inhibitors (*12*) produces behavioral effects similar to those of potent adenosine receptor agonists (*2*). These findings indicate that AD has a major role in regulating the presynaptic levels of adenosine during the process of nucleoside release and uptake. Alternatively, AD may participate in terminating the pre- or postsynaptic actions of adenosine, or both, by degrading this nucleoside to its relatively inactive metabolite, inosine (*2*). Such a putative role is consistent with the location and the high specific activity of AD in cortical synaptosomes (*4*). Finally, the application of adenosine has been shown to modify neuronal electrical activity in many brain areas including cortex and striatum (*2, 11, 13*), a finding compatible with the diverse axonal projections of AD-containing neurons in the TM, CM, and PCM nuclei.

A separate issue concerns the report by Vincent *et al.* (*9*) that many neurons in TM, CM, and PCM contain glutamic acid decarboxylase and may therefore utilize γ-aminobutyric acid as a neurotransmitter. Whether AD-containing neurons in these nuclei form a population separate from those containing glutamic acid decarboxylase or whether the two enzymes occur in the same cells remains to be determined. Some neurons in the TM and CM that were retrogradely labeled after fast blue injections into the cortex or striatum did not show staining for AD (Fig. 1, e and f). This indicates the possibility for some degree of segregation of hypothalamic neuronal populations containing AD and glutamic acid decarboxylase. Nevertheless we cannot rule out the presence of both of these enzymes in at least some hypothalamic neurons.

J. I. NAGY*
L. A. LABELLA
M. BUSS
Department of Physiology, Faculty of Medicine, University of Manitoba, Winnipeg, Manitoba R3E OW3, Canada
P. E. DADDONA
Department of Internal Medicine, University of Michigan Medical School, Ann Arbor, 48109

References and Notes

1. G. Burnstock, *Pharmacol. Rev.* **24**, 509 (1972); B. B. Fredholm and P. Hedqvist, *Biochem. Pharmacol.* **29**, 1635 (1980); S. H. Snyder, *Trends Neurosci.* **4**, 242 (1981).
2. J. W. Phillis and P. H. Wu, *Prog. Neurobiol. (N.Y.)* **16**, 187 (1981).
3. J. W. Daly, R. F. Burns, S. H. Snyder, *Life Sci.* **28**, 2083 (1981); J. A. Ribeiro, *ibid.* **22**, 1373 (1978); S. H. Snyder, J. J. Katims, Z. Annau, R. F. Burns, J. W. Daly, *Proc. Natl. Acad. Sci. U.S.A.* **78**, 3260 (1981).
4. I. Pull and H. McIlwain, *Biochem. J.* **144**, 37 (1974).
5. P. E. Daddona and W. N. Kelley, *J. Biol. Chem.* **252**, 116 (1977).
6. P. E. Daddona, unpublished observations.
7. J. I. Nagy, M. Goedert, S. P. Hunt, A. Bond, *Neuroscience* **7**, 3137 (1982).
8. R. Bleier, P. Cohn, I. R. Siggelkow, in *Handbook of the Hypothalamus*, P. J. Morgane and J. Panksepp, Eds. (Dekker, New York, 1979), vol. 1, p. 137.
9. S. R. Vincent, T. Hökfelt, L. R. Skirboll, J.-Y. Wu, *Science* **220**, 1309 (1983).
10. J. I. Nagy and S. P. Hunt, *Neuroscience* **7**, 89 (1982).
11. J. W. Phillis, J. P. Edstrom, G. K. Kostopoulos, J. R. Kirkpatrick, *Can. J. Physiol. Pharmacol.* **57**, 1289 (1979).
12. W. B. Mendelson, A. Kuruvilla, T. Watlington, K. Goehl, S. M. Paul, P. Skolnick, *Psychopharmacology* **79**, 126 (1983).
13. G. K. Kostopoulos, J. J. Limacher, J. W. Phillis, *Brain Res.* **88**, 162 (1975); G. K. Kostopoulos and J. W. Phillis, *Exp. Neurol.* **55**, 719 (1977).
14. Supported by grants from the Manitoba Mental Health Foundation and the University of Manitoba Faculty Fund (to J.I.N.) and by grant CA26284 from the National Cancer Institute (to P.E.D.). Summer student support to M.B. was provided through the University of Manitoba bachelor of science (medicine) program; L.A.L. is a recipient of a Canadian Medical Research Council studentship; J.I.N. is a scholar of the Canadian Medical Research Council.
* Address requests for reprints to J.I.N.

27 October 1983; accepted 25 January 1984

Sequence of the Human Somatostatin I Gene

Abstract. *Two human genomic DNA fragments containing alleles for the gene coding for somatostatin I were isolated and sequenced. This gene contains a single intron that interrupts the coding sequence in the propeptide portion of the somatostatin moiety. The site of initiation of transcription of the gene was located by transcription experiments in HeLa cell extracts, and the putative regions for controlling the initiation of transcription were identified.*

Somatostatin is a 14–amino acid neuropeptide and hormone (somatostatin-14) found in the brain and spinal cord, and in the pancreas, stomach, and intestine. It suppresses the release of many pituitary, pancreatic, and gastrointestinal polypeptide and glycoprotein hormones and regulates some aspects of gastrointestinal function, including gastric acid and pepsin secretion as well as smooth muscle contractility (*1*). In addition, it may function as a neurotransmitter (*1*). Previous studies suggest that somatostatin-14 is derived from a larger polypeptide, prosomatostatin (*1*). The sequence of the somatostatin precursors, including their signal peptides, have been deduced from the nucleotide sequence of the cloned somatostatin complementary DNA's (cDNA's) from anglerfish (*2, 3*), human (*4*), catfish (*5–8*), and rat (*9*). The nucleotide sequences of two independently isolated cDNA's from the anglerfish endocrine pancreas (*2*) revealed the amino acid sequences of two different prosomatostatin molecules, each of which contained distinct 14– and 28–amino acid somatostatin moieties

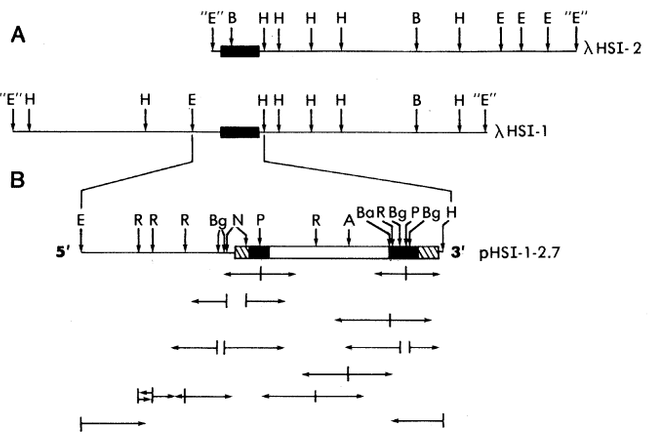


Fig. 1. Organization of the human somatostatin gene. (A) Linkage map of the gene and adjacent regions. One million recombinant phage of a λ Charon 4A library prepared from a partial Hae III–Alu I digest of human fetal liver DNA were screened with a nick-translated human preprosomatostatin cDNA probe (*22*). Phage containing somatostain gene sequences were plaque-purified, and then DNA was prepared from phage grown in liquid culture (*23*). The location of the somatostatin gene (indicated by the filled box) was determined by hybridization of the ^{32}P-labeled cDNA to Southern blots of single and double restriction endonuclease digests of DNA from λHSI-1 and λHSI-2. The restriction enzymes indicated are as follows: A, Acc I; B, Bam HI; Ba, Bal I; Bg, Bgl II; E, Eco RI; H, Hind III; N. Nar I; P, Pst I; and R, Rsa I. The ordering of the three Eco RI sites at the 3′ end of λHSI-2 has not been determined. The quotation marks indicate Eco RI sites produced by linker addition during the construction of the library. The Eco RI site of λHSI-1 on the right appears close to or identical to the natural site. (B) Organization of the gene. The 2.7-kbp Eco RI–Hind III fragment of λHSI-1 was cloned into plasmid pBR322 (pHSI-1-2.7). The box delimits the structural portion of the gene. The hatched section indicates the region encoding the untranslated portion of the mRNA, the filled region indicates the protein coding portion, and the open region represents the intervening sequence. The sequencing strategy and the restriction sites at which sequence determinations were initiated are shown.

(somatostatin I and somatostatin II). Subsequent studies have shown that somatostatin I and II have distinct biological activities (*10*). An additional somatostatin gene from catfish not clearly related to anglerfish somatostatin I or II has also been inferred from the amino acid sequence of the isolated peptide (*11*). These observations suggest that there may be a family of somatostatin genes, with individual moieties having distinct biological activities. However, only a single somatostatin peptide has been isolated from mammals. The amino acid sequence of somatostatin-14 and somatostatin-28 predicted from the cloned somatostatin cDNA derived from a human somatostatinoma of the pancreas and from a rat medullary carcinoma are identical and in precise agreement with the 14– and 28–amino acid peptides isolated from porcine and ovine hypothalamus and porcine intestine (*12–14*).

Hybridization with a radioactively labeled cloned human preprosomatostatin cDNA was used to isolate the somatostatin gene from a λ Charon 4A library containing human chromosomal DNA (*4*). Four distinct recombinants were isolated; three of these contained two Eco RI fragments of 6.0 and 12.0 kilobase pairs (kbp) (λHSI-1) (Fig. 1A), and the other contained Eco RI fragments of 1.1 and 10.3 kbp (λHSI-2) (Fig. 1A). In both λHSI-1 and λHSI-2, the larger Eco RI fragment hybridized with the cDNA probe. This probe also hybridized to an Eco RI fragment of 12 to 13 kbp in digests of genomic DNA. Mapping studies with restriction endonucleases indicated that λHSI-1 and λHSI-2 contain congruent DNA segments from the region of the somatostatin gene (Fig. 1A). However, λHSI-2 contains an additional Bam HI site within the somatostatin gene (Fig. 1A). Studies of the organization of the somatostatin gene in human populations have established that this Bam HI site and an Eco RI site not present in either of these clones are polymorphic (*15*). We infer that λHSI-1 and λHSI-2 are somatostatin alleles. Dispersed middle repetitive sequences are present in both of these clones, but their locations have not been precisely determined.

We have sequenced the 2667-base-pair (bp) Eco RI–Hind III fragment of λHSI-1 containing the somatostatin gene (Figs. 1 and 2), as well as the other allelic somatostatin gene sequence, λHSI-2 (not shown). In addition, a 1953-bp fragment was sequenced from the Hae III site at nucleotide 411 to the nucleotide at position −408, to which an Eco RI linker had been added during the cloning, to the Hind III site at nucleotides 1537 to 1542 (Fig. 2). Only three changes were found in this 1953-bp region of λHSI-2: an insertion in the 5′ flanking region of two additional T's (T, thymine) in the oligoT tract between nucleotides −164 and −152; a substitution of a C (C, cytosine) for T at position 274 which generates the Bam HI site, and a deletion of three TG (G, guanine) pairs in the tandem repeat region between nucleotides 846 and 868. Comparison with the cDNA sequence (*4*) indicates that the structural region of the gene is 1480 bp and that it contains a single intervening sequence of 877 bp which interrupts the protein coding regions of the gene between glutamine at amino acid position −57 and glutamic acid at position −56 in the propeptide portion of the somatostatin precursor. The sequence of the anglerfish somatostatin I gene (*16*) indicates that this gene contains an intron in the same location. Thus, as in other genes, the location of

```
GAATTCAAGGACAGGTTTTCTTAAACTTTCTTTGTTTCTAGGAGATCAGGCAGAGCTGAATTTAACCAAGA -1055
ATCTTTTGATCCTTTCCACATATAGATATACAATAGTGGTCACATATGTTCTGGGAGTTCCTAGACCTTAT -984
ATGTCTAAACTGGGGCTTCCTGACATAAAACTATGCTTACCGGCAGGAATCTGTTAGAAAACTCAGAGCTC -913
AGTAGAAGGAACACTGGCTTTGGAATGTGGAGGTCTGGTTTTGCTCAAAGTGTGCAGTATGTGAAGGAGAA -842
CAATTTACTGACCATTACTCTGCCTTACTGATTCAAATTCTGAGGTTTATTGAATAATTTCTTAGATTGCC -771
TTCCAGCTCTAAATTTCTCAGCACCAAAATGAAGTCCATTTCAATCTCTCTCTCTCTTTCCCTCCCGTA -700
CATATACACACACTCATACATATATATGGTCACAATAGAAAGGCAGGTAGATCAGAAGTCTCAGTTGCTGA -629
GAAAGAGGGAGGGAGGGTGAGCCAGAGTACTTCTCCCCCATTGTAGAGAAAAGTGAAGTTCTTTTAGAGCC -558
CCGTTACATCTTCAAGGCCTTTTATGAGATAATGGAGGAAATAAAGAGGGCTCAGTCCTTCTACCGTCCAT -487
ATTTCATTCTCAAATCTGTTATTAGAGGAATGATTCTGATCTCCACCTACCATACACATGCCCTGTTGCTT -416
GTTGGGCCTTACACTAAAATGTTAGAGTATGATGACAGATGGAGTTGTCTGGGTACATTTGTGTGCATTTA -345
AGGGTGATAGTGTATTTGCTCTTTAAGAGCTGAGTGTTTGAGCCTCTGTTTGTGTGTAATTGAGTGTGCAT -274
GTGTGGGAGTGAAATTGTGGAATGTGTATGCTCATAGCACTGAGTGAAAATAAAAGATTGTATAAATCGTG -203
GGGCATGTGGAATTGTGTGTGCCTGTGCGTGTGCAGTATTTTTTTTTTTTTAAGTAAGCCACTTTAGATCT -132
TGTCACCTCCCCTGTCTTCTGTGATTGATTTTGCGAGGCTAATGGTGCGTAAAAGGGCTGGTGAGATCTGG -61
                                                                  *
GGGCGCCTCCTAGCCTGACGTCAGAGAGAGAGTTTAAAACAGAGGGAGACGGTTGAGAGCACACAAGCCGC  11
TTTAGGAGCGAGGTTCGGAGCCATCGCTGCTGCCTGCTGATCCGCGCCTAGAGTTTGACCAGCCACTCTCC  82
                        -102
                         Met Leu Ser Cys Arg Leu Gln Cys Ala Leu Ala Ala
AGCTCGGCTTTCGCGGCGCCGAG ATG CTG TCC TGC CGC CTC CAG TGC GCG CTG GCT GCG 141
Leu Ser Ile Val Leu Ala Leu Gly Cys Val Thr Gly Ala Pro Ser Asp Pro Arg
CTG TCC ATC GTC CTG GCC CTG GGC TGT GTC ACC GGC GCT CCC TCG GAC CCC AGA 195
                                                        -57
Leu Arg Gln Phe Leu Gln Lys Ser Leu Ala Ala Ala Ala Gly Lys Gln
CTC CGT CAG TTT CTG CAG AAG TCC CTG GCT GCT GCC GCG GGG AAG CAG  GTAAGG 249
AGACTCCCTCGACGTCTCCCGGATTCTCCAGCCCTCCCTAAGCCTTGCTCCTGCCCCATTGGTTTGGACGT 320
AAGGGATGCTCAGTCCTTCTAAAGAGTTTTGGTGCTTTTCTGGGTCCCTCAGCTCCCGAAGCTCTTGAGAA 391
AACTATCAAAGGCTAGAATCCCCTTCTAACTCTTTTTTTCCCCCATGATAAGCGCAGTCGGTCACAGTTCA 462
GGTGAGTTCTTACTTGGCATTCAAGAAAATTACAAAATCTGGGTAGTTGTCTGGGCACGAAGCGACAATGG 533
CGTCTATCCCTGGTGCTGACCCTGGGAAGCGCTGACCCAGGTGCTGAAACGCAGACCTCTGAAGCTGCTAC 604
CTCTTAGCGTACCTCACTTCCAAACGTCGGGACTAGGGCAAAGGGGCAATCTAAAGACCGAACGCCGTATG 675
TTTGAGATTGTGAGAAGCCTCGTTCCCCTACAGTTTTACTTGGTAAAAATGGTAAAACAATTCTACTTTGT 746
AGCTCGTGATGTGAAAATTGAATTAAACTGTTGGCACACACTTTATCTTACCAGAACGGTCTTTATGTGTG 817
TGTGTGTGTGTGTGTGTGTTTGTGCGTGTGTGTGTGTGTGTGTGTGTTAAGTCTACAGGGACAGAAA 888
GGTTGCAGAAACATTTGAGCTCTTAAAGCCTTTTTGTGTAACTTGGTAATTATAGCAACTATCCTTATTTT 959
TATATCCTTGATTGATTTTAAATGTGACAAAAAATGCGCAGCTGTAAAAACTGGATTTTGTGTGTGACCAA 1030
ATCTGTTCTTTAATTTAGGCTTTTCAAATTTTTTCCATTGTCCTCCCCACTTCTCTTTCTCTCTTTTTCTA 1101
                    -56
                    Glu Leu Ala Lys Tyr Phe Leu Ala Glu Leu Leu Ser Glu
TCCCTTCTGCCCTATACAG GAA CTG GCC AAG TAC TTC TTG GCA GAG CTG CTG TCT GAA 1159
Pro Asn Gln Thr Glu Asn Asp Ala Leu Glu Pro Glu Asp Leu Ser Gln Ala Ala
CCC AAC CAG ACG GAG AAT GAT GCC CTG GAA CCT GAA GAT CTG TCC CAG GCT GCT 1213
Glu Gln Asp Glu Met Arg Leu Glu Leu Gln Arg Ser Ala Asn Ser Asn Pro Ala
GAG CAG GAT GAA ATG AGG CTT GAG CTG CAG AGA TCT GCT AAC TCA AAC CCG GCT 1267
                          1
Met Ala Pro Arg Glu Arg Lys Ala Gly Cys Lys Asn Phe Phe Trp Lys Thr Phe
ATG GCA CCC CGA GAA CGC AAA GCT GGC TGC AAG AAT TTC TTC TGG AAG ACT TTC 1321
    14
Thr Ser Cys AM
ACA TCC TGT TAG  CTTTCTTAACTAGTATTGTCCATATCAGACCTCTGATCCCTCGCCCCCACACCC 1387
CATCTCTCTTCCCTAATCCTCCAAGTCTTCAGCGAGACCCTTGCATTAGAAACTGAAAACTGTAAATACAA 1458
AATAAAATTATGGTGAAATTATGAAAAATGTGAATTTGGTTTCTATTGAGTAAATCTTTTTGTTCAATAAT 1529
                    ▲
ACATAATAAGCTT
```

Fig. 2. Sequence of the human somatostatin gene in pHSI-1-2.7. The DNA sequence was determined by the method of Maxam and Gilbert (*24*) as described (*4*). The cap site is indicated by an asterisk. The TATA box is underlined and an arrow indicates the 3′ end of the structural portion of the gene. The cap site is nucleotide 1. Upstream nucleotides are designated with minus signs. The predicted amino acid sequence of preprosomatostatin is also indicated. Alanine, the first amino acid of somatostatin-14, is residue 1.

intron junctions in the gene appears to be relatively conserved during evolution, even though boundary sequences may be quite divergent.

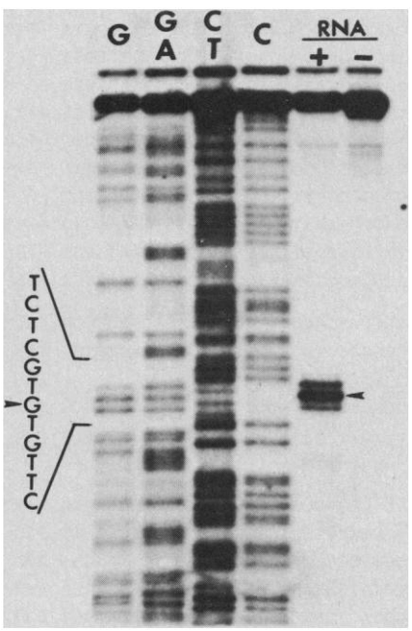
The sequences of the coding region of both alleles are identical to that of the isolated cDNA. The cap site nucleotide, A1 (A, adenine) (asterisk in Fig. 2), present in the sequence of one of the cDNA clones, was identified by S1 nuclease mapping (*17*) (Fig. 3). Four fragments of differing lengths were protected by the RNA; they comigrated with DNA sequence-specific fragments in the region of the sequencing ladder 5′-TTGTGTGCT-3′ in the noncoding strand, corresponding to the RNA transcript sequence 5′-AGCACACAA-3′. The results of the S1 mapping experiment suggest that transcription begins at nucleotides A1, G2, A3, or A5 (Fig. 2); thus somatostatin messenger RNA (mRNA) probably has heterogeneous 5′ ends. We have obtained cDNA clones corresponding to an mRNA starting at A1 and at A3. A comparison of the gene sequences with the cDNA sequence and the S1 mapping results indicates that there are no intervening sequences in the region of the gene encoding the 5′ untranslated portion of the mRNA.

The results of transcription experiments in vitro (Fig. 4) support the assignment of transcription initiation sites and suggest that the sequence TTTAAA, a variant of the TATA box (*18*) found at positions −28 to −23 from the cap site (Fig. 2), is part of the somatostatin gene promoter. DNA from pHSI-1-2.7 (Fig. 1B) was digested with Sac I (coordinates 905 to 910) and Sac II (coordinates 230 to 235) and mixed with a HeLa cell extract. Runoff transcripts of around 950 and 230 bases, indicated by the arrowheads in Fig. 4, were observed with the Sac I– and Sac II–digested templates, respectively. The synthesis of both transcripts was sensitive to low levels of α-amanitin (200 ng/ml), indicating that they were produced by RNA polymerase II. The length of the shorter (more precisely measurable) runoff transcript, in particular, is in good agreement with the length expected if transcription is initiated at the cap site suggested by the cDNA sequence and the S1 mapping experiments.

The 877-bp intervening sequence contains many short homopolymeric sequences, including a 54-bp repeat of the dinucleotide TG (nucleotides 812 to 865). Similar regions of repeating dinucleotides have been observed in the flanking regions and intervening sequences of other genes (*19, 20*).

The location of the intron in the central region of the transcribed portion of the somatostatin gene is somewhat analogous to the central position of the intron in the insulin gene. However, in contrast to the situation in the insulin gene, where the intron is in the connecting peptide region, the intron in the somatostatin gene does not obviously bifurcate functional domains. The somatostatin intron position exists within—but on the edge of—a region conserved between the anglerfish and the human genes (−59 to −31), and next to a region of high variability, probably caused by additions and deletions (−65 to −60). Thus in line with other genes, the introns may exist proximal to a region of length polymorphism (*21*).

The availability of the human gene will allow studies of its expression in heterologous cells. The putative prosomatostatin can be produced and its biological and biochemical properties studied. The human gene sequences should also aid in studies of the regulation of the somatostatin gene in various cells. This is of particular physiological interest in view of the many biological roles of this hormone.



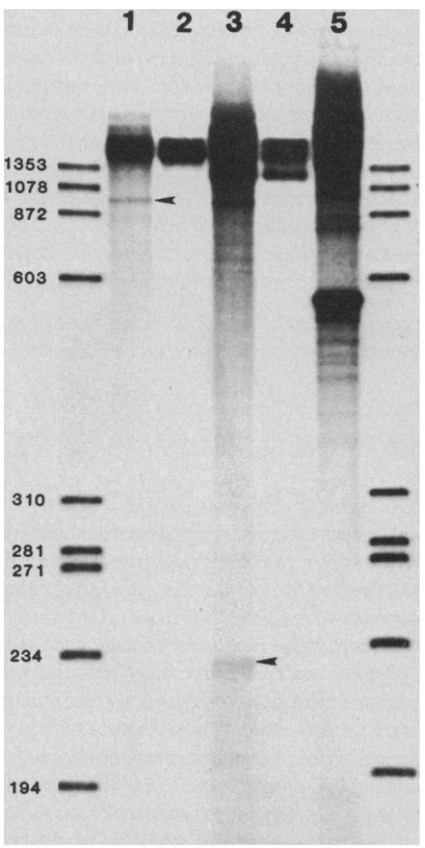


Fig. 3 (left). S1 mapping of the 5′ end of somatostatin mRNA. A 276-bp Bgl II–Pst I fragment [coordinates −68 to −63 and 208 to 213, respectively (Fig. 2)], labeled at the 5′ end of the Pst I site was hybridized overnight with or without 1 μg of total human somatostatinoma RNA in 20 μl of a mixture of 80 percent formamide, 0.4*M* NaCl, and 10 m*M* Pipes, *p*H 6.5, at 52°C (the melting temperature of this DNA fragment). The hybridization was terminated by diluting the reaction mixture with 200 μl of ice-cold 0.3*M* NaCl, 3 m*M* $ZnCl_2$, and 30 m*M* sodium acetate, *p*H 4.5, containing 1000 units of S1 nuclease (Miles). After incubation at 37°C for 1 hour, the digest was extracted with phenol and precipitated with ethanol. The sizes of the S1 nuclease–resistant DNA fragments were determined on an 8 percent polyacrylamide-urea gel (*24*). The sequencing ladder from the coding strand of the 276-bp Bgl II–Pst I fragment that was used to estimate the sizes of the S1-resistant fragments is indicated. The arrowhead in the +RNA lane indicates the S1-resistant fragments protected by mRNA molecules initiating at A3 (Fig. 2). The fragment protected by mRNA molecules initiating at G2 is faint. Fig. 4 (right). Transcription of the somatostatin gene in vitro. An extract was prepared from HeLa cells as described by Manley *et al.* (*25*). The reaction mixture (50 μl) contained 30 μl of extract, 60 μ*M* each of adenosine, uridine, and cytidine triphosphates, 25 μ*M* guanidine triphosphate (GTP), 10 μCi of [α − ^{32}P]GTP, and 2 μg of DNA. (Lanes 1 and 2) Sac I–pHSI-1-2.7; (lanes 3 and 4) Sac II–pHSI-1-2.7; (lane 5) Sma I–pSmaF. [A derivative of pBR313 containing the SmaF fragment of adenovirus that includes the adenovirus major late promoter (*26, 27*) and the runoff transcript from a Sma I digest is 536 bases (*27*)]. Lanes 2 and 4 also contained α-amanitin (0.2 μg/ml). After 1.5 hours of incubation at 30°C the reaction was stopped by adding 350 μl of 6*M* guanidinium hydrochloride, and the samples were prepared as described (*28*). The runoff transcripts were analyzed by electrophoresis in a 5 percent polyacrylamide-urea gel (*24*). ^{32}P-Labeled fragments of Hae III–digested φX174 were included as size markers (outside lanes). The identity of the higher molecular weight α-amanitin–insensitive transcripts is unknown.

LU-PING SHEN*
WILLIAM J. RUTTER
Department of Biochemistry and Biophysics, University of California, San Francisco 94143

References and Notes

1. A. Arimura, *Biomed. Res.* **2**, 233 (1981).
2. P. Hobart, R. Crawford, L. P. Shen, R. Pictet, W. J. Rutter, *Nature (London)* **288**, 137 (1980).
3. R. H. Goodman, J. W. Jacobs, W. W. Chin, P. K. Lund, P. C. Dee, J. F. Habener, *Proc. Natl. Acad. Sci. U.S.A.* **77**, 5869 (1980).
4. L.-P. Shen, R. L. Pictet, W. J. Rutter, *ibid.* **79**, 4575 (1982).
5. W. L. Taylor, K. J. Collier, R. J. Deschenes, H. L. Weith, J. E. Dixon, *ibid.* **78**, 6694 (1981).
6. M. Magazin, C. D. Minth, C. L. Funckes, R. Deschenes, M. A. Tavianini, J. E. Dixon, *ibid.* **79**, 5152 (1982).
7. C. D. Minth *et al.*, *J. Biol. Chem.* **257**, 10372 (1982).
8. P. C. Andrews and J. E. Dixon, *ibid.* **256**, 8267 (1981).
9. R. H. Goodman, J. W. Jacobs, P. C. Dee, J. F. Habener, *ibid.* **257**, 1156 (1982).
10. J. E. Shields, M. A. Root, W. J. Rutter, unpublished data.
11. H. Oyama, R. A. Bradshaw, O. J. Bates, A. Permutt, *J. Biol. Chem.* **255**, 2251 (1980).
12. A. V. Schally *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **77**, 4489 (1980).
13. F. Esch, P. Bohlen, N. Ling, R. Benoit, P. Brazeau, R. Guilleman, *ibid.*, p. 6827.
14. L. Pradayrol, H. Jornavall, V. Mutt, A. Ribert, *FEBS Lett.* **109**, 5558 (1980).
15. S. L. Naylor, A. Y. Sakaguchi, L.-P. Shen, G. I. Bell, W. J. Rutter, T. B. Shows, *Proc. Natl. Acad. Sci. U.S.A.* **80**, 2686 (1983).
16. R. C. Crawford, P. Hobart, W. J. Rutter, in preparation.
17. R. F. Weaver and C. Weissmann, *Nucleic Acids Res.* **7**, 1175 (1979).
18. R. Breathnach and P. Chambon, *Annu. Rev. Biochem.* **50**, 349 (1981).
19. I. Sures, J. Lowry, L. H. Kedes, *Cell* **15**, 1033 (1978).
20. G. L. Slightom, A. E. Blechl, O. Smithies, *ibid.* **21**, 627 (1980).
21. C. S. Craik, W. J. Rutter, R. Fletterick, *Science* **220**, 1125 (1983).
22. G. I. Bell, J. H. Karam, W. J. Rutter, *Proc. Natl. Acad. Sci. U.S.A.* **78**, 5759 (1981).
23. F. R. Blattner *et al.*, *Science* **202**, 1279 (1978).
24. A. Maxam and W. Gilbert, *Methods Enzymol.* **65**, 499 (1980).
25. J. L. Manley, A. Fire, A. Cano, P. A. Sharp, M. L. Gefter, *Proc. Natl. Acad. Sci. U.S.A.* **77**, 2855 (1980).
26. P. A. Weil, D. S. Luce, J. Segall, R. G. Roeder, *Cell* **18**, 469 (1979).
27. T. Matsui, J. Segall, P. A. Weil, R. G. Roeder, *J. Biol. Chem.* **255**, 11992 (1980).
28. D. N. Standring, A. Venegas, W. J. Rutter, *Proc. Natl. Acad. Sci. U.S.A.* **78**, 5963 (1981).

29. We thank G. Bell and D. Standring for their contributions to this research and for their help in the preparation of the manuscript, V. Watt for aid in determining the restriction map of λHSI-1 and λHSI-2, and R. Crawford, W. Swain, and P. Hobart for valuable advice during this work. We also thank L. Spector for manuscript preparation. Supported by NIH grant AM 21344.

* Present address: Institute of Biochemistry, Shanghai, People's Republic of China.

23 February 1983; accepted 16 February 1984

High Incidence of "Leapfrog" Pattern of Geographic Variation in Andean Birds: Implications for the Speciation Process

Abstract. *Many species of birds in the humid forests of the Andes show a pattern of geographic variation in color that is virtually unknown in other regions of the world. This pattern, here termed "leapfrog," is one in which two populations very similar in appearance are geographically separated from each other by very different, intervening populations of the same species. Approximately 21 percent of all Andean bird species and superspecies with three or more differentiated populations show the leapfrog pattern, and several of these show multiple cases of leapfrogging color patterns. Lack of concordance in the geographic distribution of taxa showing the leapfrog pattern suggests that there is a strongly random component in phenotypic differentiation with respect to direction, geography, and timing.*

Patterns of geographic variation in birds have been documented and analyzed for a half-century or more, and the interpretation of these patterns has provided much of the basis for speciation theory (*1*). The clinal nature of most patterns of color variation has been interpreted as evidence for the importance of gene flow, environmentally induced selection, or both, in determining population structure (*1*). Syntheses of overall patterns of geographic variation produced "Gloger's rule"—the tendency for populations from more humid areas to be more heavily pigmented than conspecific populations from less humid areas—and what could be called "Mayr's rule"—the association between increasingly marked geographic isolation and increasingly marked phenotypic differentiation.

I now report a pattern of geographic variation in color in Andean birds, the generality of which has heretofore not been recognized. This counterclinal pattern, here labeled the "leapfrog" pattern, is one in which, within a single biotope, two phenotypically very similar populations are geographically separated from each other by very different intervening populations of the same species (see cover). Geographic variation of this type has been reported for a few bird species (*2*), and two cases from the Andes have been studied extensively (*3*); however, such cases have received little theoretical attention.

To quantify the frequency of occurrence of the leapfrog pattern in Andean birds, I analyzed geographic variation in color patterns of all bird species in humid forest and forest edge in the Andes from northern Colombia and Venezuela to northwestern Argentina, the southern limit of humid montane forest. This region was selected because of the relative homogeneity in habitats at any given elevation over a broad latitudinal range (*4*). The sample consisted of 386 species and an additional 30 superspecies assembled from a subset of the species sample.

Geographic variation in color pattern was analyzed within the framework of current subspecies limits. Although the subspecies concept has been attacked repeatedly on conceptual and practical grounds (*5*), subspecies were used as the unit of analysis simply because no alternative existed; a quantitative, comprehensive assessment of color variation in all 386 species would be a life-long task. The study skin collection of the Museum of Zoology, Louisiana State University, was the primary source of data for the analysis. These data were supplemented by compendiums of subspecies descriptions (*6*) and recent taxonomic revisions. A species or superspecies was considered to show the leapfrog pattern if two geographically nonadjacent taxa were more similar in plumage pattern and color to one another than either was to the intervening taxon.

A conservative bias in the analysis was that only major, conspicuous features of coloration and pattern were analyzed; potential leapfrog patterns in subtle, less obvious plumage characters were ignored. Another conservative bias was that many described subspecies from the Andes cannot be readily distinguished from adjacent populations with taxonomically acceptable (75 percent), much less statistically acceptable (95 percent) (*7*), certainty; inclusion of invalid subspecies artificially inflates the number of species in which a leapfrog pattern can be detected.

By definition, the leapfrog pattern can be detected only in species with three or more subspecies. Of the 386 species examined, 127 were monotypic, 45 had only one, and 85 had only two subspecies within the geographic limits of the study. Thus, 129 species (33.4 percent) remained for inclusion in the analysis. Of these, 25 (about 19 percent) (*8*) showed the leapfrog pattern. An additional nine species showed the leapfrog pattern when subspecies from outside the main Andes were included; for example, from the tepuis of southeastern Venezuela, coastal ranges of Venezuela, and the highlands of Middle America. As for superspecies, only six of the 30 examined contained the necessary minimum of three component allospecies. Of these, three superspecies (50 percent) displayed a leapfrog pattern of color variation (*9*). Thus, combining species and superspecies, of 135 taxa in which the leapfrog pattern is possible (that is, those with three or more component taxa), 28 (about 21 percent) displayed leapfrog color variation (Table 1). Furthermore, there are multiple cases of the leapfrog pattern within three species and