressing and total transgenic embryos analyzed is indicated, as are the expression characteristics of each construct.

Mitchell, 1996). For example, it was initially anticipated that endothelin-1 may be critical to the regulation of blood pressure due to its expression in endothelial cells and its effect on vascular smooth muscle. However, in addition to encountering the anticipated blood pressure defect, endothelin-1 null-mutant mice also showed a significant reduction in the size of the lower jaw (Kurihara *et al.,* 1994), and later, it was surmised that the jaw defect is likely due to the fact that endothelin-1 and its receptors were expressed in the osteogenic mesenchyme of the mandibular prominence (Barni *et al.,* 1995). Interestingly, GATA-2 has been shown to be important for the regulation of preproendothelin-1 transcription in endothelial cells (Dorfman *et al.,* 1992; Lee *et al.,* 1991). It is therefore tempting to speculate that the homologous GATA-2 and GATA-3 transcription factors may be responsible for separately regulating the expression of both the endothelin-1 ligand and its receptor in the branchial arches.

Similarly to the emerging endothelin-1 example, we observed GATA-3 expression in the branchial arches only after noticing that the targeted germ-line GATA-3 mutants exhibited prominent craniofacial deformities (Pandolfi *et al.,* 1995). With the description of the enhancer that is necessary for branchial arch-specific expression presented here, and the identification of possible binding sites for transcription factors that thereby become implicated in the coregulation

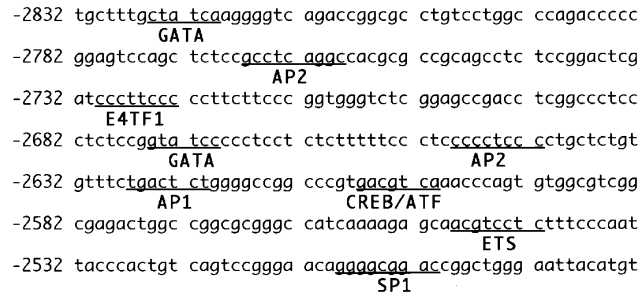


FIG. 7. Nucleotide sequence and potential transcription factor-binding sites in the mGATA-3 gene branchial arch enhancer. The DNA sequence of the locus was extended from the previous 5′ boundary at nt position −2052 (George *et al.*, 1994) to approximately position −3400 (data not shown). Potential transcription factor-binding sites are underlined with the name of the transcription factor written below the sequence of the mGATA-3 genomic DNA fragment that confers branchial arch-specific expression in the enhancer trap assay [(−2832/−2462)tklacZ, Fig. 5]. Many other incomplete (partial) matches to published binding-site consensus DNA sequences (revealed using MacVector software) are not shown.

of this enhancer early in jaw development, we can now pursue the presumably collaborative role that GATA-3 (as well as these other factors) may play in regulating this enhancer.

ACKNOWLEDGMENTS

We thank Steve Fiering and Mark Groudine (Fred Hutchinson Cancer Research Center) for providing detailed DNase I HS site mapping protocols. We also thank Jie Fan for excellent technical assistance and Dave Balwanz for help in preparing the −3400ΔlacZ construct. This work was supported by an institutional NIH MSTP Training Award (GM 08152; K.H.L.), a NATO Collaborative Research Grant (921110; F.G. and J.D.E.), and an NIH grant (GM 28896; J.D.E.).

REFERENCES

Allen, N. D., Cran, D. G., Barton, S. C., Hettle, S., Reik, W., and Surani, M. A. (1988). Transgenes as probes for active chromosomal domains in mouse development. *Nature* **333,** 852–855.

Arceci, R. J., King, A. A. J., Simon, M. C., Orkin, S. H., and Wilson, D. B. (1993). Mouse GATA-4: A retinoic acid-inducible GATA-binding transcription factor expressed in endodermally derived tissues and heart. *Mol. Cell. Biol.* **13,** 2235–2246.

Arnosti, D. N., Barolo, S., Levine, M., and Small, S. (1996). The eve stripe 2 enhancer employs multiple modes of transcriptional synergy. *Development* **122,** 205–214.

Barni, T., Maggi, M., Fantoni, G., Serio, M., Tollaro, I., Gloria, L., and Vanelli, G. B. (1995). Identification and localization of endothelin-1 and its receptors in human fetal jaws. *Dev. Biol.* **169,** 373–377.

Bedell, M. A., Brannan, C. I., Evans, E. P., Copeland, N. G., Jenkins, N. A., and Donovan, P. J. (1995). DNA rearrangements located over 100 kb 5′ of the Steel (*Sl*)-coding region in Steel-panda and Steel-contrasted mice deregulate *Sl* expression and cause female sterility by disrupting ovarian follicle development. *Genes Dev.* **9,** 455–470.

Cai, H., and Levine, M. (1995). Modulation of enhancer–promoter interactions by insulators in the Drosophila embryo. *Nature* **376,** 533–376.

Chazaud, C., Oulad-Abdelghani, M., Bouillet, P., Decimo, D., Chambon, P., and Dolle, P. (1996). AP-2.2, a novel gene related to AP-2, is expressed in the forebrain, limbs and face during mouse embryogenesis. *Mech. Dev.* **54,** 83–94.

Cheng, T.-C., Wallace, M. C., Merlie, J. P., and Olson, E. N. (1993). Separable regulatory elements governing myogenin transcription in mouse embryogenesis. *Science* **261,** 215–218.

Dillon, N., and Grosveld, F. (1993). Transcriptional analysis using transgenic animals. *In* ''Gene Transcription: A Practical Approach,'' pp. 153–188.

Dorfman, D. M., Wilson, D. B., Bruns, G. A. P., and Orkin, S. H. (1992). Human transcription factor GATA-2: Evidence for regulation of preproendothelin-1 gene expression in endothelial cells. *J. Biol. Chem.* **267,** 1279–1285.

Edmonson, D. G., and Olson, E. N. (1993). Helix–loop–helix proteins as regulators of muscle-specific transcription. *J. Biol. Chem.* **268,** 755–758.

Forrester, W. C., Epner, E., Driscoll, M. C., Enver, T., Brice, M., Papayannopoulou, T., and Groudine, M. (1990). A deletion of the human β-globin locus activation region causes a major alteration in chromatin structure and replication across the entire β-globin locus. *Genes Dev.* **4,** 1637–1649.

George, K. M., Leonard, M. W., Roth, M. E., Lieuw, K., Kioussis, D., Grosveld, F., and Engel, J. D. (1994). Embryonic expression and cloning of the murine GATA-3 gene. *Development* **120,** 2673–2686.

Glynn, J. P., Bianco, A. R., and Goldin, A. (1964). Studies on induced resistance against isotransplantants of virus-induced leukemia. *Cancer Res.* **24,** 502–508.

Groudine, M., and Weintraub, H. (1981). Activation of globin genes during chicken development. *Cell* **24,** 393–401.

Hambor, J. E., Mennone, J., Coon, M. E., Hanke, J. H., and Kavathas, P. (1993). Identification and characterization of an *Alu*-containing, T-cell-specific enhancer located in the last intron of the human CD8α gene. *Mol. Cell. Biol.* **13,** 7056–7070.

Hogan, B., Costantini, F., and Lacy, E. (1986). ''Manipulating the Mouse Embryo.'' Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Ito, E., Toki, T., Ishihara, H., Ohtani, H., Gu, L., Yokoyama, M., Engel, J. D., and Yamamoto, M. (1993). Erythroid transcription factor GATA-1 is abundantly transcribed in mouse testis. *Nature* **362,** 466–469.

Joulin, V., Bories, D., Eleouet, J.-F., Labastie, M.-C., Chretien, S., Mattei, M.-G., and Romeo, P.-H. (1991). A T-cell specific TCR δ DNA binding protein is a member of the human GATA family. *EMBO J.* **10,** 1809–1816.

Kelley, C., Blumberg, H., Zon, L. I., and Evans, T. (1993). GATA-4 is a novel transcription factor expressed in endocardium of the developing heart. *Development* **118,** 817–827.

Kermekchiev, M., Pettersson, M., Matthias, P., and Schaffner, W. (1991). Every enhancer works with every promoter for all the

combinations tested: Could new regulatory pathways evolve by enhancer shuffling? *Gene Expression* **1,** 71–81.

Ko, L. J., Yamamoto, M., Leonard, M. W., George, K. M., Ting, P., and Engel, J. D. (1991). Murine and human T lymphocyte GATA-3 factors mediate transcription through a *cis*-regulatory element within the human T cell receptor δ gene enhancer. *Mol. Cell. Biol.* **11,** 2778–2784.

Kurihara, Y., Kurihara, H., Suzuki, H., Kodama, T., Maemura, K., Nagai, R., Oda, H., Kuwaki, T., Cao, W.-H., Kamada, N., Jishage, K., Ouchi, Y., Azuma, S., Toyoda, Y., Ishikawa, T., Kumada, M., and Yazaki, Y. (1994). Elevated blood pressure and craniofacial abnormalities in mice deficient in endothelin-1. *Nature* **368,** 703–710.

Landry, D. B., Engel, J. D., and Sen, R. (1993). Functional GATA-3 binding sites within murine CD8a upstream regulatory sequences. *J. Exp. Med.* **178,** 941–949.

Lavernere, A. C., MacNeill, C., Mueller, C., Poehmann, R., Burch, J., and Evans, T. (1994). GATA-4, -5, and -6 constitute a new subfamily of transcription factors in developing heart and gut. *J. Biol. Chem.* **269,** 23177–23184.

Lee, M.-E., Temizer, D. H., Clifford, J. A., and Quertermous, T. (1991). Cloning of the GATA-binding protein that regulates endothelin-1 gene expression in endothelial cells. *J. Biol. Chem.* **266,** 16188–16192.

Leonard, M. W., Lim, K.-C., and Engel, J. D. (1993). Expression of the GATA transcription factor family during early erythroid development. *Development* **119,** 519–531.

Li, X., and Noll, M. (1994). Compatibility between enhancers and promoters determines the transcriptional specificity of gooseberry and gooseberry neuro in the Drosophila embryo. *EMBO J.* **13,** 400–406.

Lieuw, K. H., Roth, M. E., Dzierzak, E., George, K. M., Karis, A., Leonard, M. W., Lim, K.-C., Pandolfi, P. P., Grosveld, F., and Engel, J. D. (1995). Expression and regulation of GATA-2 and GATA-3 in hematopoietic and other cell lineages. *In* "Biology of Hematopoiesis and Stem Cell Gene Transfer" (G. Stamatoyannopoulos, Ed.), pp. 15–35. Intercept, Andover, UK.

Maeda, M. (1994). Gastric proton pump (H^+/K^+-ATPase): Structure and gene regulation through GATA DNA-binding protein(s). *J. Biochem.* **115,** 6–14.

Marine, J., and Winoto, A. (1991). The human enhancer-binding protein GATA3 binds to several T-cell receptor regulatory elements. *Proc. Natl. Acad. Sci. USA* **88,** 7284–7288.

Ng, A., George, K. M., Engel, J. D., and Linzer, D. I. H. (1994). A GATA factor is necessary and sufficient for transcriptional regulation of the murine placental lactogen I gene promoter. *Development* **120,** 3257–3266.

Oosterwegel, M., Timmerman, J., Leiden, J., and Clevers, H. (1992). Expression of GATA-3 during lymphocyte differentiation and mouse embryogenesis. *Dev. Immunol.* **3,** 1–11.

Pandolfi, P. P., Roth, M. E., Karis, A., Leonard, M. W., Dzierzak, E., Grosveld, F., Engel, D., and Lindenbaum, M. H. (1995). Targeted disruption of the GATA-3 gene causes severe abnormalities in the nervous system and fetal liver hematopoiesis. *Nature Genet.* **11,** 40–44.

Pevny, L., Simon, M. C., Robertson, E., Klein, W. H., Tsai, S.-F., D'Agati, V., Orkin, S. H., and Costantini, F. (1991). Erythroid differentiation in chimaeric mice blocked by a targeted mutation in the gene for transcription factor GATA-1. *Nature* **349,** 257–260.

Pons, G., O'Dea, R. F., and Mirkin, B. L. (1982). Biological characterization of the C1300 murine neuroblastoma: An in vivo neural crest tumor model. *Cancer Res.* **42,** 3719–3723.

Richman, J., and Mitchell, P. J. (1996). Cranofacial development: Knockout mice take one on the chin. *Curr. Biol.* **6,** 364–367.

Sawada, S., Scarborough, J. D., Killeen, N., and Littman, D. R. (1994). A lineage-specific transcriptional silencer regulates CD4 gene expression during T-lymphocyte development. *Cell* **77,** 917–929.

Schorle, H., Meiert, P., Buchert, M., Jaenisch, R., and Mitchell, P. J. (1996). Transcription factor AP-2 essential for cranial closure and craniofacial development. *Nature* **381,** 235–237.

Siegel, M. D., Zhang, D. H., Ray, P., and Ray, A. (1995). Activation of the interleukin-5 promoter by cAMP in murine EL-4 cells requires the GATA-3 and CLE0 elements. *J. Biol. Chem.* **270,** 24548–24555.

Ting, C.-N., Olson, M. C., Barton, K. P., and Leiden, J. M. (1996). Transcription factor GATA-3 is required for development of the T-cell lineage. *Nature* **384,** 474–478.

Tsai, F.-Y., Keller, G., Kuo, F. C., Weiss, M., Chen, J., Rosenblatt, M., Alt, F. W., and Orkin, S. H. (1994). An early haematopoietic defect in mice lacking the transcription factor GATA-2. *Nature* **371,** 221–226.

Tsai, S.-F., Martin, D. I. K., Zon, L. I., D'Andrea, A. D., Wong, G. G., and Orkin, S. H. (1989). Cloning of cDNA for the major DNA-binding protein of the erythroid lineage through expression in mammalian cells. *Nature* **339,** 446–451.

Voronova, A. F., Buss, J. E., Patschinsky, T., Hunter, T., and Sefton, B. M. (1984). Characterization of the protein apparently responsible for the elevated tyrosine protein kinase activity in LSTRA cells. *Mol. Cell. Biol.* **4,** 2705–2713.

Whiting, J., Marshall, H., Cook, M., Krumlauf, R., Rigby, P. W. J., Stott, D., and Allemann, R. K. (1991). Multiple spatially specific enhancers are required to reconstruct the pattern of *Hox-2.6* gene expression. *Genes Dev.* **5,** 2048–2059.

Yamamoto, M., Ko, L. J., Leonard, M. W., Beug, H., Orkin, S. H., and Engel, J. D. (1990). Activity and tissue-specific expression of the transcription factor NF-E1 [GATA] multigene family. *Genes Dev.* **4,** 1650–1662.

Zimmerman, L., Lendahl, U., Cunningham, M., McKay, R., Parr, B., Gavin, B., Mann, J., Vassileva, G., and McMahon, A. (1994). Independent regulatory elements in the nestin gene direct transgene expression to neural stem cells or muscle precursors. *Neuron* **12,** 11–24.

Received for publication November 4, 1996
Accepted February 13, 1997