

# Transgenic mice containing a 248-kb yeast artificial chromosome carrying the human $\beta$ -globin locus display proper developmental control of human globin genes

(developmental regulation/locus control region)

KENNETH R. PETERSON\*, CHRISTOPHER H. CLEGG†, CLARE HUXLEY‡§, BETTY M. JOSEPHSON\*, HARALD S. HAUGEN†, TATSUO FURUKAWA\*, AND GEORGE STAMATOYANNOPOULOS\*¶

\*Division of Medical Genetics, Department of Medicine, University of Washington, Seattle, WA 98195; †Bristol-Myers Squibb Pharmaceutical Research Institute, Seattle, WA 98121; and ‡Department of Genetics, Washington University, St. Louis, MO 63110

Communicated by Stanley M. Gartler, May 7, 1993 (received for review March 26, 1993)

**ABSTRACT** Transgenic mice were generated using a purified 248-kb yeast artificial chromosome (YAC) bearing an intact 82-kb human  $\beta$ -globin locus and 148 kb of flanking sequence. Seventeen of 148 F<sub>0</sub> pups were transgenic. RNase protection analysis of RNA isolated from the blood of 13  $\gamma$ - and  $\beta$ -globin-positive founders showed that only the human  $\beta$ -globin gene was expressed in the adult founders. Studies of F<sub>1</sub> and F<sub>2</sub> fetuses demonstrated that the genes of the  $\beta$ -locus YAC displayed the proper developmental switches in  $\beta$ -like globin gene expression. Expression of  $\epsilon$ - and  $\gamma$ -globin, but not  $\beta$ -globin, was observed in the yolk sac, there was only minor  $\gamma$  and mostly  $\beta$  expression in the 14-day liver, and only  $\beta$  mRNA in the blood of the adult animals. Structural data showed that the locus was intact. These results indicate that it is now possible to dissect regulatory mechanisms within the context of an entire locus *in vivo* by using the ability to perform mutagenesis efficiently in yeast via homologous recombination, followed by purification of the altered YAC and its introduction into mice.

The functional genes of the human  $\beta$ -globin locus are arranged 5' to 3' in the order in which they are expressed during development:  $\epsilon$ ,  $\gamma$ ,  $\delta$ ,  $\beta$ . Upstream of the  $\epsilon$ -globin gene is the locus control region (LCR), a set of four erythroid-specific and developmentally stable DNase I hypersensitive sites (1–3), designated 5' HS1 to 5' HS4, that confer high-level, integration site-independent, copy number-dependent expression on globin genes in cells of the erythroid lineage (3). Another developmentally stable DNase I hypersensitive site, 3' HS1, is found 20 kb downstream from the  $\beta$ -globin gene (1–3). The developmental control of the genes of the  $\beta$  locus has been investigated with various approaches, mainly using transgenic mice. In those studies, recombinant human globin genes were used and inferences about developmental control were made from the patterns of human globin expression during development of the animals (4–11). Due to technical limitations in the size of constructs that can be manipulated *in vitro* for production of transgenic mice, only portions of the  $\beta$  locus have been used, the organization of the locus usually has been distorted, and regulatory elements which are normally at great distance from the genes were abnormally juxtaposed to the genes under investigation. The situation closest to normal is a 70-kb fragment containing the globin genes and the LCR, constructed by ligating two cosmids used to produce transgenic mice (12). However, the 3' end of this fragment lies  $\approx$ 2.4 kb 3' to the  $\beta$  gene and lacks the developmentally stable, erythroid-specific 3' HS1.

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

To analyze the control of human globin gene switching in an intact  $\beta$ -globin locus, we developed transgenic mice carrying 230 kb of the human  $\beta$  locus contained in a yeast artificial chromosome (YAC). We report here that the human  $\epsilon$ -,  $\gamma$ -, and  $\beta$ -globin genes of this YAC display correct developmental regulation; i.e., expression of the embryonic/fetal genes occurs in the embryonic and early fetal stages of development, whereas  $\beta$  gene expression is restricted to the adult stage of development. Transgenic mice carrying the human  $\beta$ -globin locus domain in a YAC provide another approach to investigate the control of globin gene switching in the context of the intact  $\beta$ -globin locus.

## EXPERIMENTAL PROCEDURES

**YAC Purification and Injection.** Growth of yeast strain AB1380 containing the human  $\beta$ -locus YAC ( $\beta$ -YAC)  $\gamma$ neoglobin, and gel purification and concentration of the  $\beta$ -YAC were performed essentially as described (13). Agarose blocks containing YAC DNA were prepared as described and DNA was separated in preparative 1% low-melting-point agarose (SeaPlaque GTG, FMC) gels by pulsed-field electrophoresis (CHEF gel, Bio-Rad) in 0.5 $\times$  TBE (44.5 mM Tris/44.5 mM boric acid/1 mM EDTA) at 14°C. The gels were run at 160 V with a switching time of 12 sec for 40 hr. An unstained YAC-containing gel slice was excised after ethidium bromide staining of marker lanes on either side of the gel to locate the YAC in the gel. The gel slice was cut into small pieces and equilibrated in 10 mM Tris-HCl, pH 7.5/250  $\mu$ M EDTA/100 mM NaCl. Wide-bore pipette tips were used to pipette YAC DNA solutions from this point on, to avoid shearing the DNA. Agarose pieces weighing 0.4 g were digested with 1 unit of  $\beta$ -agarase I (New England Biolabs) per 100 mg of agarose for 2 hr essentially as described by the manufacturer. Undigested agarose was removed by centrifugation at 12,500  $\times$  g for 15 min at room temperature and the supernatant was transferred to a new tube and stored at 4°C. The YAC DNA was concentrated by loading 400  $\mu$ l of supernatant into an Ultrafree-MC filter unit (nominal molecular weight limit, 30,000; Millipore) and centrifuging at 3000  $\times$  g for 5 min at room temperature. The volume of liquid that had passed through the filter was measured and centrifugation was continued until  $\approx$ 320  $\mu$ l had passed through. The remaining 80  $\mu$ l of DNA solution was incubated at room temperature for 1 hr in the filter, pipetted

Abbreviations: YAC, yeast artificial chromosome; LCR, locus control region; HS<sub>n</sub>, hypersensitive site *n*; PFGE, pulsed-field gel electrophoresis.

§Present address: St. Mary's Hospital Medical School, Department of Biochemical and Molecular Genetics, Norfolk Place, London, W2 1PG, United Kingdom.

¶To whom reprint requests should be addressed.



ately adjacent the vector–insert junction on the left, the other specific for vector sequence immediately flanking the right vector–insert junction, demonstrated that 10 of the 13 founders with both the  $\gamma$  and  $\beta$  genes also contained both left and right vector-arm sequences (Table 1).

**$\beta$ -YAC Founders Express Only Adult Human Globin.** RNase protection analysis of RNA isolated from the blood of the 13  $\gamma\beta$ -positive founders and the one  $\gamma$ -positive-only founder (no. 13) was done with probes specific for human  $\epsilon$ -,  $\gamma$ -, and  $\beta$ -globin mRNA and for mouse  $\zeta$ - and  $\alpha$ -globin mRNA. Only  $\beta$ -globin gene expression was detected (Fig. 1B).  $\gamma$  mRNA was not detected and probably amounted to <0.5% of human globin mRNA. Thus, the  $\beta$ - and  $\gamma$ -globin genes of the  $\beta$ -YAC displayed the proper developmental regulation, since there was predominant  $\beta$ -globin gene expression in the adult animal.

Three founders (nos. 5, 6, and 9; see Fig. 1B) which retained the  $\beta$  gene had minimal  $\beta$  gene expression (1% or lower), perhaps indicating that these transgenes contained rearranged  $\beta$ -YACs that had lost the LCR or other regulatory sequences.  $\beta$ -Globin mRNA in the remaining founders ranged from 6% to 80% of endogenous mouse  $\alpha$ -globin mRNA (uncorrected for copy number) (Table 1).

**Structure of the Integrated  $\beta$ -YAC.** Of the founders which appeared to harbor an intact human  $\beta$ -globin locus (by having both vector–insert junctions,  $\gamma$  and  $\beta$  gene sequences, and  $\beta$ -globin expression), we chose the line of founder 10 for a detailed analysis to confirm the structural integrity and copy number of the  $\beta$ -YAC insert. To examine whether the  $\beta$  locus of the YAC is intact, PFGE of *Sfi* I digests and diagnostic Southern blot analysis for LCR and globin gene fragments were done. DNA prepared from livers of F<sub>1</sub> transgenic mice was subjected to PFGE and Southern blot analysis using a human  $\Delta\gamma$  fragment probe. The predicted 140-kb *Sfi* I fragment encompassing the entire  $\beta$ -globin cluster and most of the upstream controlling region was detected (Fig. 2A). The presence of this large fragment, coupled with the results of PCR analysis that confirmed the presence of YAC sequences immediately adjacent to the left and right insert–vector junctions, indicated that the YAC was not internally deleted. The  $\approx$ 60-kb band present in lanes 5 and 6 of Fig. 2A may be due to fragmentation of DNA during handling or a second partially deleted copy of the YAC. Additional Southern blot analyses were performed to further confirm the structural integrity of the  $\beta$  locus. Expected 10.6- and 6.5-kb *Eco*RI LCR, 3.8-kb *Eco*RI  $\epsilon$ , 5.5-kb *Eco*RI  $\beta$ , and 2.4-kb *Eco*RI  $\gamma$  fragments were found (Fig. 2B). YAC copy number was ascertained by quantitation (PhosphorImager, Molecular Dynamics) of the  $\gamma$  bands in transgenic DNA and in samples

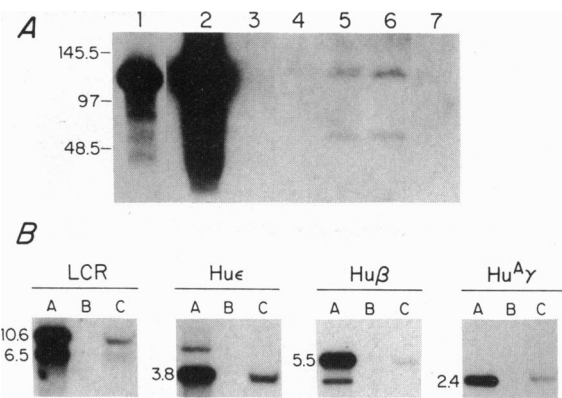


FIG. 2. Structure of the human  $\beta$ -globin locus of the  $\beta$ -globin YAC. (A) PFGE analysis confirms the presence of the 140-kb *Sfi* I fragment bearing the  $\beta$ -globin locus. Lanes 1 and 2, yeast containing the  $\beta$ -YAC (two exposures); lane 3, mouse erythroleukemia (MEL) cell negative control; lane 4, nontransgenic 17-day F<sub>1</sub> liver; lanes 5 and 6, 17-day livers from  $\beta$ -YAC F<sub>1</sub> transgenics; lane 7, undigested DNA of F<sub>1</sub> transgenic liver. DNA of lanes 1–6 was digested with *Sfi* I. A 2.4-kb *Eco*RI 3'  $\Delta\gamma$  fragment was used as a probe. (B) Southern blot analysis showing diagnostic fragments for different regions of the human  $\beta$ -YAC. All samples were digested with *Eco*RI. The probe utilized in hybridization is indicated above each panel. The expected fragment sizes in kilobases for each digest/probe combination are shown to the left of each panel. Lanes: A, yeast containing human  $\beta$ -globin YAC; B, 9-day nontransgenic F<sub>1</sub>; C, 9-day transgenic F<sub>1</sub>.

containing 1, 10, 100, or 1000 copies of the human  $\Delta\gamma$  gene. Comparison of total counts in the transgenic mouse  $\gamma$  band to a standard curve of total counts versus copy number gave an estimated copy number of 1 for three F<sub>1</sub> fetuses, 2 for one F<sub>1</sub> fetus, and 3 for the founder of line 10. In this line human globin expression, per copy of transgene, was 69% of the expression of the endogenous murine  $\alpha$  genes.

**Correct Developmental Regulation of the Globin Genes of the  $\beta$ -YAC.** To investigate the developmental control of the genes of the  $\beta$ -YAC, F<sub>1</sub> and F<sub>2</sub> fetuses were obtained from timed pregnancies of line 10. RNA isolated from yolk sac and blood of 9-, 9.5-, and 11-day embryos and liver and blood of 14-day fetuses was subjected to RNase protection analysis. There was  $\epsilon$  and  $\gamma$  mRNA but no  $\beta$  mRNA in the cells of 9-, 9.5-, and 11-day yolk sac samples (Fig. 3 *Top* and *Middle*). In the 14-day fetus,  $\epsilon$  mRNA was present in the peripheral blood (which still contains embryonic erythrocytes) but not in the liver (which is the organ of adult erythropoiesis) (Fig. 3 *Middle*). There was some  $\gamma$  mRNA but mostly  $\beta$  mRNA in the erythroid cells of the 14-day fetal liver and predominantly  $\beta$  mRNA in the adult. Thus the human  $\epsilon$  and  $\gamma$  genes behaved like the embryonic  $\beta^{H1}$  and  $\epsilon$  genes of the mouse whereas the human  $\beta$  gene behaved like the adult  $\beta^{maj}$  gene of the mouse (Fig. 3 *Bottom*). These results demonstrated that all the globin genes of the YAC displayed correct developmental regulation.

To assess globin gene expression at the cellular level, we stained yolk sac, fetal liver, and adult blood preparations with fluorescent anti- $\epsilon$ , anti- $\gamma$ , or anti- $\beta$  monoclonal antibodies.  $\epsilon$ -Globin was restricted to the cells of primitive erythropoiesis, whereas  $\beta$  gene expression was restricted to the cells of definitive erythropoiesis. There was abundant  $\gamma$ -globin in yolk sac red cells (Fig. 4A), as well as in a proportion of definitive erythroblasts in the fetal liver (Fig. 4B).  $\beta$ -Globin was present in the majority of fetal liver erythroblasts (Fig. 4C) and also the red cells of the adult animals (Fig. 4E). A few of the red cells of the adult animal also contained  $\gamma$ -globin (Fig. 4D), resembling the pattern of  $\gamma$  gene expression in the adult humans who have a small number of red cells containing  $\gamma$ -globin, called F cells.

Table 1. Characteristics of  $\beta$ -YAC transgenic founders

Line	Vector–insert junction		DNA		$\beta/\alpha$ mRNA, %
	Left	Right	$\gamma$	$\beta$	
1	+	+	+	+	10
2	+	+	+	+	48
3	–	–	+	+	6
4	+	–	+	+	18
5	+	–	+	+	0.5
6	+	+	+	+	0.5
7	+	+	+	+	7
8	+	+	+	+	80
9	+	+	+	+	1.6
10	+	+	+	+	12
11	+	+	+	+	22
12	+	+	+	+	33
13	–	–	+	–	0
14	+	+	+	+	21

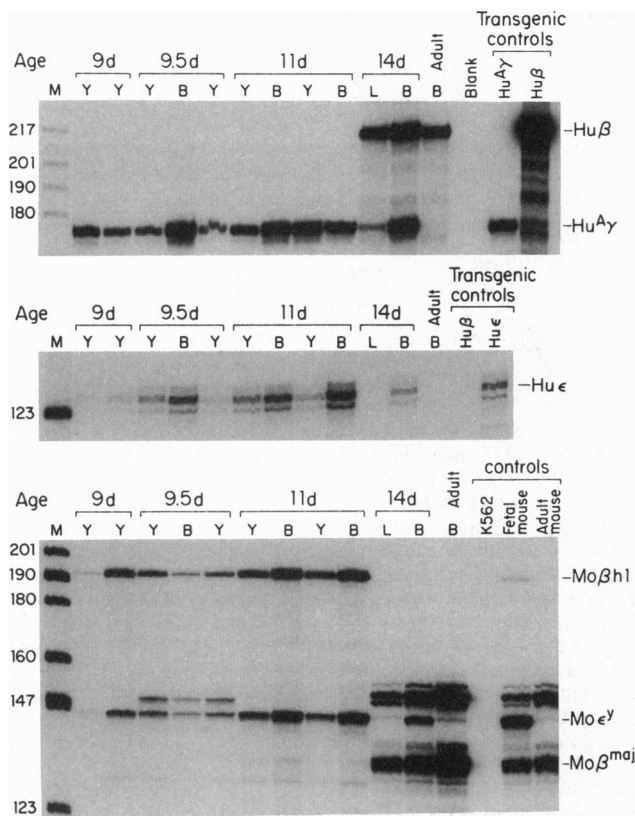


FIG. 3. Human  $\beta$ -globin gene expression is developmentally regulated in  $\beta$ -YAC  $F_1$  and  $F_2$  transgenic mice. RNase protection analysis was used to analyze expression of human  $\beta$  and  $\gamma$  (A), human  $\epsilon$  (B), and mouse embryonic  $\beta^{h1}$  and  $\epsilon^\gamma$  and adult  $\beta^{maj}$  genes (C). The age, in days (d), of prenatal samples and the tissue source are indicated above each lane (Y, yolk sac; L, liver; B, blood). Lane M, size markers (length in base pairs at left).

## DISCUSSION

Studies in transgenic mice have provided several insights on the developmental regulation of the genes of the  $\beta$ -globin locus. In these studies, constructs containing LCR sequences linked to  $\epsilon$ -,  $\gamma$ -, or  $\beta$ -globin genes or to cosmids containing these genes have been used (1, 9–12, 16, 17–20). The general concept that has emerged from this work is that two mechanisms are involved in the control of hemoglobin switching: gene silencing, represented by the turnoff of the  $\epsilon$  gene in definitive cells (17–19), and gene competition, represented by the silencing of the  $\beta$  gene in embryonic cells (10, 11, 21); both mechanisms may be required to turn off  $\gamma$  gene expression in the adult stage of development (9–11, 16, 20). All the work done so far has employed artificial constructs that lack sequences known to comprise major regions of the  $\beta$ -globin domain. Thus, it is possible that conclusions based on such artificial constructs may not always be physiologically relevant. Ideally, one would like to analyze the control of hemoglobin switching in transgenic mice in the context of the whole locus. This has been impossible for technical reasons, mainly due to the inability to manipulate, *in vitro*, the  $\approx 200$ -kb DNA sequence encompassing the human  $\beta$ -globin locus. As we show here, such experiments are now possible with the experimental system we have applied. In this paper, we demonstrate that a purified YAC containing the entire human  $\beta$ -globin locus and extensive 5' and 3' flanking sequences can be successfully introduced into fertilized mouse eggs by microinjection and transgenic mice generated. Furthermore, we demonstrate proper developmental regulation of the genes within the  $\beta$  locus contained in the YAC.

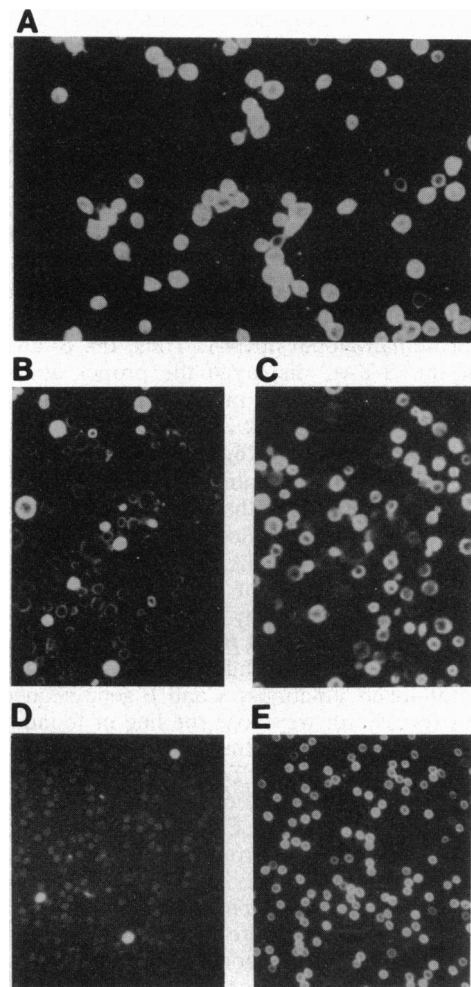


FIG. 4. Staining of erythroid cells derived from  $\beta$ -YAC  $F_1$  mice with human globin chain-specific fluorescent monoclonal antibodies. (A) Yolk sac cells at day 9 stained with anti- $\gamma$ . Virtually all embryonic erythroblasts are positive. Labeling of this population was also positive with anti- $\epsilon$ , but negative with anti- $\beta$  (data not shown). (B and C) Fetal liver cells at day 14 labeled with anti- $\gamma$  (B) or anti- $\beta$  (C). Note that only a portion of the definitive erythroblasts present as well as a few circulating embryonic erythroblasts are positive for  $\gamma$  (B), while a much higher proportion of the population is labeled with anti- $\beta$  (C). (D and E) Labeling of adult blood with anti- $\gamma$  (D) or anti- $\beta$  (E). There is uniform expression of  $\beta$ -globin (E), but only a few adult red cells express  $\gamma$ -globin (D), reminiscent of a pattern of a few "F cells" in normal humans.

Successful transfer of YACs covering multigenic loci or large single genes would allow investigation of control of gene expression with confidence that the manipulation of genes during the production of the recombinant constructs would not alter or obfuscate the control of gene expression.

Developmental expression of the globin genes of the  $\beta$ -YAC in transgenic mice showed a pattern similar, but not identical, to that of the endogenous murine globin genes (Figs. 3 and 4). The human  $\epsilon$  gene displayed exclusively embryonic expression in the yolk sac cells, while the human  $\beta$  gene displayed exclusively adult expression in cells of definitive erythropoiesis in the fetal liver and in adult red cells.  $\gamma$  gene expression was mostly embryonic, as previously observed in constructs lacking an LCR (6); this can be attributed to the evolutionary homology of the human  $\gamma$  and murine  $\beta^{h1}$  genes, both of which derive from an ancestral  $\gamma$ -globin gene. The expression of the  $\gamma$  genes of the YAC, however, differs from  $\beta^{h1}$  gene expression because it continues in the erythroblasts of adult (definitive) erythropoiesis

in the fetal liver (Fig. 4B) and in a small population of adult red cells (Fig. 4D). This suggests that the murine liver contains trans-acting factors that can activate the  $\gamma$  gene, a conclusion that is supported by previous work in transgenic mice (9–11, 16), and that the  $\gamma$  gene contains cis elements that respond to these factors. Thus, fetal recruitment of the  $\gamma$  genes probably evolved by a combination of sequence changes surrounding the genes and changes in the fetal liver trans-acting environment during evolution (6, 12).

The correct developmental control of the genes of this  $\beta$ -YAC suggests that this system can have multiple applications in the analysis of globin gene expression during development and differentiation. A major advantage of the system is the ability to manipulate YAC DNA in yeast by taking advantage of the yeast's inherent high frequency of recombination. Specific mutations can be introduced into the  $\beta$  locus and the effect of these directed mutations on developmental regulation of the globin genes can be studied in transgenic mice. For example, the physiological relevance of silencers, enhancers, and other regulatory motifs that have been defined with other expression systems (including the use of short recombinant genes in transgenic mice) can be evaluated by introducing such mutations into the whole  $\beta$ -YAC and assaying how these mutations affect developmental control. Furthermore, several questions of the developmental regulation of the human  $\beta$ -globin locus can best be answered in the context of a normally organized locus. For example, the function of each hypersensitive site of the LCR, the developmental role of each individual hypersensitive site, the relationship between gene order and temporal order of gene expression, and the control of globin gene switching through gene competition can best be investigated with manipulations that do not affect the overall organization of the locus or the distance between genes and regulatory elements. Such manipulations can best be done in the context of the YAC.

Efforts to dissect developmental regulatory mechanisms in complex multigene loci have been hampered by the inability to manipulate, *in vitro*, the large, often hundred- and sometimes thousand-kilobase-sized, stretches of DNA encompassing these loci. As we show here [and also shown recently by others (22–24)] this can be done by producing transgenic mice carrying DNA from YACs. Transgenic mice produced by injection of YACs will allow functional analysis of other large human loci whose investigation is limited by the size requirements of current technologies used for production of transgenic mice.

This work was supported by National Institutes of Health Grants

DK45365, DK30852, and HL20899 and by a Lucille P. Markey visiting fellowship to C.H.

1. Tuan, D., Solomon, W., Li, Q. & London, J. M. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 6384–6388.
2. Forrester, W. C., Thompson, C., Elder, J. T. & Groudine, M. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 1359–1363.
3. Grosveld, F., van Assendelft, G. B., Greaves, D. R. & Kollias, G. (1987) *Cell* **51**, 975–985.
4. Magram, J., Chada, K. & Costantini, F. (1985) *Nature (London)* **315**, 338–340.
5. Townes, T. M., Lingrel, J. B., Chen, H. Y., Brinster, R. L. & Palmiter, R. D. (1985) *EMBO J.* **4**, 1715–1723.
6. Chada, K., Magram, J. & Costantini, F. (1986) *Nature (London)* **319**, 685–689.
7. Kollias, G., Wrighton, N., Hurst, J. & Grosveld, F. (1986) *Cell* **46**, 89–94.
8. Kollias, G., Hurst, J., deBoer, E. & Grosveld, F. (1987) *Nucleic Acids Res.* **15**, 5739–5747.
9. Enver, T., Ebens, A. J., Forrester, W. C. & Stamatoyannopoulos, G. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 7033–7037.
10. Enver, T., Raich, N., Ebens, A. J., Papayannopoulou, T., Costantini, F. & Stamatoyannopoulos, G. (1990) *Nature (London)* **344**, 309–313.
11. Behringer, R. R., Ryan, T. M., Palmiter, R. D., Brinster, R. L. & Townes, T. M. (1990) *Genes Dev.* **4**, 380–389.
12. Strouboulis, J., Dillon, N. & Grosveld, F. (1992) *Genes Dev.* **6**, 1857–1864.
13. Gnirke, A., Huxley, C., Peterson, K. & Olson, M. V. (1993) *Genomics* **15**, 659–667.
14. Gaensler, K. M. L., Burmeister, M., Brownstein, B. H., Tailon-Miller, P. & Myers, R. M. (1991) *Genomics* **10**, 976–984.
15. Elder, J. T., Forrester, W. C., Thompson, C., Mager, D., Henthorn, P., Peretz, M., Papayannopoulou, T. & Groudine, M. (1990) *Mol. Cell. Biol.* **10**, 1382–1389.
16. Dillon, N. & Grosveld, F. (1991) *Nature (London)* **350**, 252–254.
17. Raich, N., Enver, T., Nakamoto, B., Josephson, B., Papayannopoulou, T. & Stamatoyannopoulos, G. (1990) *Science* **250**, 1147–1149.
18. Raich, N., Papayannopoulou, T., Stamatoyannopoulos, G. & Enver, T. (1992) *Blood* **79**, 861–864.
19. Shih, D. M., Wall, R. J. & Shapiro, S. G. (1990) *Nucleic Acids Res.* **18**, 5465–5472.
20. Lloyd, J. A., Krakowsky, J. M., Crable, S. C. & Lingrel, J. B. (1992) *Mol. Cell. Biol.* **12**, 1561–1567.
21. Hanscombe, O., Whyatt, D., Fraser, P., Yannoutsos, N., Greaves, D., Dillon, N. & Grosveld, F. (1991) *Genes Dev.* **5**, 1387–1394.
22. Strauss, W. M., Dausman, J., Beard, C., Johnson, C., Lawrence, J. B. & Jaenisch, R. (1993) *Science* **259**, 1904–1907.
23. Jakobovits, A., Moore, A. L., Green, L. L., Vergara, G. J., Maynard-Currie, C. E., Austin, H. A. & Klapholz, S. (1993) *Nature (London)* **362**, 255–258.
24. Schedl, A., Montoliu, L., Kelsey, G. & Shutz, G. (1993) *Nature (London)* **362**, 258–261.