Proc. Natl. Acad. Sci. USA
Vol. 92, pp. 594–598, January 1995
Genetics

Seminalplasmin: Recent evolution of another member of the neuropeptide Y gene family

HERBERT HERZOG*†, YVONNE HORT*, RAINER SCHNEIDER‡, AND JOHN SHINE*

*Garvan Institute of Medical Research, St. Vincent's Hospital, Darlinghurst, New South Wales 2010, Sydney, Australia; and †Institute of Biochemistry, University of Innsbruck, Peter-Mayr-Strasse 1a, A-6020 Innsbruck, Austria

Communicated by William J. Rutter, University of California, San Francisco, CA, August 31, 1994

ABSTRACT **Seminalplasmin, the major basic protein of bull semen, an important regulator of calcium transport in bovine sperm and a positive modulator of the zona pellucida-induced acrosome reaction, is shown to be a recently created member of the neuropeptide Y gene family. Sequence analysis of the bovine peptide YY–pancreatic polypeptide gene cluster reveals an unexpected and extensive homology between seminalplasmin and the neuropeptide Y gene family, at the level of both gene structure and primary amino acid and nucleotide sequences. The extremely high degree of homology to the peptide YY gene, in both coding and especially noncoding regions, suggests that the seminalplasmin gene has arisen by a very recent gene duplication of the bovine peptide YY gene. Despite the more than 95% nucleotide sequence identity, a few specific mutations in the seminalplasmin gene have resulted in both the loss of the amino- and carboxyl-terminal cleavage sites characteristic of all other members of the neuropeptide Y family and the acquisition of a function apparently unrelated to the neurotransmitter/endocrine role of peptide YY.**

The neuropeptide Y (NPY) gene family, which also includes peptide YY (PYY) and pancreatic polypeptide (PP), is an example of multiple gene duplication events giving rise to a range of structurally related but functionally distinct gene products (1). NPY is one of the most highly conserved peptides known (for example, with only three amino acid differences between human and shark NPY), suggesting that this peptide subserves evolutionary old and important functions. In the mammalian nervous system, NPY is one of the most abundant neuropeptides and acts both centrally and peripherally to regulate the cardiovascular system. It also modulates a wide range of other important physiological activities, including appetite, central endocrine secretion, anxiety, and reproduction (2, 3). On the other hand, PYY is secreted from endocrine cells in the lower small intestine, colon, and pancreas. It acts in an inhibitory nature on the gastrointestinal tract, including inhibition of gastric acid secretion, gastric emptying, pancreatic exocrine secretion, and gut motility (4). The third member of the NPY family, PP, is secreted by cells within the endocrine and exocrine pancreas and specifically inhibits the secretion of enzymes and bicarbonate from the exocrine pancreas (5).

Analysis of the structure and localization of the genes encoding NPY, PYY, and PP has suggested that these genes arose from an initial gene duplication event that generated the NPY and PYY genes, followed by a subsequent duplication of the PYY gene to create the PP gene. The human NPY gene has been mapped to chromosome 7, while the PYY and PP genes lie only 10 kb apart on chromosome 17q21.1 (6). Consistent with this evolution by gene duplication, all three genes share a common intron/exon structure, although the three introns of the NPY gene (in all species studied to date) are much larger than the corresponding introns in the PYY and PP genes. The overall nucleotide sequence similarity between the three members of the gene family is restricted to the coding regions and ranges from approximately 55% (NPY vs. PYY) to 38% (NPY vs. PP) and 30% (PYY vs. PP).

Recently we determined the gene structure and amino acid sequence of the human PYY gene (6). In the course of this work we noted an unusually high similarity between the signal peptide of human PYY and the major basic protein of bull semen, bovine seminalplasmin. A detailed statistical analysis of similarities between these proteins along their entire length confirmed them as being significantly homologous, suggesting that seminalplasmin may be a member of the NPY family, possibly having evolved from PYY. As no protein equivalent to seminalplasmin has been identified in humans, we determined the sequences of the members of the bovine NPY family to clarify the evolutionary origin of seminalplasmin.§

MATERIALS AND METHODS

Genomic Library Screening. A bovine genomic DNA library in Lambda DASH II (Stratagene) was screened with a 0.6-kb α-^{32}P-labeled fragment (nucleotides −16 to 572) of the human PYY gene, obtained by PCR. Phage λ DNA was transferred to Hybond N$^+$ filters (Amersham) and hybridized with the probe in a solution containing 6× SSC, 5× Denhardt's solution, 0.1% SDS, and 100 μg of denatured and sheared salmon sperm DNA per ml at 55°C for 16 hr (1× SSC = 0.15 M NaCl/0.015 M sodium citrate, pH 7; Denhardt's solution = 0.02% polyvinylpyrrolidone/0.2% Ficoll/0.02% bovine serum albumin). Filters were washed twice for 15 min in 2× SSC containing 0.1% SDS at 55°C followed by a 15-min wash in 0.1× SSC containing 0.1% SDS and exposed to x-ray film (Kodak, X-Omat) with an intensifying screen at −70°C for 16 hr. Phage λ DNA from positive plaques was isolated by standard procedures (7). The DNA was digested with *Eco*RI, *Hin*dIII, *Bam*HI, and *Pst* I and then subcloned into the Bluescript SK vector (Stratagene), generating clones covering all of the bovine PYY and PP genes.

Nucleotide Sequence Determination. Supercoiled plasmid DNA was alkaline-denatured and sequenced by the dideoxy chain-termination method using T7 polymerase (Promega) (7). The oligonucleotide primers used initially were complementary to the flanking region of the vector and then were based on genomic sequences obtained to complete the sequence analysis.

Restriction Map Determination. Phage λ DNA was digested with restriction enzymes *Eco*RI, *Not* I, *Bam*HI, *Hin*dIII, and *Sma* I alone and in all possible combinations, electrophoresed

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

Abbreviations: NPY, neuropeptide Y; PYY, peptide YY; PP, pancreatic polypeptide.

†To whom reprint requests should be addressed at: Garvan Institute of Medical Research, 384 Victoria Street, Darlinghurst, New South Wales 2010, Sydney, Australia.

§The nucleotide sequences reported in this paper have been deposited in the GenBank data base (accession nos. L37369 for bovine PYY gene and L33970 for bovine PP gene).

on a 0.8% agarose gel, denatured with alkali (0.4 M NaOH), transferred via capillary with 0.4 M NaOH to Hybond N^+ membranes, and hybridized with several specific oligonucleotides and genomic DNA fragments obtained from the subcloning.

RESULTS AND DISCUSSION

Cloning and characterization of the bovine PYY and PP gene cluster (Fig. 1), while confirming the close linkage of these members of the gene family, revealed extensive homology to the previously described bovine seminalplasmin precursor (8). Seminalplasmin, also known as bovine caltrin (calcium transport inhibitor), is a 47-amino acid peptide that has been shown to exhibit several different properties including antimicrobial activity and inhibition of lymphocyte proliferation (9). In the presence of Ca^{2+} seminalplasmin (or smaller fragments from the central region of the peptide) forms a high-affinity complex with calmodulin and inhibits its action (10). The seminalplasmin gene is apparently exclusively expressed in the male bovine reproductive tract, where it regulates Ca^{2+} transport in bovine sperm and is a positive modulator of the zona pellucida-induced acrosome reaction (11). It has been suggested that a Na^+/Ca^{2+}-antiporter, present in mammalian sperm, is the seminalplasmin-sensitive transport pathway and that inhibition of this pathway contributes to the inhibition of spontaneous acrosomal exocytosis (12). However, there is also evidence that seminalplasmin may modulate the zona pellucida agonist signaling pathways. This process includes L-type Ca^{2+} channels, which are activated by zona pellucida agonists and therefore mediate an important component of the Ca^{2+} influx required for the acrosome reaction (13). Although this mechanism is not well understood, it has been proposed that these L-type Ca^{2+} channels are functionally regulated by seminalplasmin, either indirectly via protein kinases or directly by interaction with guanine nucleotide-binding proteins (G proteins).

The nucleotide sequence of the bovine seminalplasmin gene is unexpectedly very closely related to that of the PYY gene (Fig. 2), from which it has presumably arisen by a recent gene duplication event. Furthermore, the homology found between seminalplasmin and the other members of the NPY gene family is not restricted to nucleotide and amino acid sequences but also is manifest in a similar exon/intron organization of the genes, suggesting a common ancestral gene (Fig. 1*B*). All four peptides are synthesized as a prepropeptide, consisting of a signal peptide, followed by the active peptide NPY, PYY, PP or seminalplasmin and a carboxyl-terminal flanking peptide (Fig. 3).

The overall homology at the nucleotide level between the PYY and seminalplasmin genes over 2000 bp is >95% (Fig. 2). The strongest conservation of nucleotide sequence is found in

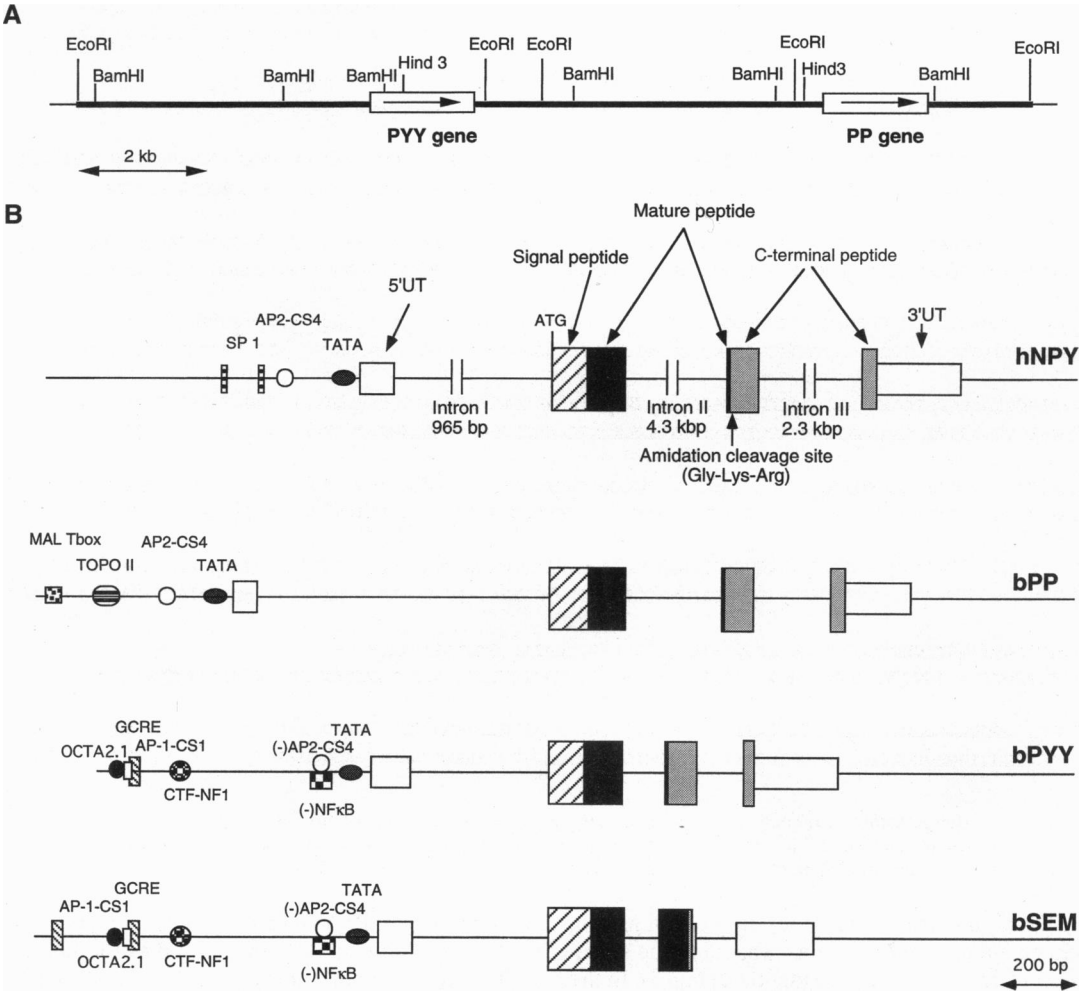

FIG. 1. (*A*) A complete restriction map including the enzymes *Bam*HI, *Eco*RI, and *Hin*dIII for the 18-kb insert of the phage λ clone 4-2 containing the bovine PYY–PP gene cluster is shown. The exact location and direction of transcription of the two genes is indicated by boxes containing an arrow. (*B*) Comparison of the exon/intron organization of the NPY gene family. A schematic representation of the genes of four members of the human and bovine NPY gene family is shown. The sequences contained in the four exons are indicated. The positions of consensus sequences for several transcription factor binding sites found in the 5′ upstream region of the genes are shown.

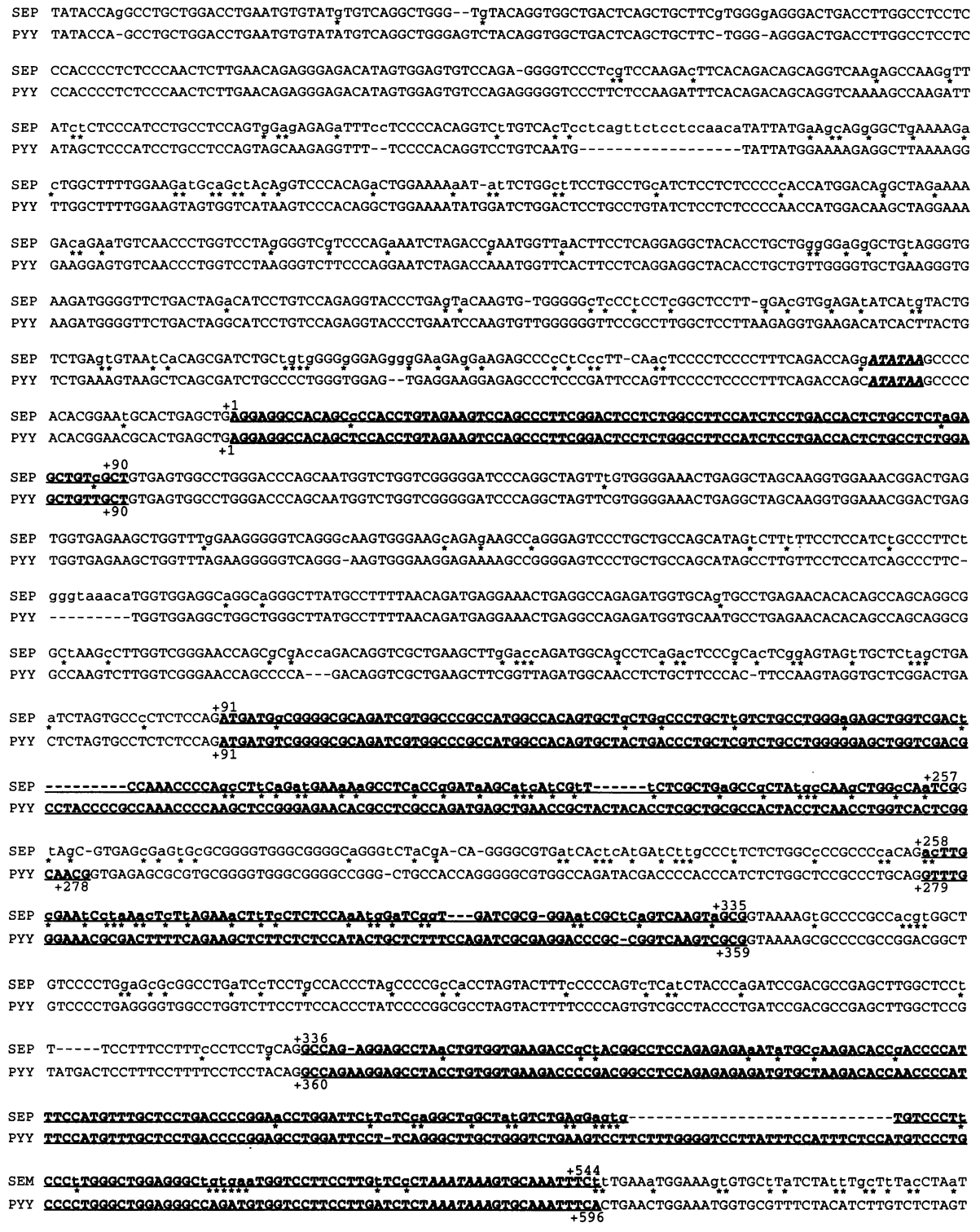

FIG. 2. Alignment of the genomic sequence of bovine seminalplasmin (14) with that of bovine PYY determined in this work. Nucleotide exchanges are indicated with an asterisk and lowercase letters. Nucleotide deletions are indicated with a dash. A potential TATA box and the exon sequences of both genes are underlined. Numbers above and below the sequences compare cDNA sequence positions.

the 5′ flanking and 5′ untranslated regions, in the first intron, and in the following nucleotide sequence encoding the signal peptide where 25 of 27 amino acids are identical (Fig. 3). In the previously reported cloning of the bovine seminalplasmin cDNA, it was suggested that the signal peptide was shorter than that shown in Fig. 3, with initiation of translation beginning at the third available methionine (14). However, in view of the amino acid identity and 98% nucleotide sequence homology to PYY in this region, it appears more likely that the first available methionine is used as the initiation site for the bovine seminalplasmin precursor. Small deletions and point mutations of the nucleotide sequence at the end of the signal peptide coding region and at several other regions within the bovine seminalplasmin gene have led to the loss of important amino acids required for the typical processing of the precursors to an active NPY, PYY, or PP peptide. For example, the

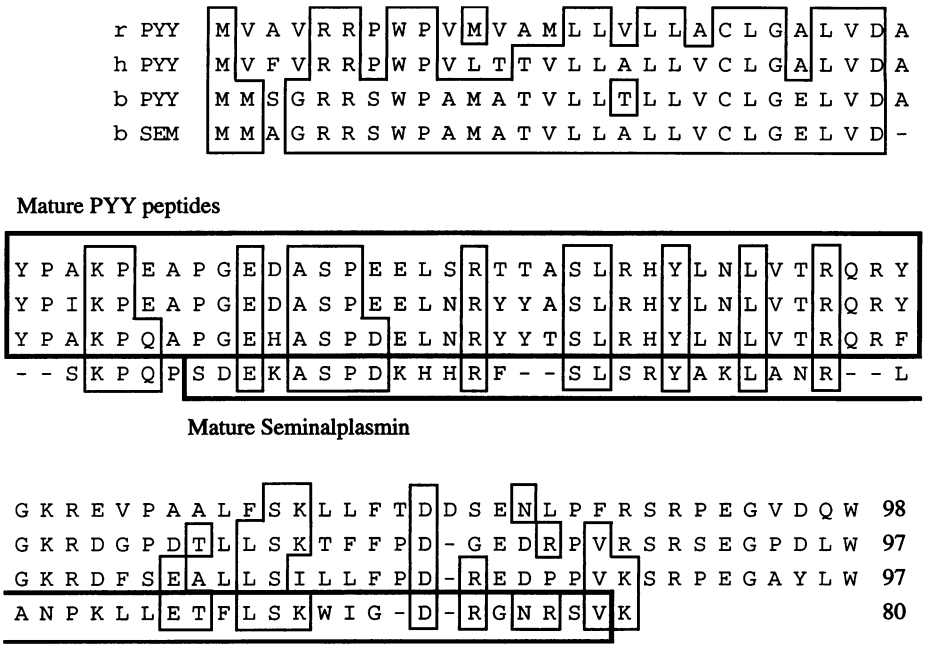

FIG. 3. Amino acid alignment of the bovine seminalplasmin precursor (b SEM; ref. 14) with all known prepro-PYY sequences: rat PYY (r PYY; ref. 4), human PYY (h PYY; ref. 6), and bovine PYY (b PYY; this work). The alignment is based on the nucleotide differences in the seminalplasmin gene, with dashes indicating deletions in the seminalplasmin gene sequence. Corresponding amino acids identical between seminalplasmin and any of the three PYY sequences are boxed. The sequence of the mature PYY peptides and of seminalplasmin is in a bold frame.

deletion of nine nucleotides in the seminalplasmin gene at the region representing the border between the signal peptide and the active peptide of the PYY gene effectively eliminates the amino-terminal tyrosine typical of NPY and PYY. This small deletion results in different processing of the seminalplasmin precursor at the amino-terminal end. It is not clear if the signal peptidase cleaves the signal peptide directly after 32 amino acids between a proline and a serine residue or if initial cleavage occurs at Ser-28 (replacing the alanine found in NPY, PYY, and PP), followed by secondary processing with a dipeptidylaminopeptidase. A further deletion of six nucleotides has led to the removal of the tyrosine and threonine codons at positions 49 and 50 of the PYY gene.

Another critical point mutation has occurred at nucleotide position 274, with the substitution of a cytosine for a thymidine, resulting in the generation of an alternative splice site in the seminalplasmin gene at the end of exon II, just five nucleotides earlier than in the PYY gene. This mutation effectively shortens the precursor peptide by two amino acids, specifically deleting the Gln-Arg-(Tyr or Phe) carboxyl-terminal region critical for receptor binding in the NPY/PYY/PP peptides (2). In both genes the same splice acceptor site at the beginning of exon III is used. However, several changes in the nucleotide sequence immediately following the splice site result in the loss of the highly conserved sequence Gly-Lys-Arg, responsible for the correct processing of the NPY, PYY, and PP precursor proteins. A point mutation from cytidine to guanine within the codon for Ser-89 in the PYY gene generates a stop codon in the seminalplasmin gene and terminates the coding sequence nine codons earlier. Both genes use the same splice sites for the end of exon III and the beginning of exon IV where, in the case of the PYY gene, the last seven amino acids of the carboxyl-terminal extension peptide are encoded.

Instead of the carboxyl-terminally amidated 36-amino-acid-long active peptide processed from NPY, PYY, and PP precursors, deletion of the processing Lys-Arg residues and introduction of an earlier termination codon results in a 47-amino-acid-long seminalplasmin. The remaining carboxyl-terminal sequence of the seminalplasmin precursor consists of only one amino acid, lysine, which is not found in the active peptide and which is possibly removed by carboxypeptidase. This is similar to the processing of the corticotropin releasing factor precursor, where only two amino acids (Gly-Lys) form the carboxyl-terminal extension (15). These changes in the seminalplasmin coding region also result in the loss of the tertiary structure, known as the PP fold, characteristic of other members of this family. Both the amino-terminal polyproline and the following amphiphilic α-helix are no longer present. Recent structural analysis of seminalplasmin suggests a random-coil organization in solution, with the ability to form helical structures in the carboxyl-terminal region only upon hydrophobic interaction with membranes (16).

The identification of bovine seminalplasmin as a very recently evolved member of the NPY gene family provides a clear example of how a gene duplication event, followed by a few selected mutations, can rapidly give rise to a new bioactive peptide with radically altered functional properties. That the creation of the bovine seminalplasmin gene is a very recent event is supported by its high degree of nucleotide sequence homology to the bovine PYY gene, together with the absence of any cross-hybridizing sequences (other than PYY) in the human genome. Although there is evidence that the human PYY–PP gene cluster localized at 17q21.1 may have undergone a duplication event (17q11), the cross-hybridizing sequences in this case are only weakly homologous and hybridize more strongly to the human PP gene than to the PYY gene (6). It thus appears that the creation of the seminalplasmin gene by duplication of the existing PYY gene occurred after the separation of the ungulates from other mammals. The observation of sequences homologous to bovine seminalplasmin in the sheep genome (8) (in addition to the PYY gene) is consistent with this view and suggests that the recently created copy of the PYY gene was conscripted to provide a new function, perhaps related to some unique features of the reproductive systems in this group of animals.

The extent of the duplication event and the involvement of other loci remain to be determined. However, it is of interest that hybridization of human PP sequences to bovine genomic DNA indicates the presence of two homologous regions, suggesting that the closely linked PP gene may also lie within the duplicated region.

With the observation that relatively few mutations in the structural regions of the newly created gene are sufficient to provide a novel function, it should be interesting to determine the effect of the similarly small number of changes to the putative regulatory regions of the gene. These are likely to be responsible for the very different patterns of tissue-specific

expression and hormonal regulation seen with PYY and seminalplasmin.

As all of the other members of the NPY gene family exert their diverse effects through activation of GTP-binding protein-coupled receptors, elucidation of the mechanism of action of seminalplasmin and isolation and characterization of its "receptor(s)" should provide further insight into the acquisition of new functional pathways during vertebrate evolution.

We gratefully acknowledge the assistance of M. Liu with DNA sequencing. This work was supported by the National Health and Medical Research Council and by the RT Hall Trust.

1. Larhammer, D., Blomqvist, A. G. & Söderberg, C. (1993) *Comp. Biochem. Physiol. C* **106,** 743–752.
2. Dumont, Y., Mantel, J. C., Fournier, A., St. Pierre, S. & Quirion, R. (1992) *Prog. Neurobiol.* **38,** 125–167.
3. Heilig, M. & Widerlöv, E. (1990) *Acta Psychiatr. Scand.* **82,** 95–114.
4. Leiter, A. B., Toder, A., Wolfe, H. J., Taylor, I. L., Cooperman, S., Mandel, G. & Goodman, R. H. (1987) *J. Biol. Chem.* **262,** 12984–12988.
5. Taylor, I. L., Solomon, T. E., Walsh, J. & Grossman, M. I. (1979) *Gastroenterology* **76,** 524–528.
6. Hort, Y. J., Baker, E., Sutherland, G., Shine, J. & Herzog, H. (1994) *Genomics*, in press.
7. Sambrook, J., Fritsch, E. F. & Maniatis, T. (1992) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab. Press, Plainview, NY), 2nd Ed.
8. Lewis, R. V., San Agustin, J., Kruggel, W. & Lardy, H. A. (1985) *Proc. Natl. Acad. Sci. USA* **82,** 6490–6491.
9. Reddy, E. S. P. & Bhargava, P. M. (1979) *Nature (London)* **279,** 725–728.
10. Comte, M., Malnoë, A. & Cox, J. A. (1986) *Biochem. J.* **240,** 567–573.
11. Florman, H. M., Corron, M. E., Kim, T. D. H. & Babcock, D. (1992) *Dev. Biol.* **152,** 304–314.
12. San Agustin, J. T. & Lardy, H. A. (1990) *J. Biol. Chem.* **265,** 6860–6867.
13. Clark, E. N., Corron, M. E. & Florman, H. M. (1993) *J. Biol. Chem.* **268,** 5309–5316.
14. Kleine-Kuhlmann, J. & Scheit, K. H. (1993) *Biochim. Biophys. Acta* **1173,** 85–86.
15. Jingami, H., Mizuno, N., Takahashi, H., Shibahara, S., Furutani, Y., Imura, H. & Numa, S. (1985) *FEBS Lett.* **191,** 63–66.
16. Maurer, T. & Rüterjans, H. (1994) *Eur. J. Biochem.* **220,** 111–116.