Retroviral and Pseudogene Insertion Sites Reveal the Lineage of Human Salivary and Pancreatic Amylase Genes from a Single Gene during Primate Evolution

LINDA C. SAMUELSON,* KARIN WIEBAUER,† CLAUDETTE M. SNOW, AND MIRIAM H. MEISLER

Department of Human Genetics, University of Michigan, Ann Arbor, Michigan 48109-0618

Received 3 November 1989/Accepted 23 February 1990

We have analyzed the junction regions of inserted elements within the human amylase gene complex. This complex contains five genes which are expressed at high levels either in the pancreas or in the parotid gland. The proximal 5'-flanking regions of these genes contain two inserted elements. A γ -actin pseudogene is located at a position 200 base pairs upstream of the first coding exon. All of the amylase genes contain this insert. The subsequent insertion of an endogenous retrovirus interrupted the γ -actin pseudogene within its 3'-untranslated region. Nucleotide sequence analysis of the inserted elements associated with each of the five human amylase genes has revealed a series of molecular events during the recent history of this gene family. The data indicate that the entire gene family was generated during primate evolution from one ancestral gene copy and that the retroviral insertion activated a cryptic promoter.

The complete human amylase gene cluster has recently been cloned in our laboratory and by others (8, 9, 11, 12, 21). The cluster spans 230 kilobases (kb) on human chromosome 1. The most common human haplotype contains five intact genes: two pancreatic genes (*AMY2A* and *AMY2B*) and three salivary genes (*AMY1A*, *AMY1B*, and *AMY1C*) (8, 9). The salivary and pancreatic genes are 93% identical in the sequence of the proximal 718 base pairs (bp) of the 5'-flanking regions (12). Nonetheless, these genes are expressed with different tissue specificities. *AMY2A* and *AMY2B* transcripts are present in the human pancreas but not in the parotid salivary gland, while *AMY1* transcripts are present in the parotid gland but not in the pancreas (28).

Two inserted elements have been identified in the 5' regions of the human amylase genes. A γ -actin pseudogene was identified 200 bp upstream of the first amylase-coding exon (4, 28). The promoter and nontranslated exon of the salivary amylase genes are derived from the γ -actin pseudogene. Retroviral long terminal repeat (LTR) and *gag* sequences were also identified in the 5'-flanking region of an amylase gene (4). We subsequently determined the locations of a total of seven LTRs within the gene cluster (28). The mouse amylase gene cluster does not contain retroviral or γ -actin-related sequences (unpublished observations), indicating that insertion of both elements occurred subsequent to the separation of human and mouse.

In the present report, we describe the junction regions of the γ -actin and retroviral inserts upstream of each human amylase gene. These inserts provide unique molecular markers for the recent evolution of this gene family. The structural data suggest a specific sequence of events by which the five human amylase genes were generated from one ancestral gene within the past 40 million years.

MATERIALS AND METHODS

Sequence analysis. Genomic DNA fragments from cosmids G6 (*AMY2B*), G45 (*AMY2A*), and N1 (*AMY1A*) (9) were

subcloned into pGEM vectors (Promega Biotec), and double-stranded plasmid DNA was sequenced with the Sequenase kit (U.S. Biochemicals, version 2.0). For most of the sequencing, the T7 and SP6 promoter primers were used. (See Fig. 2 and 3 [brackets] for sequences corresponding to other primers.) The primer for sequencing upstream of the *AMY1A* 3' LTR contained the sequence 5'ATATGATCTTAGTCCA3'. Identification of CCAAT-box and glucocorticoid consensus elements was carried out on an IBM PC computer with the aid of Pustell Sequence Analysis software (International Biotechnology Inc., version 2.02).

Southern blot analysis. Ten micrograms of genomic DNA was digested with *EcoRI*, electrophoresed on 0.7% agarose, transferred to nitrocellulose by the method of Smith and Summers (33), and hybridized with amylase or γ -actin riboprobes in a 50% formamide buffer (19) at 45°C. The final wash was at 60°C in 1× SSC (0.15 M NaCl and 0.015 M sodium citrate)–0.1% sodium dodecyl sulfate–20 mM PP_i. Restriction fragments from cosmid G21 (9) were subcloned into pGEM 1 or 2 for preparation of riboprobe templates. The amylase probe was a 470-nucleotide *HaeIII-EcoRI* fragment containing exon *a*. The γ -actin probe was a 690-nucleotide *HpaI-NdeI* fragment from the 3'-untranslated region of the γ -actin insert. Single-stranded uniformly labeled riboprobes were prepared as described by Samuelson et al. (28).

RESULTS

The γ -actin pseudogene and retroviral inserts at the 5' ends of the five human amylase genes were characterized by subcloning and sequencing. The results are summarized in Fig. 1, which presents the overall organization of genes within this complex (Fig. 1a) and the arrangement of inserted elements in the proximal upstream region of each gene (Fig. 1b).

Identification of an intact γ -actin pseudogene upstream of *AMY2B*. Sequences corresponding to 573 bp of the 3' untranslated region of the γ -actin gene were previously identified terminating approximately 200 bp upstream of the first coding exon of the human amylase genes (4, 28). The restriction map of the 5'-flanking region of *AMY2B* continues

* Corresponding author.

† Present address: Friedrich Miescher Institute, CH-4002 Basel, Switzerland.

FIG. 1. Human amylase gene complex. (a) The positions of five amylase genes and the pseudogene *AMYP1* were determined by analysis of overlapping cosmids (8, 9). Recently triplicated regions A, B, and C are bracketed (discussed in text). The approximate positions of two long interspersed repeat elements (L1) are also indicated (unpublished observations). *ACTGP3*, γ -actin pseudogene 3; *ERVA*, endogenous retrovirus, amylase associated. Symbols: $\blacksquare$, γ -actin-related sequences; $\square$, amylase genes; $\square$, retroviral sequences; +, retroviral LTRs. (b) Flanking regions of five human amylase genes. The major start sites for transcription in pancreas (Pan) and parotid gland (Sal) are indicated. The arrows beneath the LTRs indicate 5'-to-3' orientation. Symbols: $\square$, exon *a* and the nontranslated exon (NTE); $\blacksquare$, γ -actin-related sequences; $\blacksquare$, retroviral sequences.

to be related to the human γ -actin cDNA for an additional 1.4 kb upstream of this region (Fig. 2a). To determine the extent of the γ -actin-related sequences, we sequenced an additional 477 bp of the 5' end and 402 bp of an internal fragment of the γ -actin insert (Fig. 2a, regions b and c). These sequences are aligned with the human γ -actin cDNA in Fig. 2b and c. The region of similarity continuous upstream to the 5' end of the γ -actin cDNA (Fig. 2b). Overall, the sequences of regions b and c are 87% identical to the γ -actin cDNA (667 of 769 nucleotides). The γ -actin-related sequence lacks introns (Fig. 2b, arrow) and terminates with an A-rich region (28), indicating that it is a typical processed pseudogene. This full-length γ -actin pseudogene is designated *ACTGP3* (Fig. 1a), in accordance with standard nomenclature (14). A repeated sequence (TG)₂₁ is located 48 nucleotides upstream of *ACTGP3* (Fig. 2b, underlined).

LTR sequences in the amylase gene cluster. There are seven LTR-hybridizing regions in the amylase cluster, at positions upstream of *AMY2A*, *AMY1A*, *AMY1B*, and *AMY1C* (Fig. 1a, +). The *AMY1* genes are associated with a proximal and a distal LTR, while *AMY2A* is associated with one LTR (Fig. 1b). In order to determine the relationships between these

LTRs, we sequenced the *AMY2A* LTR and the distal LTR upstream of *AMY1A*. The sequencing strategy is outlined in Fig. 3a. In Fig. 3c, the sequences of these two LTRs are aligned with the previously characterized proximal LTR of *AMY1* (4). The 468-bp LTRs are closely related (89% sequence identity) and appear to be members of the 4-1 family of human endogenous retroviruses (34). Like other members of this family, the RNA primer-binding site adjacent to the 5' LTR is complementary to the 3' end of rat glutamic acid tRNA (17 of 18 nucleotides) and is separated from the LTR by the pentanucleotide TTTCT (Fig. 3c). Comparison of the *AMY2A* LTR with clone 4-14 of the 4-1 family (34) revealed 75% sequence identity (377 of 502 nucleotides). The proximal and distal LTRs upstream of the *AMY1* genes correspond to the retroviral 5' and 3' LTRs, respectively.

The LTRs contain several potential transcription control elements, including a TATA box, a CCAAT box, a polyadenylation signal, and glucocorticoid response elements (Fig. 3c). Also of interest are a 10-bp direct repeat and the overlapping 9-bp inverted and direct repeats within the U3 regions (Fig. 3c, arrows). The LTR sequences begin with TG and end with CA, as is commonly observed (23).

FIG. 2. Comparison of the 5'-flanking region of *AMY2B* with the human γ -actin cDNA. γ -Actin cDNA sequence is taken from Erba et al. (6). (a) Restriction maps and sequencing strategy. Arrows indicate the direction and extent of sequencing. Primers whose sequences are bracketed in sections b and c (○) and the previously published sequence (---) (28) are indicated. B, *Bgl*II; S, *Sca*I; H, *Hind*III; EV, *Eco*RV; Hp, *Hpa*I; E, *Eco*RI. (b) Comparison of γ -actin cDNA sequence +1 to +348 with region b of *AMY2B*. Identical nucleotides (.) and deleted nucleotides (-) are indicated. The arrow indicates the position of the first intron in the γ -actin gene (5). (c) Comparison of γ -actin cDNA sequences +1088 to +1504 with region c of *AMY2B* (-1029 to -628). The 4-nucleotide target site for retroviral insertion is boxed.

The retroviral sequences adjacent to the primer binding site are *gag* related (4), and those adjacent to the distal LTR are *env* related (data not shown). This sequence information, together with the 9-kb distance between LTRs (28), indicates the presence of an intact retroviral genome in the opposite transcriptional orientation to amylase. The restriction maps of the three *AMYI* genes are identical throughout a 27-kb

region which includes both LTRs (9, 28). Each *AMYI* gene thus appears to be associated with an intact retrovirus. These three members of the 4-1 family of endogenous retroviruses have been designated *ERVA1A*, *ERVA1B*, and *ERVA1C* (Fig. 1a, open boxes).

The retroviral inserts interrupt the 3' untranslated region of the γ -actin pseudogenes flanking the *AMY2A* and *AMYI* genes. In contrast, the *AMY2B*-associated pseudogene *ACTGP3* is intact in this region (Fig. 2c). One characteristic of retroviral insertion is the duplication of four to six nucleotides of host DNA at the site of integration. Alignment of the *ACTGP3* pseudogene with the junction regions surrounding the retroviral inserts revealed the expected duplication of four nucleotides, with one subsequent T-to-C substitution adjacent to the distal LTR (Fig. 3b).

The 5'-flanking region of *AMY2A*. The 5'-flanking region of the pancreatic amylase gene *AMY2A* contains a single LTR which appears to be the product of excision of an intact retrovirus by homologous recombination between 5' and 3' LTRs. This is demonstrated by the similarity between the sequences flanking the *AMY2A* LTR and the sequences flanking the *ERVA1A* retrovirus. The two copies of the tetranucleotide which were generated by retroviral insertion are retained in *AMY2A* (Fig. 3b). The sequences of an additional 150 bp upstream of the retroviral inserts are presented in Fig. 4. The γ -actin pseudogene sequence is truncated 76 bp upstream of both the *AMY2A* LTR and the *AMYI* distal LTR. Upstream of the truncated γ -actin pseudogene, the adjacent genomic DNA is 93% identical (77 of 83 nucleotides). It is evident from these comparisons that the sequences surrounding the *AMY2A* LTR are closely related to the corresponding sequences surrounding the intact *AMYI* retrovirus. Retroviral excision would result in this residual structure.

Estimation of the time of origin of the γ -actin and retroviral inserts. The time of origin of the γ -actin pseudogene *ACTGP3* was estimated from the extent of divergence of the pseudogene sequence from the functional human and mouse γ -actin genes. The sequence of the 3' untranslated region of the pseudogene *ACTGP3* was manually aligned with the corresponding sequences of the human (6) and mouse (24) cDNAs (alignments available upon request). We observed 63 of 686 nucleotide substitutions (9.2% divergence) between *ACTGP3* and the human cDNA; 139 of 665 nucleotide substitutions (20.9% divergence) between *ACTGP3* and the mouse cDNA; and 118 of 687 nucleotide substitutions (17.2% divergence) between the human and mouse cDNAs. The pseudogene is clearly more closely related to the human gene than to the mouse gene. The method of Fitch and Margoliash as detailed by Wu et al. (36) was used to calculate the distribution of divergence among the three sequences (Fig. 5). The 9.2% divergence between the functional human gene and the pseudogene was divided into 6.5% divergence due to base substitutions within the pseudogene since its origin and 2.7% divergence due to base substitution in the functional gene during the same period. If the average rate for divergence of ancient primate pseudogenes is 1.5×10^{-9} substitutions per site per year (15), then the observed 0.065 substitutions per site in the pseudogene would require 43 million years to accumulate. This is a minimal estimate for the time of origin of the pseudogene, since functional constraints may have reduced the rate of divergence.

The time of insertion of the retrovirus can be estimated from the 11.6% sequence divergence between the 5' and 3' LTRs (52 of 448 nucleotide substitutions [Fig. 3c]), since

FIG. 4. The 5'-flanking region of *AMY2A* is more closely related to *AMY1* than to *AMY2B*. (a) Structural comparison of the three flanking regions. Symbols: ▨, γ -actin sequence; ▩, retroviral sequence. (b) Sequences of the regions depicted in panel a. The open arrow marks the termination of the γ -actin sequence in *AMY2A* and *AMY1*. The retroviral LTR and adjacent duplicated tetranucleotide are boxed.

retroviral LTRs are identical at the time of insertion. Assuming an equal substitution rate for the two LTRs, the divergence is distributed into 0.058 substitutions per site for each LTR. Using the rate formula applied above, we estimate the time of insertion of the retrovirus to be 39 million years ago. Thus, the retrovirus integrated into the amylase gene complex soon after the γ -actin pseudogene insertion.

Association of amylase and γ -actin sequences in other

FIG. 5. Phylogenetic tree of γ -actin sequences. Sequence divergence of the 3'-untranslated regions of the human and mouse γ -actin genes and the γ -actin pseudogene *ACTGP3* were calculated as described in the text. O refers to the ancestral mammalian γ -actin gene. Myr, Million years.

primate genomes. To determine whether γ -actin-related sequences are associated with amylase genes of other primates, genomic DNA was isolated from human, chimpanzee, gorilla, orangutan, and Old World monkey (pigtail macaque) and hybridized with γ -actin and amylase probes. The γ -actin probe contained the 3'-untranslated region of the pseudogene and did not cross-hybridize with the α - and β -actin genes. The amylase probe contained exon *a* and hybridized with salivary and pancreatic amylase genes.

The γ -actin probe hybridized with numerous *EcoRI* fragments (Fig. 6b), reflecting the presence of multiple γ -actin pseudogenes in these species (26). Each *EcoRI* fragment hybridizing with the amylase probe corresponded in length to a fragment hybridizing with the γ -actin probe (Fig. 6, arrows). Identity of γ -actin- and amylase-hybridizing fragments was also observed after digestion of the genomic DNAs with *HindIII* (not shown). The apparent association of γ -actin and amylase sequences in all of these species is consistent with γ -actin pseudogene insertion before the divergence of apes and Old World monkeys, which is estimated to have occurred approximately 25 million years ago (7, 25).

Recent origin of three *AMY1* gene copies. The amylase-hybridizing fragments in human and chimpanzee DNA are identical in length and are derived from human *AMY2A*, *AMY2B*, and *AMY1* (Fig. 6, lanes H and C). These three bands were quantitated by scanning with a laser densitometer. The intensity of the bands from *AMY2A* and *AMY2B* were equivalent in the two species. However, the *AMY1* fragment in human DNA hybridized three times more intensely than the fragment in chimpanzee DNA. This obser-

FIG. 6. Association of γ -actin and amylase sequences in various primate species. Genomic DNA was digested with *Eco*RI and analyzed by the method of Southern. Duplicate blots were hybridized either with amylase exon *a* (a) or the 3'-untranslated region of the γ -actin pseudogene (b). The human *Eco*RI fragments which hybridized with the amylase probe were derived from the genes indicated at the left. Arrows (b) indicate restriction fragments which hybridized with both probes. The positions of λ *Hind*III-digested size markers are indicated to the right of each blot. H, Human; C, chimpanzee; G, gorilla; O, orangutan; M, pigtail macaque.

vation indicates that the human salivary amylase gene was amplified from one copy to three copies after the divergence of humans and chimpanzees, which has been estimated at 5 to 7 million years ago (29, 32). This recent date is consistent with the lack of sequence divergence between the three human salivary amylase genes (0 of 950 nucleotides) (9).

The copy number for *AMY1* in the other primate DNAs could not be estimated because they contain amylase-hybridizing restriction fragments of various lengths which cannot be recognized as originating from salivary or pancreatic amylase genes.

DISCUSSION

The common integration site for the γ -actin pseudogene provides strong evidence that the five human amylase genes diverged after the pseudogene insertion, which occurred approximately 40 million years ago. Comparison of the retroviral inserts indicated that the order of appearance was *AMY2B* followed by *AMY1* and then by *AMY2A*. A model which accounts for this evolutionary sequence is presented in Fig. 7. The proposed events include (i) insertion of a γ -actin pseudogene upstream of the ancestral gene copy, (ii) duplication of the amylase gene with the adjacent γ -actin pseudogene, (iii) retroviral insertion into one γ -actin pseudogene, (iv) duplication of the amylase gene with the associated retrovirus, (v) homologous recombination between LTRs to generate the residual LTR in *AMY2A*, and (vi) recent triplication of the salivary amylase gene. It is unusual for the sequence of events during molecular evolution of a gene family to be so clearly indicated by the resulting gene structures. The proposed recent origin of the

FIG. 7. Proposed sequence of events during the evolution of the human amylase gene family. The partial deletion of the γ -actin pseudogene upstream of *AMY1* and *AMY2A* occurred between the first two gene duplications.

human amylase genes is consistent with the previously observed 98% sequence identity of amylase cDNAs isolated from human pancreas and parotid gland (20).

There is significant conservation of the proximal promoter regions of the human and mouse pancreatic amylase genes (9), indicating that pancreas specificity is determined by a common mechanism in the two species. This region of the mouse gene contains an enhancer element which is recognized by the pancreas-specific nuclear protein PTF1 (3, 13). The rodent PTF1-binding sites are aligned with the human genes in Fig. 8. Conservation of this sequence in the human pancreatic amylase genes suggests that these elements are functional PTF1-binding sites.

The proposed derivation of human salivary and pancreatic amylase genes from a single precursor (Fig. 7) implies a change in tissue specificity of the *AMY1* genes during primate evolution. It is not clear whether the salivary-specific determinant of human *AMY1* arose de novo or was conserved from a preexisting element. There is no evident sequence conservation between the promoters of the human and mouse salivary amylase genes (9). It is possible that retroviral sequences are responsible for the salivary gland specificity of human *AMY1*, since the intact retrovirus is associated with salivary but not pancreatic amylase genes (Fig. 1b). Experiments are in progress to identify the salivary gland-specific determinant of the *AMY1C* gene by using expression studies in transgenic mice. With regard to the loss of expression in pancreas, it is interesting that *AMY1* contains a T-to-A substitution at a position which is invariant in functional PTF1-binding sites (Fig. 8) (3). If this nucleo-

FIG. 8. Sequence comparison of the human amylase genes (Hum) with the mouse (Mus) and rat pancreatic amylase genes in the region of the PTF1-binding site. Human *AMY1* sequence is numbered in accordance with the corresponding *AMY2* sequence. The rodent consensus was determined by analysis of the PTF1-binding sites in nine pancreas-specific genes (1, 3, 13). N_x varies from 5 to 15 bp in the rodent genes. Human amylase sequences are from Gumucio et al. (9), mouse pancreatic amylase sequences are from Osborn et al. (22), and rat pancreatic amylase sequences are from Boulet et al. (1). R, Purine residue.

tide substitution prevents binding of PTF1, it could account for the observed lack of *AMY1* transcription in pancreas (28).

Another example of a recently expanded gene family with divergent tissue specificity is the human growth hormone cluster (2). The five growth hormone-related genes were estimated to have evolved from a single precursor within the past 10 million years. Four of the genes are expressed in placenta, while the fifth is expressed exclusively in the anterior pituitary. The molecular basis for this difference in tissue specificity is not known.

The human genome contains several thousand endogenous retroviruses. The three retroviruses in the amylase gene cluster appear to be members of the 4-1 family, which contains approximately 50 members and is related to the baboon endogenous virus and the Moloney murine leukemia virus (18, 27, 34). Retroviral insertion into the amylase-associated γ -actin pseudogene occurred approximately 40 million years ago. Transcription of the *AMY1* genes is initiated within the γ -actin pseudogene at a position only 250 bp downstream of the retroviral LTR (Fig. 1b). We have not detected any transcripts originating from this position of the *AMY2B* gene, which lacks the retroviral insert (28). It thus appears that insertion of the retrovirus resulted in activation of a cryptic promoter within the γ -actin pseudogene. It is interesting that three independent insertions of retroviral elements into mouse γ -actin pseudogenes have also been reported (17).

A (TG)₂₁ repeated sequence is located 2 kb upstream of the *AMY2B* gene, close to the site of integration of the γ -actin pseudogene. There are approximately 10⁵ (TG)_n repeats in the human genome (10). Variation in the length of these repeats has recently been exploited as a source of a polymorphic variation in the human population (16, 35). Using the polymerase chain reaction, we observed six common alleles of the amylase TG repeat (N. C. Dracopoli and M. H. Meisler, Genomics, in press). This informative polymorphism provides a useful marker for the genetic and physical maps of human chromosome region 1p21.

Analysis of the 5'-flanking regions of the human amylase genes has revealed a series of molecular events during the evolution of this gene cluster. The results demonstrate the contributions of DNA insertions, deletions, and duplications to rapid molecular change in mammalian evolution.

ACKNOWLEDGMENTS

We thank D. Hewett-Emmett and W.-H. Li for advice on calculating the timing of the γ -actin pseudogene insertion, P. Venta and J. L. Slightom for providing the primate genomic DNAs, and D. Ginsberg for critical reading of an earlier version of the manuscript.

This work was supported by Public Health Service grant GM 24872 from the National Institutes of Health.

LITERATURE CITED

- Boulet, A. M., C. R. Erwin, and W. J. Rutter. 1986. Cell-specific enhancers in the rat exocrine pancreas. *Proc. Natl. Acad. Sci. USA* 83:3599-3603.
- Chen, E. Y., Y.-C. Liao, D. H. Smith, H. A. Barrera-Saldana, R. E. Gelinas, and P. H. Seeburg. 1989. The human growth hormone locus: nucleotide sequence, biology, and evolution. *Genomics* 4:479-497.
- Cockell, M., B. J. Stevenson, M. Strubin, O. Hagenbuehle, and P. K. Wellauer. 1989. Identification of a cell-specific DNA-binding activity that interacts with a transcriptional activator of genes expressed in the acinar pancreas. *Mol. Cell. Biol.* 9:2464-2476.
- Emi, M., A. Horii, N. Tomita, T. Nishide, M. Ogawa, T. Mori, and K. Matsubara. 1988. Overlapping two genes in human DNA: a salivary amylase gene overlaps with a gamma-actin pseudogene that carries an integrated human endogenous retroviral DNA. *Gene* 62:229-235.
- Erba, H. P., R. Eddy, T. Shows, L. Kedes, and P. Gunning. 1988. Structure, chromosomal location, and expression of the human γ -actin gene: Differential evolution, location, and expression of the cytoskeletal β - and γ -actin genes. *Mol. Cell. Biol.* 8:1775-1789.
- Erba, H. P., P. Gunning, and L. Kedes. 1986. Nucleotide sequence of the human γ -cytoskeletal actin mRNA: anomalous evolution of vertebrate nonmuscle actin genes. *Nucleic Acids Res.* 14:5275-5294.
- Gingerich, P. D. 1984. Primate evolution: evidence from the fossil record, comparative morphology, and molecular biology. *Yearb. Phys. Anthropol.* 27:57-72.
- Groot, P. C., M. J. Bleeker, J. C. Pronk, F. Arwert, W. H. Mager, R. J. Planta, A. W. Eriksson, and R. R. Frants. 1989. The human α -amylase multigene family consists of haplotypes with variable numbers of genes. *Genomics* 5:29-42.
- Gumucio, D. L., K. Wiebauer, R. M. Caldwell, L. C. Samuelson, and M. H. Meisler. 1988. Concerted evolution of human amylase genes. *Mol. Cell. Biol.* 8:1197-1205.
- Hamada, H., M. G. Petrino, and T. Kakunaga. 1982. A novel repeated element with Z-DNA-forming potential is widely found in evolutionarily diverse eukaryotic genomes. *Proc. Natl. Acad. Sci. USA* 79:6465-6469.
- Handy, D. E., S. H. Larsen, R. C. Karn, and M. E. Hodes. 1987. Identification of a human salivary amylase gene. *Mol. Biol. Med.* 4:145-155.
- Horii, A., M. Emi, N. Tomita, T. Nishide, M. Ogawa, T. Mori, and M. Matsubara. 1987. Primary structure of human pancreatic α -amylase gene: Its comparison with human salivary α -amylase gene. *Gene* 60:57-64.
- Howard, G., P. R. Keller, T. M. Johnson, and M. H. Meisler. 1989. Binding of a pancreatic nuclear protein is correlated with amylase enhancer activity. *Nucleic Acids Res.* 17:8185-8195.
- Human Gene Mapping 8. 1985. Eighth International Workshop

- on Human Gene Mapping. *Cytogenet. Cell. Genet.* **40**:8–10.
15. Li, W.-H., M. Tanimura, and P. M. Sharp. 1987. An evaluation of the molecular clock hypothesis using mammalian DNA sequences. *J. Mol. Evol.* **25**:330–342.
 16. Litt, M., and J. A. Luty. 1989. A hypervariable microsatellite revealed by *in vitro* amplification of dinucleotide repeat within the cardiac muscle actin gene. *Am. J. Hum. Genet.* **44**:397–401.
 17. Man, Y. M., H. Delius, and D. P. Leader. 1987. Molecular analysis of elements inserted into mouse γ -actin processed pseudogenes. *Nucleic Acids Res.* **15**:3291–3304.
 18. Martin, M. A., T. Bryan, S. Rasheed, and A. S. Khan. 1981. Identification and cloning of endogenous retroviral sequences present in human DNA. *Proc. Natl. Acad. Sci. USA* **78**:4892–4896.
 19. Melton, D. A., P. A. Krieg, M. R. Rebagliati, T. Maniatis, K. Zinn, and M. R. Green. 1984. Efficient *in vitro* synthesis of biologically active RNA and RNA hybridization probes from plasmids containing a bacteriophage SP6 promoter. *Nucleic Acids Res.* **12**:7035–7056.
 20. Nishide, T., M. Emi, Y. Nakamura, and K. Matsubara. 1986. Corrected sequences of cDNAs for human salivary and pancreatic α -amylases. *Gene* **50**:371–372.
 21. Nishide, T., Y. Nakamura, M. Emi, T. Yamamoto, M. Ogawa, T. Mori, and K. Matsubara. 1986. Primary structure of human salivary α -amylase gene. *Gene* **41**:299–304.
 22. Osborn, L., M. P. Rosenberg, S. A. Keller, and M. H. Meisler. 1987. Tissue-specific and insulin-dependent expression of a pancreatic amylase gene in transgenic mice. *Mol. Cell. Biol.* **7**:326–334.
 23. Panganiban, A. T. 1985. Retroviral DNA integration. *Cell* **42**:5–6.
 24. Peter, B., Y. M. Man, C. E. Begg, I. Gall, and D. P. Leader. 1988. Mouse cytoskeletal γ -actin: analysis and implications of the structure of cloned cDNA and processed pseudogenes. *J. Mol. Biol.* **203**:665–675.
 25. Pilbeam, D. 1984. The descent of hominoids and hominids. *Sci. Am.* **252**:84–96.
 26. Ponte, P., P. Gunning, H. Blau, and L. Kedes. 1983. Human actin genes are single copy for α -skeletal and α -cardiac actin but multicopy for β - and γ -cytoskeletal genes: 3' untranslated regions are isotype specific but are conserved in evolution. *Mol. Cell. Biol.* **3**:1783–1791.
 27. Repaske, R., P. E. Steele, R. R. O'Neill, A. B. Rabson, and M. A. Martin. 1985. Nucleotide sequence of a full-length human endogenous retroviral segment. *J. Virol.* **54**:764–772.
 28. Samuelson, L. C., K. Wiebauer, D. L. Gumucio, and M. H. Meisler. 1988. Expression of the human amylase genes: recent origin of a salivary amylase promoter from an actin pseudogene. *Nucleic Acids Res.* **16**:8261–8275.
 29. Sarich, V. M., and A. C. Wilson. 1967. Immunological time scale for hominid evolution. *Science* **158**:1200–1203.
 30. Scheidreit, C., and M. Beato. 1984. Contacts between hormone receptor and DNA double helix within a glucocorticoid regulatory element of mouse mammary tumor virus. *Proc. Natl. Acad. Sci. USA* **81**:3029–3033.
 31. Sekiya, T., Y. Kuchino, and S. Nishimura. 1981. Mammalian tRNA genes: Nucleotide sequence of rat genes for tRNA^{Asp}, tRNA^{Gly}, and tRNA^{Glu}. *Nucleic Acids Res.* **9**:2239–2250.
 32. Sibley, C. G., and J. E. Ahlquist. 1984. The phylogeny of the hominoid primates, as indicated by DNA-DNA hybridization. *J. Mol. Evol.* **20**:2–15.
 33. Smith, G. C., and M. D. Summers. 1980. The bidirectional transfer of DNA and RNA to nitrocellulose or diazobenzyloxy-methyl-paper. *Anal. Biochem.* **109**:123–129.
 34. Steele, P. E., A. B. Rabson, T. Bryan, and M. A. Martin. 1984. Distinctive termini characterize two families of human endogenous retroviral sequences. *Science* **225**:943–947.
 35. Weber, J. L., and P. E. May. 1989. Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction. *Am. J. Hum. Genet.* **44**:388–396.
 36. Wu, C.-I., W.-H. Li, J. J. Shen, R. C. Scarpulla, K. J. Limbach, and R. Wu. 1986. Evolution of cytochrome c genes and pseudogenes. *J. Mol. Evol.* **23**:61–75.