© 1990 Oxford University Press
Nucleic Acids Research, Vol. 18, No. 15 4361

Molecular mechanisms underlying the expression of the human HOX-5.1 gene

Luciano Cianetti*, Anna Di Cristofaro, Vincenzo Zappavigna, Lisa Bottero, Giovanni Boccoli, Ugo Testa, Giovanni Russo[1], Edoardo Boncinelli[2] and Cesare Peschle
Department of Hematology and Oncology, Istituto Superiore di Sanità, Viale Regina Elena 299, 00161 Rome, [1]Clinica Ostetrica e Ginecologica, Ospedale Civile di Avellino, 83100 Avellino and [2]International Institute of Genetics and Biophysics–CNR, Via G. Marconi 10, 80125 Naples, Italy

Received May 30, 1990; Accepted June 13, 1990

EMBL accession no. X17360

ABSTRACT

The complex mechanisms underlying homeobox genes expression involve regulation at transcriptional, post-transcriptional and translational levels. The multiple transcripts of the human *HOX-5.1* gene are expressed differentially in tissue- and stage-specific patterns during embryogenesis, and differentially induced by retinoic acid (RA) in human embryonal carcinoma (EC) NT2/D1 cells. We have sequenced 6.3 Kb of the genomic region containing the *HOX-5.1* gene and analyzed its mechanisms of expression. Two alternative promoters underlie the transcription of two classes of *HOX-5.1*-specific mRNAs. These classes differ in tissue and subcellular distribution, induction by RA, structure of the 5′-UT region and mRNA stability: these features are compatible with a differential function of the two classes of transcripts in embryogenesis.

INTRODUCTION

Homeobox containing genes are highly conserved and widely distributed among Metazoa (*i.e.*, Annelids, Drosophila, Echinoderms, Xenopus, mice, and humans) (1). The homeobox is a conserved 183 bp sequence encoding a sequence-specific DNA-binding domain, which admittedly regulates gene expression (2, 3).

In Drosophila these genes, mostly organized in chromosome 3 in two clusters (the *Antennapedia* (*Antp*) and *Bithorax* (*Bx*) complex), play a key role in embryogenesis, particularly in the determination of the number, polarity and identity of body segments (4). In mice and humans *Antp*-like genes are clustered in four major complexes (murine *Hox-* and human *HOX-1*, *-2*, *-3*, and *-5*), which map in man on chromosomes 7, 17, 12 and 2 respectively (5). Growing evidence suggests that these genes play a key role in mammalian development: thus, (i) they are expressed in embryonic tissues according to tissue- and/or stage-specific patterns (6, 7); (ii) the structure/expression features of the genes in the *Antp-Bx* complex show striking similarities with those of corresponding genes in the murine *Hox-1*, *-2*, and *-5* and human *HOX-2* clusters (8, 9, 10, 11 and our unpublished observations); (iii) inappropriate expression of homeobox genes in transgenic mice leads to malformations (12, 13).

A 1.3 Kb cDNA clone (*HHo.c13*), belonging to the *HOX-5* cluster, has been isolated from a SV40-transformed human fibroblast library (14). This clone, corresponding to the *HOX-5.1* gene, encodes a homeobox-containing protein of 255 residues. Northern blot analysis revealed multiple *c13*-specific transcripts, which are differentially expressed in a variety of human embryonic tissues, *i.e.* , brain, spinal cord, backbone, heart and limb buds, often according to stage-specific patterns (14). All *c13* specific transcripts are also expressed in the retinoic acid (RA)-induced NT2/D1 human embryonal carcinoma (EC) cell line (15). The murine homologue to *c13*, isolated from an embryonic mouse cDNA library (16), is similarly expressed in multiple RNA species according to a temporally-regulated pattern, with a tissue distribution analogous to that found in the human *HOX-5.1* (16, 17).

In an attempt to clarify the complex structure of the human *HOX-5.1* gene and the origin of its multiple transcripts, we have analyzed a large genomic region containing the gene. The gene was entirely sequenced, and the structure of the transcripts determined. Transcripts are generated starting from two alternative promoters and polyadenylated at the level of at least three sites. The two classes of *HOX-5.1* transcripts, related to the two different promoters, are characterized by differential tissue and subcellular distribution, induction by RA and mRNA stability.

MATERIALS AND METHODS

Embryos and cells

Human embryos were obtained virtually intact by legal curettage abortions at 5–9 weeks after fertilization. Their age was carefully established by morphologic staging as previously described (7). Different organs and body parts were sterilely dissected under an inverted microscope.

Human pluripotent EC cells NTERA-2, clone D1 (NT2/D1) (18) were grown at 37°C in Dulbecco's modified Eagle's medium (high-glucose formulation), supplemented with 10% fetal calf serum (Flow Laboratories, Glasgow, Great Britain) in 5% CO_2,

* To whom correspondence should be addressed

in air humidified atmosphere. The cells were maintained in their undifferentiated phenotype by continuous growth at high cell density ($5-50\times10^6$ cells/175 cm^2 flask) as described (19).

Human embryonic cDNA libraries

cDNA libraries were prepared from poly(A)$^+$ RNA from 5-wk whole embryos, 7-wk spinal cord in λgt10 (Clontech, Palo Alto, CA) and placenta (20).

Genomic DNA and cDNA fragments were subcloned in M13 mp10/mp11 or pGEM 3Z/4Z (Promega, Madison, WI). DNA sequencing was performed according to Sanger et al. (21) and adapted to double-stranded plasmid DNA (22).

Induction of teratocarcinoma cells differentiation

Differentiation of NT2/D1 cells was induced by seeding cells at 2×10^6 cells per 175 cm^2 flasks in 10^{-5}M RA (Sigma Co., St. Louis, MO). After 1 week the cells were refed with an equal volume of fresh culture medium containing 10^{-5}M RA. Cultures were then exposed to RA for 2 weeks.

In order to assess the induction of neuronal differentiation (23, 24) of NT2/D1 cells by RA, we evaluated both the expression of nerve growth factor receptor (NGF-R) and the reactivity of the cells with mAbs to neurofilament proteins. The expression of NGF-R was assayed by binding studies with (^{125}I)-NGF (Amersham, Buckinghamshire, England). The expression of neurofilament proteins was examined by an immunoperoxidase technique with three different mAbs to 68, 160 and 200 kDa neurofilament polypeptides (Amersham). Control NT2/D1 cells did not express these neuronal markers, whereas cells grown for 14 days in the presence of RA were clearly positive for all of them.

RNA extraction and Northern analysis

Total cellular RNA was extracted from fresh or frozen cells and tissues by the guanidinium thiocyanate technique (25).

To extract nuclear and cytoplasmic RNA, NT2/D1 cultures were trypsinized and the cell pellet lysed in 10 mM TrisHCl pH 7.5/10 mM NaCl/3 mM $MgCl_2$/0.5% NP40. Nuclei and cytoplasms were separated by centrifugation at 1020×g 10 min at + 4°C. Total RNA was extracted from these fractions by the guanidinium thiocyanate technique (25).

Poly(A)$^+$ RNA was selected by one passage on oligo(dT)-cellulose columns, run on 1.0% agarose-formaldehyde gels,

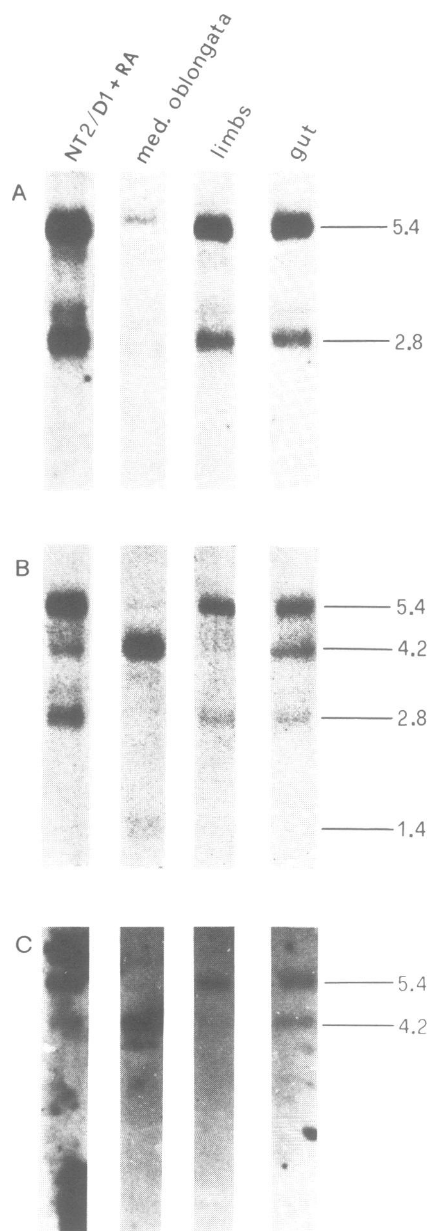


Figure 1: Northern blot analysis of poly(A)$^+$ RNA (2–3 μg/lane) from RA-induced NT2/D1 cells, human embryonal 8-wk *medulla oblongata*, 6-wk limbs, 7-wk gut. (A) Hybridization to the 0.8 Kb BamHI-HindIII fragment; the same pattern was observed for 0.5 Kb HindIII-PvuII and probe A. (B) Hybridization to probe B; probes C and D revealed the same pattern. (C) Hybridization to the 2.3 Kb XbaI-EcoRI probe; the 0.8 Kb AluI probe gave the same pattern. Sizes in Kb are shown on the right.

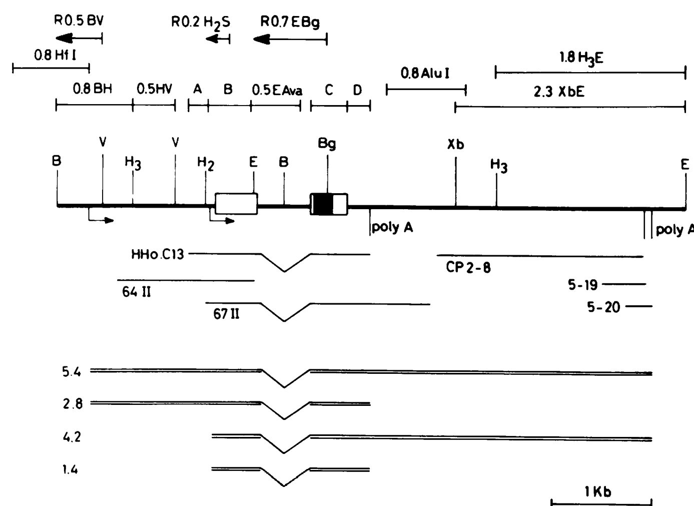


Figure 2: Genomic map of the human *HOX-5.1* gene. Boxes represent the coding region; the black box is the homeobox. Restriction sites are indicated: B, BamHI; V, PvuII; H3, HindIII; H2, HincII; E, EcoRI; Bg, BglII; Xb, XbaI; Hfl, HinfI; S, SmaI. Cap sites (bent arrows) and polyadenylation sites (vertical lines) are indicated under the map. The probes reported in the text are shown at top. Ava is AvaI. Probes A, B, C and D are derived from the *HHO.c13 cDNA* clone (see Results and ref. 12). The riboprobes (R) used in RNAase protection experiments are indicated as horizontal arrows turned in the antisense orientation. The cDNA clones analyzed are indicated below the map. Structure of the *HOX-5.1* transcripts is indicated at bottom; sizes in Kb are given on the left.

transferred to nylon membranes (Amersham, Hybond N) by Northern capillary blotting and hybridized to $1-2\times10^{7}$ cpm DNA probes labeled by nick translation or random priming. Prehybridization and hybridization were carried out as described (7). Filters were washed under stringent conditions (15 mM NaCl/1.5 mM Na citrate/0.1% SDS at 65°C) and exposed to Kodak X-OMAT S0-282 film at −70°C in an X-omatic intensifying screen cassette.

RNAase protection

RNA probes were synthesized to high specific activity from DNA fragments cloned in the pGEM 4Z vector (Promega, Madison, WI) to generate ^{32}P-labeled antisense RNAs according to standard protocols (Promega and 26). In each experiment 20 μg of total RNA were hybridized to $2-10\times10^{5}$ cpm of the probe.

Hybridization and RNAase digestion were performed according to standard procedures (26). Samples were run on a 8M urea/6−8% acrylamide gel. Gels were dried and exposed to Kodak X-OMAT S0-282 film for 2 days.

Primer extension

Total RNA (50 mg) was hybridized to 10^{6} cpm (1 pmole) of the ^{32}P-kinase labeled oligonucleotide 5′-GCGCCGGCTGCTTCT-AGTCTCCATGTCGCA-3′ in 10 μl of 10 mM Pipes pH 6.4/400 mM NaCl 16 hr at 60°C. Reverse transcription was performed for 2 h at 42°C in 100 μl of a buffer containing 50 mM TrisHCl pH 8.3/5 mM $MgCl_2$/50 mM KCl/250 μM each dNTP/10 mM DTT/10 u RNasin/20 u AMV reverse transcriptase (Boehringer). Samples were phenol extracted and run on a 8M urea/8% acrylamide gel. The gel was then dried and exposed to Kodak X-OMAT S0-282 film for 4 days.

Evaluation of HOX 5.1 transcripts half-life

mRNA half-life was evaluated in NT2/D1 cells grown for 14 days in the presence of RA. The cells were incubated for 0−8 h in the presence of 1μg/ml actinomycin D (Sigma). Control experiments showed that this concentration of actinomycin D was sufficient to inhibit > 98% (^{3}H)-uridine incorporation into trichloro-acetic-acid (TCA)-precipitable material (i.e., > 98% inhibition of RNA synthesis). At fixed times after actinomycin D addition, the culture medium was removed and cells, immediately lysed in guanidinium thiocyanate, were processed for RNA extraction as described above.

RESULTS

A 24 Kb clone (λ13G) has been isolated from a human genomic library probed with the *HHo.c13* (14). This clone largely encompasses the *c13* sequence and appears to contain the entire *HOX-5.1* gene.

The *HOX-5.1* gene is expressed in multiple transcripts, namely the 5.4, 4.2 and 2.8 Kb mRNAs (14,15). To map the position of these transcripts on genomic DNA, poly(A)$^{+}$ RNA from human embryonal tissues (14) and 14-day RA-induced NT2/D1 cells (15) were successively probed, by Northern analysis, with several restriction fragments from λ13G and *c13* (Fig. 1). The results indicate that all *HOX-5.1* mRNAs are transcribed from a region included between a BamHI site located 2.6 Kb upstream and a EcoRI site 3.5 Kb downstream to the homeobox: indeed,

```
        10        20        30        40        50        60        70        80
GGATCCTGGTGGGGGAGGGTGGTTAATAAAGCCGCCATCCTTGGGATGGATTATTTTTCTTTCTTTCTTTTTTTTCT
BamHI
        90       100       110       120       130       140       150       160
TTCTTAAGAAGAATATTCTGGTTGTTCGCCTGCTTGGTAACCCTGACCCTGGCAGAAGAATGAGGGAACTCATTGCTTCA

       170       180       190       200       210       220       230       240
AATTGTCGCCAAGCCCATTAGGCTACCTGAACTGTCTCAGAAAGTGCGGGTGGCTGCGTCGAACGGTGGTGGCTCAGAGG

       250       260       270       280       290       300       310       320
AAGAGATTGGGGCCGGCAGCGACCTAGGTACCTCACTCTGGGTGGGACCCAGAGGTTGTAACGTTGTCTATATATACCCT
                                                Ava II
       330       340       350       360       370       380       390       400
GTAGAACCGAATTTGTGTGGTATCCGTATAGTCACAGATTCGATTCTAGGGGAATATATGGTCGATGCAAAAACTTCACG
                                    Hinf I                 Δ
       410       420       430       440       450       460       470       480
TTTCTTCGGAATAGCCAGAGACCAAAGTGCGACATGGAGACTAGAAGCAGCCGGCGCTGGTCAGCCGCCTCGTTCTGTTT

       490       500       510       520       530       540       550       560
TATTACCTTGGACTCCAGGAGGATCAGCTGCGCCTGGTGACATAGAGCAGCTTTTCCTCTCCAGAAGCTCCTCACCTTTA
                            Pvu II
       570       580       590       600       610       620       630       640
AACAGAGTATCCTCTGGGTGCTGAAAAGAAAGAAAGACAGAAAGAGAGAAAGAGAGAGAGAGAGAAAGAGAGAATGCA
                                                                   └λ64II
       650       660       670       680       690       700       710       720
AGCCTAATTGGTTGCATGGATGCAGGGCCAAAGGGCTAGGTTTTGGGGTACTAGGGAGTGAGGTACAAGGCCAGCTTGCC

       730       740       750       760       770       780       790       800
CAGTCCCAGCTCTGCCCTCCAGGAACATGAGGTGCAAAGGTACCCAAATGGGGGCTTGCTTGTATTTGGGGCCTGTGGGA

       810       820       830       840       850       860       870       880
AGAAAGCAAGCTTCAAAGAAGCCCAGTGGGGAGCTCTAGGGTGCATTTTGACAAGGTGGAGGTGCCCTTGCCACCATCCC
       Hind III
       890       900       910       920       930       940       950       960
AGCCCACCCCCAGCTACATGGGCAAGGGCAGCGAGGGCCCCCTGCTATTTTGGCAGGGCCCAGCTTTGGCTGGGAACCCC

       970       980       990      1000      1010      1020      1030      1040
CGGGCCTGGGCACTGGTAGAAAGCATGGCGGTTACTCATTGCCTAATTTGATTCAAGCTGGCCAGATTCTGGTAACTTTT

      1050      1060      1070      1080      1090      1100      1110      1120
GGGTGACCCTGATGAAGACAAAGCCAGGACGGCGGCCTTTGTATGGCAGATCCCTGCTCCCGCCGGCTGCAGGCAGGGCG

      1130      1140      1150      1160      1170      1180      1190      1200
GGCAGGCAGGAACCCTCCTCGCCTGGGGCACTCTGCCCAACTCAGAGGCGAGTTCACCCACCCACCTTTCATTGCTCTGT

      1210      1220      1230      1240      1250      1260      1270      1280
ACCCCAATAGGAGGATTCATTCTCCCTTGAGCTGTGCCTACTTGGTGTCGGGGGGCGGGGGTTGCATTCAGCTGGGGGTG
                                                                      Pvu II
      1290      1300      1310      1320      1330      1340      1350      1360
AGTGGAAGGGCCACGGAAGGTTGGCAAAATCAGTGGCAGACAAAAGCTGGGATTACCTGAGGGGAATGGGGTGCTGGGGA

      1370      1380      1390      1400      1410      1420      1430      1440
CTGGAACTACATTAATATCTGGCAGGGGCTCTCAAATGTGCCATAGCAAGCTACTTGATTACACGTATGTTATTTAGTTA
        └HHo.c13
      1450      1460      1470      1480      1490      1500      1510      1520
AATTTGTGAAAATTATGAGATGCTCACCAACCCGGTGATAAACTTGCTCCCTCGCCATTGGCTGGCCTGGTCACATGGCT

      1530      1540      1550      1560      1570      1580      1590      1600
GCCCAACTTTATTCAGTTGACAGCAAGTAGGAGGGCCCTATGGAAGGAGAAAAAAAGACAACACGAGAAAAATTAGTATT
                Hinc II        Apa I              └λ67II                       Δ
      1610      1620      1630      1640      1650      1660      1670      1680
TTCTACCTTCTGAAATTAATGGTCATGAGTTCGTATATGGTGAACTCCAAGTATGTGGACCCCAAGTTCCCTCCGTGCGA
                    ΔMetValMetSerSerTyrMetValAsnSerLysTyrValAspProLysPheProProCysGl

      1690      1700      1710      1720      1730      1740      1750      1760
GGAGTATTTGCAGGGCGGCTACCTAGGCGAGCAGGGCGCCGACTACTACGGCGGCGGCGCAGGGCGCAGACTTCCAGC
uGluTyrLeuGlnGlyGlyTyrLeuGlyGluGlnGlyAlaAspTyrTyrGlyGlyGlyAlaGlnGlyAlaAspPheGlnP

      1770      1780      1790      1800      1810      1820      1830      1840
CCCCGGGGCTCTACCCACGGCCCGACTTCGGTGAGCAGCCTTTCGGAGGCAGCGGCCCCGGGCCTGGCTCGGCGCTGCCT
roProGlyLeuTyrProArgProAspPheGlyGluGlnProPheGlyGlySerGlyProGlyProGlySerAlaLeuPro
 Sma I
      1850      1860      1870      1880      1890      1900      1910      1920
GCGCGGGGTCACGGACAAGAGCCAGGCGGCCCCGGCGGTCACTACGCCGCTCCAGGAGAGCCTTGCCCAGCTCCCCCGGC
AlaArgGlyHisGlyGlnGluProGlyGlyProGlyGlyHisTyrAlaAlaProGlyGluProCysProAlaProProAl

      1930      1940      1950      1960      1970      1980      1990      2000
GCCTCCGCCGGCGCCCCTGCCTGGCGCCCGGGCCTACAGTCAGTCCGACCCCAAGCAGCCGCCCCCGGGACGGCACTCA
aProProProAlaProLeuProGlyAlaArgAlaTyrSerGlnSerAspProLysGlnProProProGlyThrAlaLeuL
                                                                    •
      2010      2020      2030      2040      2050 ▼    2060      2070      2080
AGCAGCCGGCCGTGGTCTACCCCTGGATGAAGAAGGTGCACGTGAATTCGGGTAAGGCTAGGGTCCAGTAACCTTTCTGT
ysGlnProAlaValValTyrProTrpMetLysLysValHisValAsnSer
                                                •EcoRI┘ λ64II
      2090      2100      2110      2120      2130      2140      2150      2160
CCACATCCCAGCCCGTTAGCCTGGGTCCTCTGGAAGGGGGTGCGAGTAGGTGGGGGCGTGTGGAGCTTCCATGGGCGCCG

      2170      2180      2190      2200      2210      2220      2230      2240
CAATTACTCTCCCCATAAATTTTTATAGCTGAGGGAGCAGGTCAGGACCATGTGGCTGGCTGCTCGGCTGTGGGCGCAAA

      2250      2260      2270      2280      2290      2300      2310      2320
AGGGGGTGGGGATGGGGGGGTGGGGGAGGACTCCATTTTCAGAGCAGGGGGAAGGCTGTGGAGGAGCGGGGGATTTCCAA

      2330      2340      2350      2360      2370      2380      2390      2400
AATGCTTGAGGGTTCCGGACCTGGTGGTGGGCCCAGAAGAAGGAGCACATTTGGGGATCCCGCAAGCCTGGGGTATGTGG

      2410      2420      2430      2440      2450      2460      2470      2480
GTGTGTTTGAGGAGGTGGGTGGGAGTGAGCGTGTGCGCCGGGGAGAGGGCGGGAGGGAGGAAGCAAGCGAGCTTGGGAGC

      2490      2500      2510      2520      2530      2540      2550      2560
GCGCGGGGAGGGCCGCGGGCCTCGGGGCGCGCCAGGAAGTGAGCGGCGGAGGCGAGGGGCCTAACTAGTGGCCGGGCGCT
                    Ava I
      2570      2580      2590 ▼    2600      2610      2620      2630      2640
GACCTGCCTGTCCTGTCTGTTTTGTCTCGCAGTGAACCCCAACTACACCGGTGGGGAACCCAAGCGGTCCCGAACGGCCT
                                ValAsnProAsnTyrThrGlyGlyGluProLysArgSerArgThrAlaT
                                                           homeobox
      2650      2660      2670      2680      2690      2700      2710      2720
ACACCCGGCAGCAAGTCCTAGAACTGGAAAAAGAATTTCATTTTAACAGGTATCTGACAAGGCGCCGTCGGATTGAAATC
yrThrArgGlnGlnValLeuGluLeuGluLysGluPheHisPheAsnArgTyrLeuThrArgArgArgArgIleGluIle

      2730      2740      2750      2760      2770      2780      2790      2800
GCTCACACCCTGTGTCTGTCGGAGCGCCAGATCAAGATCTGGTTCCAGAACCGGAGGATGAAGTGGAAAAAAGATCATAA
AlaHisThrLeuCysLeuSerGluArgGlnIleLysIleTrpPheGlnAsnArgArgMetLysTrpLysLysAspHisLy
                                  Bgl II
      2810      2820      2830      2840      2850      2860      2870      2880
GCTGCCCAACACTAAAGGCAGGTCATCGTCCTCATCTTCCTCCTCATCTTGCTCCTCCTCAGTCGCCCCCAGCCAGCATT
sLeuProAsnThrLysGlyArgSerSerSerSerSerSerSerSerSerCysSerSerSerValAlaProSerGlnHisL

      2890      2900      2910      2920      2930      2940      2950      2960
TACAGCCGATGGCCAAAGACCACCACACGGACCTGACGACCTTATAGAAGTGGGGACCCTGGGCCCATCTCTCCCTGCGC
euGlnProMetAlaLysAspHisHisThrAspLeuThrThrLeu                 •          Apa I

      2970      2980      2990      3000      3010      3020      3030      3040
ACCAGGCTGAGCCGAAGCTGCGGGGCAGGCCGGGCCTGCTGTCACCTCGCTGGGCTCTAAGGTACTGTGGGGTGGACCTG
                                                                   ••
```

```
      3050       3060       3070       3080       3090       3100       3110       3120
GGACAAGCAGGCCGCCCTCGGACTAGGTTAGCATCCTGCCCGAGGGCAGCCCCCTCCCTAGAGCGGGATGGGGATGGGAG
                                              ****
      3130       3140       3150       3160       3170       3180       3190       3200
GGGGGGCGGGATTCTCTCTCTAAGTATATTATATGGCAGGAGCTACTGAGAACATAAAATCTTGGCGAGTC[ATTAAA]CTT
                                 *
      3210       3220       3230       3240       3250       3260       3270       3280
ATGAAAATCACCGCTCTTGGATTTTGAATTTGCAAATGAAGGTTGGATGCTTTATCCCACTGTGAATTTGGACATTCTCC
      ** HHo.c13
      3290       3300       3310       3320       3330       3340       3350       3360
CCCACTCCACCCCTCCAGGGTGCTTTGTGGCTTAATAATGTGGGGGAGTTGAGGCAGAAGGTTGGCCACCCCTGTGCTAG

      3370       3380       3390       3400       3410       3420       3430       3440
GTGCTTTCAGTGGAAGCCAGAGAGCTGGGTCAGGATTTCTGGACTTTCTGGGTTGTCTATGGAATTTCATGTGATTAAAA
                        AluI
      3450       3460       3470       3480       3490       3500       3510       3520
AATATATATTTTGCTCCCAGTGGCCCCACCTCCAAAGAAATGGGTCTAAGAAGGAAGTAAAAATGGGTTATTTTATGTTT

      3530       3540       3550       3560       3570       3580       3590       3600
AGATATTTGCTTAAATATTTATTTGTTGGGAAATGTGGTACAGAAATAACTGACACCTTCATGCCAAAAATCTTAAAAAG

      3610       3620       3630       3640       3650       3660       3670       3680
GTGAAAGGGGCTGAACTTCAGGGAGCAGAATCAGAGATATGTGCACTTACTTCTGAACTCCACCCCTCCCCACTCTCTGG

      3690       3700       3710       3720       3730       3740       3750       3760
AAATGTATATAGGGGGGCCTTAACCCTTCCAGAAGGAAGCAAAGGATTCACTCAAAGTTGCATTCTTGAAAATATATTTC

      3770       3780       3790       3800       3810       3820       3830       3840
CACATGTGTTTTTTTCAGCACTGTGCTTACAACCAGTTCTGGGTGATTAAAGGAAAGGAAAAAAACCAACAAATGGTCC
                              λ67II┘                 └CP 2-8
      3850       3860       3870       3880       3890       3900       3910       3920
AACATTTTCCTTCTGGGGAAAGAAAACAAAACCTCTATGCACTGGGTCATTAGATAATGACTGAATTTTCTGTTCCAACT

      3930       3940       3950       3960       3970       3980       3990       4000
GGATTCCAAATGCCCTAAATACCCTCATATAGCAGTGTTTTACAGGAATTAGTGTATGGCCCGTGTAGGGGAGGGGCTGT

      4010       4020       4030       4040       4050       4060       4070       4080
GCAGTGGGGAGAAAGTGGGAAGGTGAGGAACTCTTGCTTTAAGAAGGAAAAAAAAAAAACCCTAATTGAATCTAGAAGTC
                                                                          Xba I
      4090       4100       4110       4120       4130       4140       4150       4160
CACAAAAGTTAGCCTTAGAGTTTTTTTCCCCCTGAAGTTTTAATTTTTTTAAAAACCAAATCTAAGGAAGTTTTCCTCAG
                                                                               Al
      4170       4180       4190       4200       4210       4220       4230       4240
CTCATTAATTAGAAGCAGAATTTGTAAAAGTATAAAAGTTTTCAAGCACTCGTCTTTGCCTTGAGAATAGTGGTTTTTTA
ul
      4250       4260       4270       4280       4290       4300       4310       4320
AAGAATCACTCTCAACAGGGGAGATGTCCTCTAGTCGTTTTTCTTCTGCCTCTCCTGGGAAGGGTTCAAAGTTCATTTTT

      4330       4340       4350       4360       4370       4380       4390       4400
CTAAAATGCTGACCCTCAAGCATAAGGAGGAAGAAGTCAAAGTTAATGGCCAGAGTTCATATACTCAGATGAAACCAGTC

      4410       4420       4430       4440       4450       4460       4470       4480
TTCCCAAGGCCTCAGGCTCCAAAAAAGGTTGTAGCTATCAAAAAGTGACCAAAGTGGGAAAGGGAGAAAGGATAAGCTTA
                                                                           HindIII
      4490       4500       4510       4520       4530       4540       4550       4560
AAATTTAATTTTAAGATCCAGAAGGGGGGTATTTTTTTCAGTACTTCAAAAACACTTTAGAAGGTTTCTGTTGTAATTTA

      4570       4580       4590       4600       4610       4620       4630       4640
AAAAATATATTTAAGTGGGAGGGAAAAAAGAGTTTCTCTGTAGGCTTGTTCTTTGGCTGTGTCTCCTGAGAGCTGAGGGC

      4650       4660       4670       4680       4690       4700       4710       4720
AGGTATTCACTGCAGTCCCTAGGCTGAAATTCCGCTTCTCTGAAGTGTCTTCCAAGCCTTGGTCTTTTGTATTAGACCCT

      4730       4740       4750       4760       4770       4780       4790       4800
GGGGACTGCTCTTTGTTTCTCCTTGGGGTGAGCCTGGCTCTCAGACTTGCACATGGCAATACTTGAATGTCACCACGTCG

      4810       4820       4830       4840       4850       4860       4870       4880
GGATATTAAAGATGGATATTCGTGCATTATTCACATCATTGTTTCTATGACAAAAAGCACAGAGTTCATACATAGTCAAG

      4890       4900       4910       4920       4930       4940       4950       4960
ACGTCTTTTTCTGACGCCCTCACGTTGAGAAGCTGAAAAGGTATTTTACCGAAGTTCGGGTAAATTACAGAATCAGGTTC

      4970       4980       4990       5000       5010       5020       5030       5040
ATCCAGAGGACAAATTTTCTATTTGATTAGCTGTATTTCAGCCGGAGGACTGACCTCTAAACCCCTAACCTTTTGGACT

      5050       5060       5070       5080       5090       5100       5110       5120
CTAACTACCCTTCTCTTCTTTTTCCTCTCTAACATGGAGAGCAGTCTTTGGATGTACCATTTGAAAGGAGCCGCTATCC

      5130       5140       5150       5160       5170       5180       5190       5200
TTAGGCAAGTTGGAAAGTTGCTAAGCTGCTTTCCTAAAACCCAAATCTGTCTATACATTGAACCTTCTCTTTGAGAAGGG

      5210       5220       5230       5240       5250       5260       5270       5280
GAAAAAGGTATATATTTTCACAACATCCAATTACATATATATAATAGAGATTTGTTGTATAGATTTCCCCCACCTCAGA

      5290       5300       5310       5320       5330       5340       5350       5360
AGTTCAGGTTACTCACCCCCAGTTTCATACCAAATGCCACACAGGCTTAACTGACTGCATCCCTGCCCCAGAGGAAAGCC

      5370       5380       5390       5400       5410       5420       5430       5440
AAGAAACATGTTTTCATGAGGAAAACCCAAGCTCCTTCTCAAACATAGCCCACTACTTTGGAAAGTAACTTAATCAGAG

      5450       5460       5470       5480       5490       5500       5510       5520
AAACAACTTCTTTGTTTATAAGTCTCAGCTCTCCTTCCTTCAGCTTGGAGGGATTCTTTTGAAATGTTAATGGAGCCTGGAT

      5530       5540       5550       5560       5570       5580       5590       5600
GGCCCAGAGTGCAGCCCCCAACCCTGAGGTCCCAGTCGGACCCCAGCATCCATTTGGGCCCACAGGAGTGGGCCAGGGAA
                      └λ5-19
      5610       5620       5630       5640       5650       5660       5670       5680
GGGGTAGGGCCCCGTAACCACTTAGGGCAGGGAAGGAAATGGGTTTCCATCTGAGAACGTGCTTTGGAGAAAGCTAGGTG

      5690       5700       5710       5720       5730       5740       5750       5760
TGGAAAAGCTCCAATGCCCATTTGCTATTATTTGTTTCCAGTTTGTTCCTTTAAATATGAGCCAGAAGTGTTTGTGTTGG

      5770       5780       5790       5800       5810       5820       5830       5840
TGTTTTAAAAACAAAAACAAAAACCGTGTTGGGGTCCTGACTGGGGGAGGGGGAGAGTGAAGTGTTTGCTGAGGACATTG
     └λ5-20
      5850       5860       5870       5880       5890       5900       5910       5920
CTCCTCTGACTCCCATCTCACTTTGTCCATCGCAGCCTTTTGTTGGGAGATGACACTGTCAGTCAGCCCATGATGTCTGT

      5930       5940       5950       5960       5970       5980       5990       6000
TCACACGAGATGCTTTTTTAATAGAATTGACCAATGTTTTGCTGCCACTG[ATTAAA]GTATTATTTATACTAATTGTTGCT
                                                               CP 2-8┘ ˈλ5-19
      6010       6020       6030       6040       6050       6060       6070       6080
TGTAGTTTTGATGTAATTCATTGATCTATATTTAAAAT[AATAAA]AGGTGTAGCAAAATCTCCCTCCTGTTTGGTGCCTTA
                                                      ˈλ5-20
      6090       6100       6110       6120       6130       6140       6150       6160
ACAGAAGCATTCATCCTTTGTTAAGTCTTCTAAAAGCTAACATTTAAACATAAACAAGTTATTATTTTCTGC[AATAAA]TTA

      6170       6180       6190       6200       6210       6220       6230       6240
GGCACCATTTTTTGGGGGGTGCCTAAAGTGTGAAGGTTAAACCATGTAAGGCTTAGCAATTCTATTATTACCACCTCCTT

      6250       6260       6270       6280       6290       6300
AATGTACACACACTCCCAGTTGGCCACCATATTTTGTGAGCATTGGGAAGCCTGGGGTTGAATTC
                                                               EcoRI
```

Figure 3: Nucleotide sequence of the *HOX-5.1* gene. Restriction sites reported in the text are shown. cDNA boundaries are indicated. TATA box, polyadenylation signals and the homeobox are boxed; poly-A addition sites are indicated by vertical arrowheads, followed by the name of the cDNA clones polyadenylated at those sites. Black triangles above the sequence indicate the splice sites. The open triangles below the sequence indicates the cap sites. The oligonucleotide used in the primer extension experiment is underlined. Nucleotide positions that do not agree with those reported for the *HHo.c13* clone (14) are indicated by asterisks.

probes external to this region failed to hybridize to any *HOX-5.1* transcript. The 0.8 Kb BamHI-HindIII (1−808, in Fig. 3), 0.5 Kb HindIII-PvuII (808−1269) and A (0.2 Kb from the 5′ end of *c13* to ApaI, 1370−1553) (14) probes detect both the 5.4 and 2.8 Kb transcripts (Fig. 1A). Using the B (0.5 Kb ApaI-EcoRI, 1553−2044), C (0.35 Kb EcoRI-ApaI, 2044−2941 excluding the intron) and D (0.25 Kb ApaI-3′ end of *HHo.c13*, 2941−3209) (14) probes both the above mentioned RNAs and an additional 4.2 Kb transcript were detected (Fig. 1B). A faint 1.4 Kb band was detected using the same probes, but only in RA-induced NT2/D1 cells and embryonal medulla oblongata. When probes downstream to the *c13* poly-A site were used, *i.e.*, 0.8 Kb Alu I (3383−4159) and 2.3 Kb XbaI-EcoRI (4071−6300), only the 4.2 and 5.4 Kb bands were observed (Fig. 1C): these data suggest that the 5.4 and 2.8 Kb transcripts contain a long region upstream to the reported ATG, while the 5.4 and 4.2 Kb transcripts largely extend 3′ to the *c13* poly-A site (Fig. 2).

The 6.3 Kb region encompassed by the BamHI and EcoRI sites was completely sequenced (Fig. 3). The genomic sequence of the region corresponds to the *c13* sequence, with the exception of: (i) two sites in the 123 and 142 codons in the first coding exon (positions 1985 and 2043 in Fig. 3), where nucleotide substitutions determine respectively changes of Pro and Val in the Ser and Ala residues previously reported (14), and (ii) 10 positions in the 3′ untranslated (UT) region (see asterisks in Fig. 3). The aminoacid 123 is also Ser in the λ64II and λ67II cDNAs (see below). All of the 12 new positions in the human genomic sequence agree with the mouse *Hox-5.1* sequence (16). Three cDNA libraries were screened with *HOX-5.1*-specific probes. When a 3′ probe was used (1.8 Kb HindIII-EcoRI, 4474−6300) a 2169 bp clone, *cp2−8*, from a placenta library and two shorter clones, λ5−20 and λ5−19, 288 and 449 bp long respectively, from a 5-wk human embryo library were isolated. By restriction mapping and partial or total sequencing, the clones were found to be colinear to the genomic sequence in the 3 ‘ region (Fig. 2). Moreover, λ5−20 and λ5−19 were polyadenylated at two different positions (6053 and 5990 in Fig. 3), 63 bp apart from one another, 9 and 14 bp respectively downstream from canonical polyadenylation signals (27). An additional AATAAA sequence is located at position 6152 in Fig. 3, possibly representing the signal for a third poly-A site in a multiple poly-A region. All three poly-A signals are followed, at a 10−20 bp distance, by a typical GT-rich stretch (27). This region appears to constitute the 3′ termination of the 5.4 and 4.2 mature transcripts. When probe B was used, two clones, λ64II and λ67II, were isolated from a 7-wk spinal cord library. These clones were characterized as described above (Fig. 2). λ64II is a 1424-bp clone, colinear to the genomic sequence in the 5′ UT and first coding exon. λ67II is a 1685-bp clone spliced as *c13*, but extending 587 bp 3′ to it. Neither clone is polyadenylated. As indicated by the colinearity

of the described cDNA to the genomic sequence, no other introns appear to be spliced out in primary transcripts from the *HOX-5.1* transcription unit other than the 541 bp (Figs. 2 and 3).

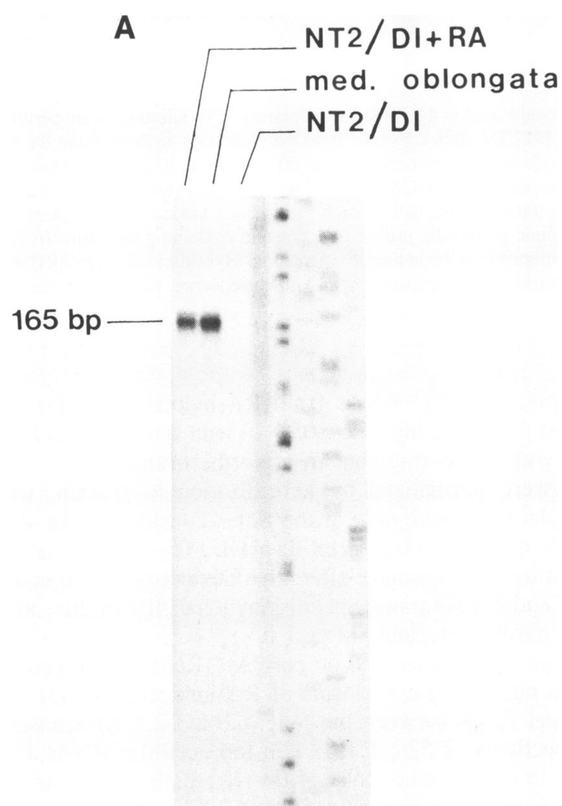
To verify that this intron is spliced in the same way in all transcripts, a 711 bp EcoRI-BglII RNA probe encompassing the intron sequence was hybridized in a RNAase protection experiment to 20 μg total RNA samples from: (a) RA-induced NT2/D1 cells, (b) 8-wk embryo *medulla oblongata*, (c) NT2/D1 stem cells. In both the (a) and (b) samples, the first one expressing all *HOX-5.1* transcripts and the other only the 4.2 kb messenger, we detected a 165 bp band (Fig. 4A), corresponding to the distance from the *c13* splice site and the BglII site in the homeobox. No bands were detected in sample (c). If we further consider that *c13* and λ67II appear to correspond to partial copies of the 2.8 and 5.4 Kb transcripts respectively, we may conclude that the 541 bp intron is similarly spliced in all mature transcripts of the *HOX-5.1* gene.

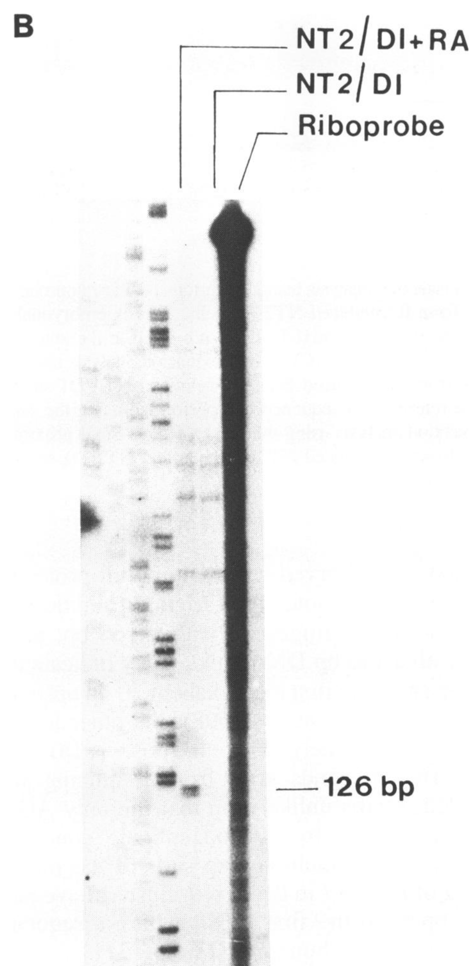
Nevertheless, when we used a 457 bp EcoRI-AvaI (2044−2501) probe containing most of the intron sequence, we detected only two faint bands of approximately 6.0 and 3.2 Kb in poly(A)$^+$ RNAs from RA-induced NT2/D1 cells, embryo limbs and gut (data not shown). These bands are present in nuclear but not in cytoplasmic poly(A)$^+$ RNA from RA-induced NT2/D1 cells (results not presented): thus, they may represent precursor forms, polyadenylated but not yet spliced, of the 5.4 and 2.8 Kb mRNAs.

Positions of the distal cap site pertaining to 5.4 and 2.8 Kb transcripts were determined by both RNAase protection and primer extension assays. Total RNA from RA-induced or uninduced NT2/D1 cells was used for both assays. The RNAase protection experiment, performed with a BamHI-PvuII 504 bp (1−505) or a AvaII-PvuII 220 bp (285−505) RNA probe, showed in both cases an RNA-protected fragment co-migrating with a 126-bp DNA band in the RA-induced but not in the stem cell sample (Fig. 4B). In this regard, evaluation of the size of RNA fragments may be up to 10% approximate with respect to DNA size markers (26). In the primer extension experiment, when a 30 bp oligonucleotide complementary to the 428−457 sequence (underlined in Fig. 3) was used and hybrids were extended with AMV reverse transcriptase, a band of 81 bp was detected in RA-induced samples (Fig. 4C). Therefore, the distal cap site may be located at the T at residue position 377 (Fig. 3) or, more likely, at the preceding A: indeed, the cap site is most often an A residue surrounded by pyrimidines (28) and a one residue shorter cDNA may be generated by methylation of the first nucleotide of the corresponding mRNA (29).

This A residue is 26 bp 3′ to a TATA sequence. Furthermore, no bands were detected when poly(A)$^+$ RNA from RA-induced NT2/D1 cells was hybridized to an excess of a ^{32}P-labeled HinfI-HinfI 0.8 Kb probe (0.8 HfI in Fig. 2), spanning from a HinfI site 0.45 Kb upstream from the BamHI site (pos. 1) to the HinfI site at position 357 (data not shown).

From hybridization experiments (see the A probe in Fig. 1) the position of the proximal cap site, pertaining to the 4.2 and possibly the 1.4 Kb transcripts, appears to be located close to the first ATG. An RNAase protection experiment was performed: a 254 bp RNA probe containing the 226 bp antisense sequence from the HincII site (pos.1536) to the SmaI site (pos.1762) was hybridized to 20 μg of total RNA from RA-induced NT2/D1 cells and 20 μg from NT2/D1 stem cells (Fig. 4D). In the RA-induced





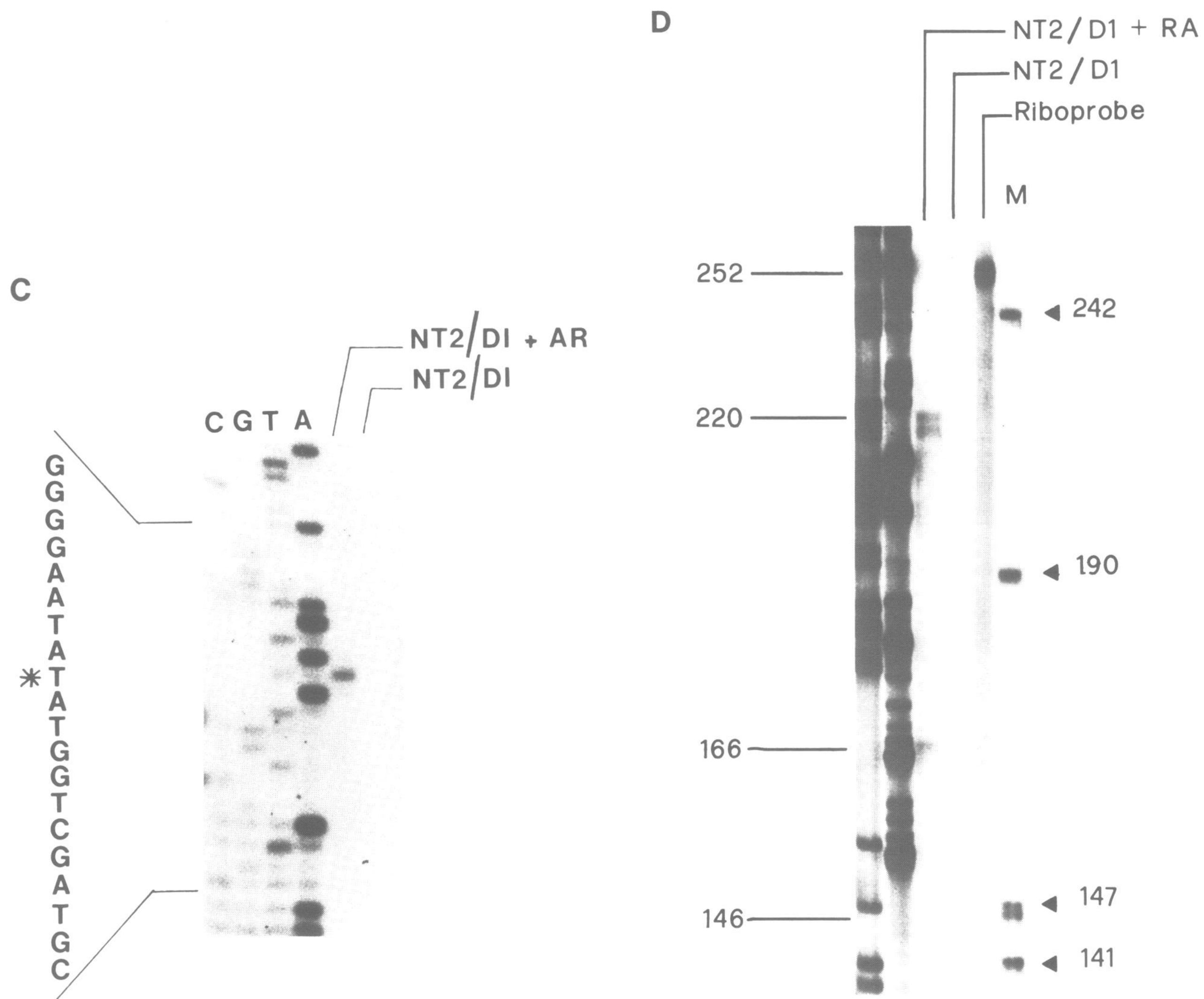


Figure 4: (A) RNAase protection analysis using the intron-specific riboprobe 0.7 Kb EcoRI-BglII synthesized in the antisense orientation. Hybridizations were performed to 20 μg of total RNA from RA-induced NT2/D1 cells, 8-week embryonal *medulla oblongata*, NT2/D1 stem cells. (B) RNAase protection analysis using the distal cap site-specific riboprobe 0.5 Kb BamHI-PvuII, synthesized in the antisense orientation. Hybridizations were performed to 20 μg of total RNA from RA-induced NT2/D1 cells and NT2/D1 stem cells. (C) Primer extension analysis using the 30 bp oligonucleotide (position 428−457 in Fig. 3). Hybridizations were performed to 50 μg of total RNA from RA-induced NT2/D1 cells and NT2/D1 stem cells. After reverse transcription, half of each sample was loaded on a 8M urea / 8% acrylamide gel. As size reference, a sequence was performed using the same 30 bp oligonucleotide as specific primer on a plasmid containing the entire *HOX-5.1* gene. (D) RNAase protection analysis using the 226 bp HincII-SmaI proximal cap site-specific riboprobe in the antisense orientation. Hybridizations were performed to 20 μg of total RNA from RA-induced NT2/D1 cells and NT2/D1 stem cells. M is a size marker. In all experiments a known sequence was run together with the samples as size reference.

sample a 220 bp band was observed representing full-protection by the probe of the distally-promoted transcripts. In addition two bands are present: the first comigrating with a 166 bp, and a second fainter one with a 146 bp DNA band. They indicate two possible proximal cap sites, the first located about 20 bp upstream from the ATGs (probably at positions 1598), the other located at the first ATG or immediately 5′ to it (pos. 1618). The possibility that the 166 bp bands arise from a splicing site, although not excluded, seems unlikely, in that the only AG in the region is not surrounded by a good splicing consensus sequence (30). Furthermore, multiple cap sites of the murine *Hox-1.4*, the paralog of *Hox-5.1* in the *Hox-1* cluster, have been mapped 21 and 13 bp 5′ to the first ATG, within a sequence region 80% similar to that of human *HOX-5.1* (31).

To characterize the differential half-lives of *HOX 5.1* transcripts, poly(A)$^+$ RNAs from RA-induced NT2/D1 cells, incubated for increasing times (0−8 h) with actinomycin D, were probed with the B fragment in a Northern analysis.

Data were normalized by hybridization to β-actin, whose mRNA decay is negligible at the times considered. As shown in Fig. 5, the 4.2 Kb transcript displays a 60−75 min half-life, and is virtually undetectable after 4 h of actinomycin D treatment. The 5.4 and 2.8 Kb transcripts display a half-life of 20−30 min and are hardly detectable after 1 h.

The same probe was used on poly(A)$^+$ RNAs from separately prepared nuclei and cytoplasms of RA-induced NT2/D1 cells. Reciprocal ratios between the 5.4, 4.2 and 2.8 Kb transcripts are respectively: 2.52 : 1.19 : 1 in total cellular RNA; 1.14 : 2.28 : 1 in cytoplasmic RNA; 4.04 : 1 : 2.30 in nuclear RNA (Fig. 6). From these results the 5.4 and 2.8 Kb transcripts appear

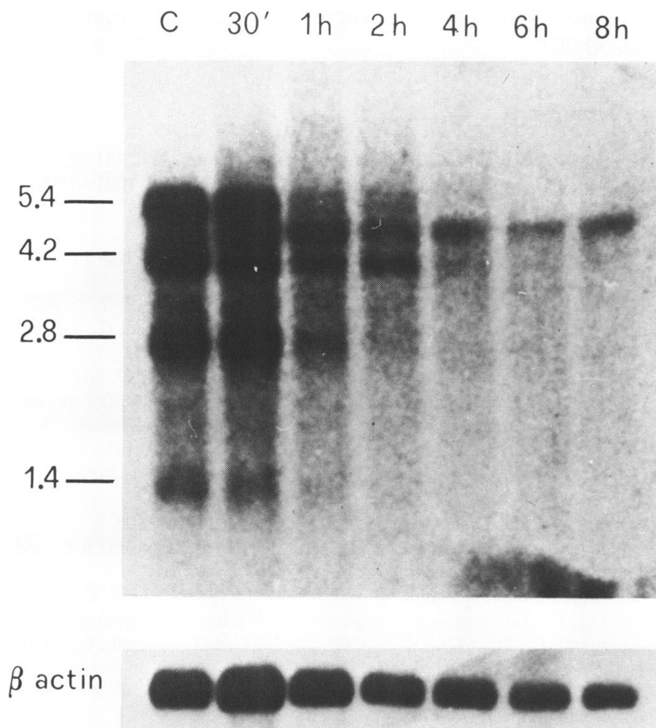


Figure 5: Northern blot analysis of poly(A)$^+$ RNA (2–3 μg/lane) from RA-induced NT2/D1 cells treated for different times (0–8 h) with 1 μg/ml actinomycin D hybridized to the B probe (see Fig. 2). C corresponds to time 0. Sizes in kb are indicated at left. Hybridization to the β-actin probe is shown. The 5.0 Kb band present in all lanes represents a cross hybridization of the GC-rich (70%) B probe to a 28S ribosomal RNA residue.

to be mainly accumulated in the nucleus, whereas the 4.2 Kb RNA is distributed between the nuclear and cytoplasmic compartments, representing the main band in the cytoplasm.

DISCUSSION

We previously reported the sequence of a 1.3 Kb homeobox-containing cDNA, *HHo.c13*, apparently comprising the entire coding sequence (14). This clone was used to probe poly(A)$^+$ RNA samples from a variety of human embryonic body parts: several transcripts with differential stage- and/or tissue-specific expression were detected in limbs, heart, spinal cord, brain and backbone (14). We now report its expression in gut (see Fig. 1), as described for the murine *Hox-5.1* (16, 17). The expression in the brain appears to be restricted to a region between the cervical and pontine flexures (*medulla oblongata*), as the region anterior to the pontine flexure fails to express any of the *HOX-5.1* transcripts (data not shown). Accordingly, the anterior boundary of the expression of the murine *HOX-5.1* has been located in the mid-myelencephalon (17).

The present work also includes an extensive analysis of the structure of the human *HOX-5.1* gene, as well as the mechanisms responsible for the expression of its mRNAs.

Production of multiple transcripts is a feature shared by a number of genes, including the homeobox (9, 32, 33, 34). The multiple transcripts, often showing a stage- and/or tissue-specific expression pattern in development, arise from multiple promoters and/or alternative processing. *HOX-5.1* appears to use two promoters in both embryonic tissues and teratocarcinoma cells: the 5.4 and 2.8 Kb transcripts are driven from the distal promoter and the 4.2 from the proximal one. The two promoters appear

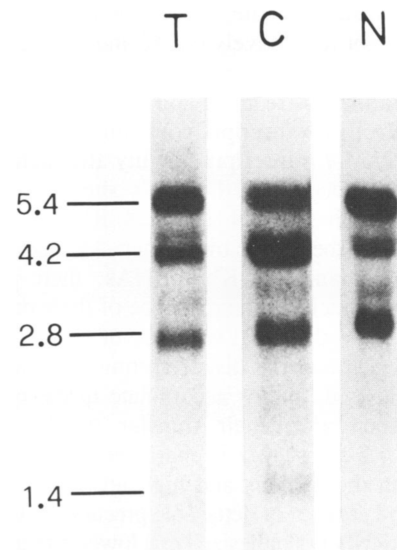


Figure 6: Northern blot analysis of poly(A)$^+$ RNA (2–3 μg/lane) from total cellular (T), cytoplasmic (C) and nuclear (N) fractions of RA-induced NT2/D1 cells, hybridized to the B probe (see Fig. 2). Sizes in Kb are indicated at left.

to be differentially regulated in a tissue- and stage-specific manner (14) and respond differently to RA induction (15). Likewise, a proximal and at least two distal, clustered but independent, polyadenylation sites exist: the first pertaining to the 2.8 Kb, the others to the 4.2 and 5.4 Kb transcripts (Fig. 2). We report the presence of a 1.4 Kb transcript, which is expressed as a faint band detected by only the B, C and D *c13*-specific probes (Fig. 1B). This transcript has been observed only in RA-induced NT2/D1 cells and in embryonic *medulla oblongata* poly(A)$^+$ RNA after prolonged exposure (see Figs. 1, 5 and 6). Remarkably, its expression is strictly related to that of the 4.2 Kb transcript, being clearly detectable only when the 4.2 kb band is abundantly expressed. The 1.4 Kb transcript appears to utilize the proximal cap site(s) and the proximal polyadenylation site.

Analysis of the sequence upstream from the transcription start points revealed that the distal promoter displays two TATA sequences, 29 and 67 bp, and an inverted CCAAT 129 bp upstream from the cap site (positions 347, 309 and 246 respectively). A 32 bp sequence composed of a tandemly repeated CTTT module located 322 bp upstream from the distal cap site (position 54), is invertedly repeated 209 bp downstream from it (position 584) in a 53-bp GA-rich sequence: the cap site is located between these inverted repeats. In this regard, poly(pyrimidine)-poly(purine) stretches have been implicated in gene regulation (35).

The proximal promoter does not show a TATA box, but displays an inverted CCAAT and two GGGCGG located at positions 1497, 1116 and 1253 respectively.

Regulatory elements acting as enhancers with time-, tissue- and region-specificity have been demonstrated in transgenic mice in the 2830 bp fragment upstream the Hincll site (position 1536 in Fig. 3) using a *lacZ* gene fusion construct (36). These elements specifically direct β-galactosidase expression to the upper cervical region of the central nervous system and appear to regulate *HOX-5.1* expression at transcriptional level from the proximal promoter (36).

The 5'-UT-region, included in the distally promoted transcripts, contains three potential short open reading frames at 3, 678 and

709 bp 3′ to the distal cap site, starting with an ATG, which are able to encode for respectively 68, 52 and 85 aminoacid long polypeptides. It has been suggested (37,38) that short open reading frames in the upstream region may be involved in the translational control of transcripts containing that region.

Analysis of *HOX-5.1* transcripts stability after actinomycin D treatment of RA-induced NT2/D1 cells shows a half-life of 20−30 min for both the 5.4 and 2.8 Kb transcripts and approximately 1 h for the 4.2 Kb messenger. Altogether, the short half-life of the 5.4 and 2.8 Kb mRNAs, their preferential accumulation in the nucleus, the presence of their precursors in the nuclear RNA and the rapid turnover of mature transcripts in cytoplasm suggest that the distally-promoted transcripts are abundantly synthesized, and/or accumulate in the nucleus, but rapidly decay soon after their transfer to the cytoplasm. Conversely, the 4.2 messenger is characterized by a relatively low abundance in the nucleus and high level in cytoplasm, a longer half-life and absence of detectable precursors: this suggests that the 4.2 Kb mRNA is synthesized at a lower rate as compared to the distal transcripts, but has a slower catabolic rate in the cytoplasm.

Control of these variables may be related in part to alternative functional features of the two differentially regulated upstream regions, and/or to the intrinsic stability of different mRNA molecules.

Determinants of mRNA stability may reside in the 5′ or 3′-UT region of the molecule (39). The determinant responsible for the differential stability of the 2.8 and 5.4 Kb transcripts versus the 4.2 Kb mRNA may be located in the long 5′-UT region, which structurally discriminates between the 5.4 and the 4.2 Kb mRNAs. In this regard, it has been reported that the 5′-UT region of c-myc oncogene is responsible for the rapid degradation of c-myc mRNA (40).

The 5.4 and 4.2 Kb transcripts largely extend 3′ to the proximal poly-A addition site in a long AT-rich sequence. The AT-stretches are reported to be responsible for rapid degradation of short-lived mRNAs (41): in this case they might partially contribute to the rapid turnover of the distally polyadenylated *HOX-5.1* transcripts.

An alternative explanation may be considered. RA-induced NT2/D1 cells differentiate into heterogeneous cell lineages that include neuron-like cells (23). It has been suggested (15) that the *HOX-5.1* transcripts may be differentially expressed in different cell subpopulations, *e.g.* the 4.2 Kb transcript, specifically expressed in embryonic *medulla oblongata* and spinal cord could be selectively synthesized in the neuron-like cells in RA-induced NT2/D1 cells. In this case the differential regulation and mRNA stability of *HOX-5.1* mRNAs may be attributed to the different intracellular environment in which each transcript is expressed.

In conclusion, analysis of the molecular mechanisms of *HOX-5.1* expression shows the presence of two independent classes of transcripts, one including the 5.4 and 2.8 Kb transcripts and the other the 4.2 and 1.4 Kb ones. The two classes appear to be differentially regulated at transcriptional and post-transcriptional level. However, all *HOX-5.1* transcripts share the same major open reading frame, thus suggesting they encode for the same protein.

The functional significance of the two classes of transcripts remains to be elucidated.

ACKNOWLEDGEMENTS

This work was supported by Progetto Finalizzato CNR 'Oncologia' to C.P. (Grant No. 88.00809.44) and Progetto Finalizzato CNR 'Biotecnologie' to C. P. (Grant No. 89.00250.70).

REFERENCES

1. Holland, P.W.H. and Hogan, B.L.M. (1986) Nature, *321*, 251−253.
2. Desplan, C., Theis, J. and O'Farrell, P. (1988) Cell, *54*, 1081−1090.
3. Hoey, T. and Levine, M. (1988) Nature, *332*, 858−861.
4. Gehring, W.J. (1987) Science, *236*, 1245−1252.
5. Boncinelli, E., Somma, R., Acampora, D., Pannese, M., D'Esposito, M., Faiella, A. and Simeone, A. (1988) Hum. Reprod., *3*, 880−886.
6. Holland, P.W.H. and Hogan, B.L.M. (1988) Gene Dev., *2*, 773−782.
7. Simeone, A., Mavilio, F., Bottero, L., Giampaolo, A., Russo, G., Faiella, A., Boncinelli, E. and Peschle, C. (1986) Nature, *320*, 773−775.
8. Graham, A., Papalopulu, N. and Krumlauf, R. (1989) Cell, *57*, 367−378.
9. Giampaolo, A., Acampora, D., Zappavigna, V., Pannese, M., D'Esposito, M., Caré, A., Faiella, A., Stornaiuolo, A., Russo, G., Simeone, A. , Boncinelli, E. and Peschle, C. (1989) Differentiation, *40*, 191−197.
10. Gaunt, S.J., Sharpe, P.T. and Duboule, D. (1988) Development, *104* (Suppl.), 169−179.
11. Duboule, D. and Dollè, P. (1989) EMBO J., *8*, 1497−1505.
12. Wolgemuth, D.J., Behringer, R.R., Mostoller, M.P., Brinster, R.L. and Palmiter, R.D. (1989) Nature, *337*, 464−467.
13. Balling, R., Mutter, G., Gruss, P. and Kessel, M. (1989) Cell, *58*, 337−347.
14. Mavilio, F., Simeone, A., Giampaolo, A., Faiella, A., Zappavigna, V., Acampora, D., Poiana, G., Russo, G., Peschle C. and Boncinelli, E. (1986) Nature, *324*, 664−668.
15. Mavilio, F., Simeone, A., Boncinelli, E. and Andrews, P.W. (1988) Differentiation, *37*, 73−79.
16. Featherstone, M.S., Baron, A., Gaunt, S.J., Mattei, M.G. and Duboule, D. (1988) Proc. Natl. Acad. Sci. USA, *85*, 4760−4764.
17. Gaunt, S.J., Krumlauf, R. and Duboule, D. (1989) Development, *107*, 131−141.
18. Andrews, P.W., Goodfellow, P.N., Shevinsky, L., Bronson, D.L., and Knowles B.B. (1982) Int. J. Cancer, *29*, 523−531.
19. Andrews, P.W., Damjanov, I., Simon, D., Banting, D.S., Carlin, C., Dracopoli, N.C. and Fogh, J. (1984) Lab. Invest., *50*, 147−162.
20. Simeone, A., Pannese, M., Acampora, D., D'Esposito, M. and Boncinelli, E. (1988) Nucl. Acids Res., *16*, 5379−5390.
21. Sanger, F., Nicklen, S. and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. USA, *74*, 5463−5467.
22. Chen, E.Y. and Seeburg, P.H. (1985) DNA, *4*, 165−170.
23. Andrews, P.W. (1984) Dev. Biol., *103*, 285−293.
24. Andrews, P.W., Gonczol, E., Plotkin, S.A., Dignazio, M. and Oosterhuis, J.W. (1986) Differentiation, *31*, 119−126.
25. Chirgwin, J.M., Przybyle, A.E., MacDonald, R.J. and Rutter, W.J. (1979) Biochemistry, *18*, 5294−5300.
26. Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor.
27. Birnstiel, M.L., Busslinger, M. and Strub, K. (1985) Cell, *41*, 349−359.
28. Lodish, H., Darnell, J. and Baltimore, D. (1986) Molecular Cell Biology. Scientific American Books, Inc. New York., pp. 343−349.
29. Calzone, F.J., Britten, R.J. and Davidson, E.H. (1987) Methods Enzymol., *152*, 611−632.
30. Mount, S.M. (1982) Nucl. Acids Res., *10*, 459−472.
31. Galliot, B., DollÄ, P., Vigneron, M., Featherstone, M.S., Baron, A. and Duboule, D. (1989) Development, *107*, 343−359.
32. Gehring, W.J. and Hiromi, Y. (1986) Ann. Rev. Genet., *20*, 147−173.
33. Simeone, A., Mavilio, F., Acampora, D., Giampaolo, A., Faiella, A., Zappavigna, V., D'Esposito, M., Pannese, M., Russo, G., Boncinelli, E. and Peschle C. (1987) Proc. Natl. Acad. Sci. USA, *84*, 4914−4918.
34. Cho, K.W., Goetz, J., Wright, C.V., Fritz, A., Hardwicke, J. and De Robertis, E.M. (1988) EMBO J., *7*, 2139−2149.
35. Lee, J.S., Woodsworth, M.L., Latimer, L.J.P. and Morgan, A.R. (1984) Nucl. Acids Res., *12*, 6603−6614.
36. Tuggle, C.K., Zakany, J., Cianetti, L., Peschle, C. and Nguyen-Huu, M.C. (1990) Gene Dev., 4, 180−189.
37. Burglin, T.R., Wright, C.V.E. and De Robertis, E.M. (1987) Nature, *330*, 701−702.
38. Kessel, M. and Gruss, P. (1988) Nature, *332*, 117−118.
39. Brawerman, G. (1987) Cell, *48*, 5−6.
40. Rabbits, P.H., Forster, A., Stinson, M.A. and Rabbitts, T.H. (1985) EMBO J., *4*, 3727−3733.
41. Shaw, G. and Kamen, R. (1986) Cell, *46*, 659−667.