J Mol Evol (2005) 60:606–614
DOI: 10.1007/s00239-004-0239-9

JOURNAL OF MOLECULAR EVOLUTION

Molecular Evolution of Prolactin in Primates

O. Caryl Wallis, Akofa O. Mac-Kwashie,* Georgia Makri,† Michael Wallis

Biochemistry Department, School of Life Sciences, University of Sussex, Brighton BN1 9QG, UK

Received: 30 July 2004 / Accepted: 22 November 2004 [*Reviewing Editor:* Dr. Wen-Hsiung Li]

Abstract. Pituitary prolactin, like growth hormone (GH) and several other protein hormones, shows an episodic pattern of molecular evolution in which sustained bursts of rapid change contrast with long periods of slow evolution. A period of rapid change occurred in the evolution of prolactin in primates, leading to marked sequence differences between human prolactin and that of nonprimate mammals. We have defined this burst more precisely by sequencing the coding regions of prolactin genes for a prosimian, the slow loris (*Nycticebus pygmaeus*), and a New World monkey, the marmoset (*Callithrix jacchus*). Slow loris prolactin is very similar in sequence to pig prolactin, so the episode of rapid change occurred during primate evolution, after the separation of lines leading to prosimians and higher primates. Marmoset prolactin is similar in sequence to human prolactin, so the accelerated evolution occurred before divergence of New World monkeys and Old World monkeys/apes. The burst of change was confined largely to coding sequence (nonsynonymous sites) for mature prolactin and is not marked in other components of the gene sequence. This and the observations that (1) there was no apparent loss of function during the episode of rapid evolution, (2) the rate of evolution slowed toward the basal rate after this burst, and (3) the distribution of substitutions in the prolactin molecule is very uneven support the idea that this episode of rapid change was due to positive adaptive selection. In the slow loris and marmoset there is no evidence for duplication of the prolactin gene, and evidence from another New World monkey (*Cebus albifrons*) and from the chimpanzee and human genome sequences, suggests that this is the general position in primates, contrasting with the situation for GH genes. The chimpanzee prolactin sequence differs from that of human at two residues and comparison of human and chimpanzee prolactin gene sequences suggests that noncoding regions associated with regulating expression may be evolving differently from other noncoding regions.

Key words: Prolactin — Primates — Episodic evolution — Adaptive molecular evolution — Marmoset — Slow loris — Chimpanzee

Introduction

Prolactin is a protein hormone comprising a single polypeptide chain of about 199 residues and three disulfide bridges. It is produced and secreted by lactotrophs in the anterior pituitary gland, and in mammals it plays a major role in the regulation of lactation and mammary growth. However, it is now clear that it is also produced in many extrapituitary tissues, where it may act locally in an autocrine or paracrine manner (Ben-Jonathan et al. 1996). Prolactin is structurally similar to growth hormone (GH), and these, together with various related placental proteins and the fish hormone somatolactin, form the GH-prolactin protein family (Forsyth and Wallis 2002). This family in turn is a member of the

**Present address:* Department of Clinical Microbiology, University Hospital, SE-751 85 Uppsala, Sweden
†*Present address:* Hellenic Pasteur Institute, Vas. Sofias 127, Athens, 11521, Greece
Correspondence to: O. Caryl Wallis; *email:* C.Wallis@sussex.ac.uk

cytokine superfamily, all members of which have a characteristic folding pattern, comprising a four-helix bundle with up–up–down–down connectivity (Kossiakoff and De Vos 1999). The tertiary structure of human prolactin has been determined recently and conforms to this pattern (Keeler et al. 2003).

In humans the prolactin gene is located on chromosome 6 and has an overall organization similar to that of the GH gene on chromosome 17, with a division into five exons and four introns (Truong et al. 1984). However, the prolactin gene is much larger than the GH gene, extending over more than 10 kb of DNA, compared with 2 kb for the GH gene. Also, the prolactin gene has an alternative exon (exon 1a), which is usually utilized when the gene is expressed at extrapituitary sites. Exon 1a is about 5.7 kb upstream of exon 1b, which is used for pituitary expression (Berwaer et al. 1994). In humans there is just a single prolactin-like gene, compared with a cluster of five GH-like genes, including genes for placental lactogen and a GH variant expressed in the placenta (Chen et al. 1989). Thus human placental lactogens are structurally more similar to GH than to prolactin. In some other mammalian groups (rodents and ruminants), however, there are clusters of prolactin-like genes, many of which are expressed in the placenta, so here placental lactogens are prolactin-like rather than GH-like proteins.

The main hormonal action of prolactin in mammals is on the mammary gland, where it promotes mammary growth and milk production. Prolactin produced at extrapituitary sites has a number of actions, including actions on the immune system (Ben-Jonathan et al. 1996), but the physiological significance of these remains incompletely understood. Prolactin interacts with its target tissues via a membrane-bound receptor which is a member of the cytokine receptor superfamily (Edery et al. 2001). Binding of prolactin to the extracellular domain of the receptor leads to the recruitment of a second receptor molecule. The resulting receptor dimerization leads to interaction of intracellular domains and initiation of intracellular signalling events, particularly activation of protein kinase Jak2 and transcription factor STAT5a, followed by expression of specific genes (Goffin et al. 2002). Site-directed mutation studies have identified a number of residues on prolactin in the binding site for the first receptor molecule (binding site 1 [Goffin et al. 1992, 1996; Fuh et al. 1993]) but the second site is less well defined.

Prolactin-like hormones are found in most vertebrates, and in all but possibly the cyclostomes (Kawauchi et al. 2002) they are distinct from GH, indicating that the gene duplication event that gave rise to GH and prolactin must have occurred early in vertebrate evolution or, possibly, preceded the appearance of vertebrates. In nonmammals the function of prolactin is of course not concerned with regulation of the mammary gland, and a very broad range of actions has been described for the hormone (Nicoll 1993). Despite this, the sequence of prolactin is strongly conserved in most tetrapod groups (2000). In mammals the sequence is conserved quite strongly in many groups, but in a number of lineages a burst of rapid evolution has occurred, leading to substantial sequence variation. Such episodes of rapid change occurred during the evolution of prolactin in lineages leading to rat and mouse, elephant, ruminants and higher primates (Wallis 1981, 2000). In rodents and ruminants there were also multiple duplications of the prolactin gene leading to large families of prolactin-like proteins, many of which are expressed in the placenta (Wallis 1992; Wiemers et al. 2003).

In order to explore further the burst of rapid evolution that occurred during the evolution of prolactin in primates, we have determined the coding sequence of the prolactin genes of a New World monkey, the marmoset (*Callithrix jacchus*), and of a prosimian, the slow loris (*Nycticebus pygmaeus*). The results can be assessed in the light of parallel studies on GH genes, in which we have shown that slow loris has a single GH gene coding for a protein that is very similar to the GH of most nonprimates, but that the marmoset has a cluster of GH-like genes, comparable with, though differing in detail from, the cluster found in humans and Old World monkeys (Wallis et al. 2001; Wallis and Wallis 2002). Similar results have been reported for other prosimians and higher primates (Adkins et al. 2001; Revol de Mendoza 2004). The availability of the draft genomic sequence for the chimpanzee (*Pan troglodytes*) provides information about an additional primate prolactin gene, and this is also examined here.

Materials and Methods

Genomic DNA. Slow loris genomic DNA was a kind gift from Professor Y.-P. Zhang (Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming, China) and was prepared from the liver of a female slow loris (*Nycticebus pygmaeus*) obtained in the Yunnan province of China. Genomic DNA from a male marmoset (*Callithrix jacchus*) was the kind gift from Dr David Hunt (Institute of Opthalmology, University of London).

Polymerase Chain Reaction (PCR). Primers for PCR were obtained from MWG-Biotech (Milton Keynes, Bucks, UK) and from Sigma-Genosys (Haverhill, Suffolk, UK). Sequences of primer pairs used for amplification are given at the Web site http://www.biols.sussex.ac.uk/Home/Mike_Wallis/Prolactin/ and the regions of the prolactin genes amplified are shown in Fig. 1. Most PCR reactions were carried out using Pfu Turbo DNA polymerase or Herculase Enhanced DNA polymerase (Stratagene, La Jolla, CA, USA). Some PCR reactions with slow loris DNA were carried out using Taq DNA Polymerase (Roche Diagnostics Ltd., Lewes,

East Sussex, UK). Reactions contained 50 pmol of each primer, 100–200 ng of genomic DNA, 10 µl of 10× buffer (as provided with the enzyme used), a 0.2 m*M* concentration of each dNTP, and 2.5 units of Pfu Turbo DNA polymerase or 5 units of Herculase Enhanced DNA polymerase or Taq DNA polymerase, adjusted to a final volume of 100 µl with H_2O. PCR was carried out for 31 cycles using conditions recommended by the enzyme suppliers. Annealing temperatures ranged from 51 to 65°C, and extension times from 1 to 4 min.

Cloning and Sequencing of the Slow Loris and Marmoset Prolactin Genes. PCR products amplified using pfu Turbo or Herculase Enhanced DNA polymerase were cloned into the phagemid pPCR-Script Amp(+) vector, according to the instructions supplied with the PCR-Script AMP cloning kit (Stratagene), and transformed into ultracompetent *E. coli* cells supplied with the kit. PCR products amplified using Taq DNA polymerase were cloned into the phagemid pCR II-TOPO using the TOPO TA Cloning kit (Invitrogen Ltd., Paisley, Renfrewshire, UK) and transformed into TOP10F′ competent *E. coli* provided with the kit. Double-stranded DNA for sequencing was prepared using the QIAprep Spin Miniprep Kit (Quiagen Ltd., Crawley, West Sussex, UK). Initially DNA was subjected to sequencing using an ABI 343A automatic sequencer and dye-primer (Amersham, Little Chalfont, Bucks, UK) and dye-terminator (Perkin Elmer, Warrington, Cheshire, UK) kits. Later sequencing was by the sequencing services provided by Genetix Ltd. (New Milton, Hants, UK), SeqLab (Gottingen, Germany), or MWG-Biotech (Ebersberg, Germany).

Cloning and Characterization of Part of the Prolactin Gene of the White-Faced Capuchin. A genomic library for the white-faced capuchin (*Cebus albifrons*) in phage Charon 4A was a kind gift from Dr. Jeremy Slightom (Pharmacia & Upjohn Inc, Kalamazoo, MI, USA). Approx 2.4×10^5 plaques produced on a lawn of *E. coli*, strain LE392, were screened using a probe containing a human prolactin DNA sequence derived from cDNA (a kind gift from Dr. Joseph Martial, Universite de Liege, Liege, Belgium [Cooke et al. 1981]). *Eco*R1 fragments derived from positive plaques were cloned into phagemid pBluescript II KS(−) (Stratagene) and transformed into competent *E. coli*, strain JM109. Double-stranded DNA for sequencing was prepared using the Wizard *Plus* SV Miniprep kit (Promega UK Ltd., Southampton, UK) and sequenced as described above.

Sequence Analysis. Sequences were aligned using the Clustalw program (Higgins and Sharp 1988) with gap introduction and extension penalties of 10 and 5, respectively, followed by manual adjustment, with no increase in number of gaps. For analysis of rates of prolactin evolution a phylogenetic tree for mammalian evolution was based on the tree of Kumar and Hedges (1998), which was derived from a large amount of molecular evidence. Trees based solely on prolactin sequences deviate markedly from conventional expectations for mammalian evolution, presumably because of the very variable rates of evolution (Felsenstein 1978). Prolactin gene sequence alignments were used to determine branch lengths on this defined tree, using the neighbor-joining method in PAUP (Swofford 1998) and the parsimony-based method MacClade 3 (Maddison and Maddison 1992). Nonsynonymous (d_N) and synonymous (d_S) substitution rates in coding sequence were determined using the method of Nei and Gojobori (1986), with correction for transition/transversion ratio as given by Zhang et al. (1998) and Zhang and Nei (2000); matrices of these were used as input for the neighbor-joining option in PAUP. Significance of differences between ratios was tested using Fisher's exact test (Zhang et al. 1998). Analysis of d_N/d_S ratio (sometimes referred to as K_A/K_S) was also carried out using the paml method of Yang and Nielsen (2002). Sequence alignments and accession numbers of sequences used can be found at the Web site http://www.biols.sussex.ac.uk/Home/Mike_Wallis/Prolactin/.

Analysis of 3D Structure. The 3D structure of prolactin was analyzed using the visualization program RasMol 2.6 (Sayle and Milner-White 1995) and coordinates for human prolactin determined by Keeler et al. (2003) (pdb accession, 1N9D; the first of the nmr-derived structures was used). Numbers of intramolecular contacts for each residue within this structure were determined using RasMol or the dist program (available at the Web site given above).

Results and Discussion

Amplification and Cloning of Exons 2–5 of Prolactin Genes from Slow Loris and Marmoset DNA

Primers suitable for amplifying exons 2–5 were designed on the basis of the published sequences of human (Cooke et al. 1981), pig (Schulz-Aellen et al. 1989), and sheep (Varma et al. 1989) prolactin cDNAs and the partial sequence determined here for prolactin from *C. albifrons*. At later stages primers were designed based on the sequences derived from overlapping clones. These primers enabled production of overlapping clones encoding the region from the end of intron A to the 3′ end of the gene (containing exons 2–5). Attempts to clone exon 1b, which is expected to encode the first 10 residues of the signal peptide, were unsuccessful. For each clone the region corresponding to the exon sequence and small lengths of flanking intron were sequenced. Combining the exon sequences gave the complete coding sequences for mature prolactins from slow loris and marmoset, but only part of the corresponding signal peptides (Fig. 1). The data have been deposited with the embl/genbank/ddbj database, with accession numbers AJ784991 and AJ786353.

Cloning of Part of the Prolactin Gene from the White-Faced Capuchin

The availability of a genomic library for the white-faced capuchin (*Cebus albifrons*) enabled us to take an alternative approach to characterizing the prolactin gene of a New World monkey. Approximately 240,000 plaques were screened using labeled human prolactin cDNA probe. Three hybridizing plaques were identified and cloned, each containing an apparently identical insert of ~12 kbp. Sequence analysis showed that the clone(s) contained the 5′ end of the prolactin gene (exons 1–3), with derived amino acid (aa) sequence differing at only two residues from that obtained for the corresponding part of marmoset prolactin. Clones corresponding to the 3′ end of the

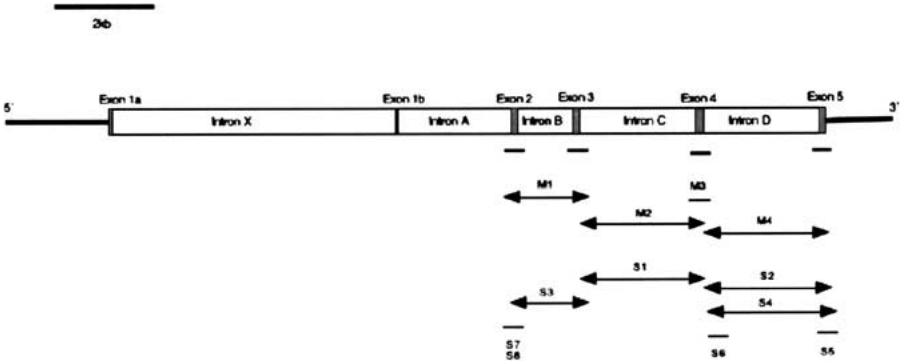

Fig. 1. Schematic representation of the primate prolactin gene, based on the human gene. Thick bars just below the figure show the regions characterized from the slow loris and marmoset prolactin genes. Arrows and thin bars below them show the regions of the gene amplified using primer pairs M1–M5 (marmoset) and S1–S8 (slow loris). Exons 2–5 encode mature prolactin; exon 1a is used for most extrapituitary expression.

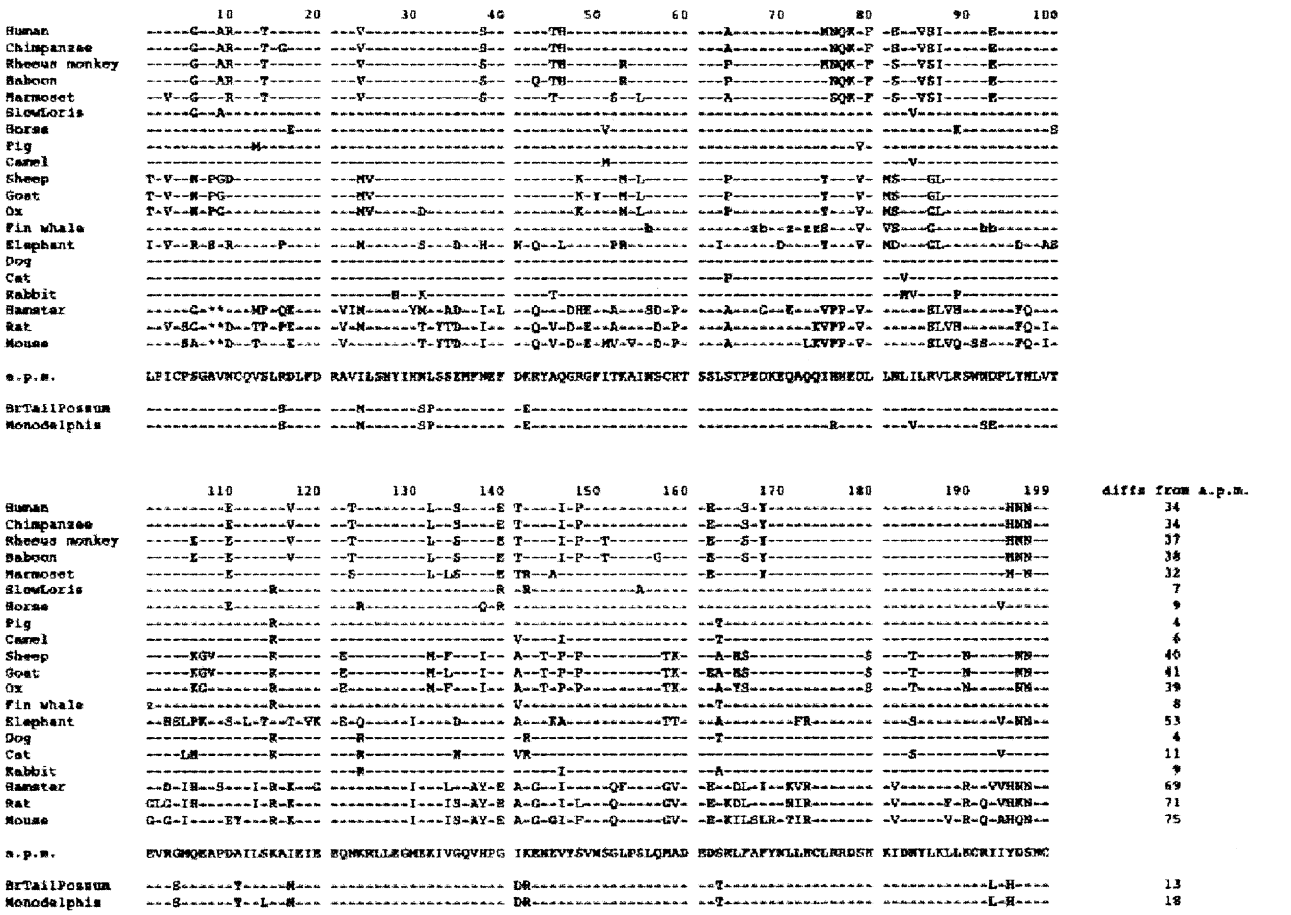

Fig. 2. Alignment of the amino acid sequences of mammalian prolactins. Sequences are compared with that deduced for the ancestral placental mammal (a.p.m.) (Wallis 2000). A hyphen indicates identity, an asterisk indicates a gap, b indicates D or N, and z indicates E or Q.

prolactin gene were not obtained for *C. albifrons*, possibly because they were underrepresented in the library. Nevertheless, this incomplete sequence provides useful confirmation that there is only a single copy of the prolactin gene in New World monkeys. The situation contrasts markedly with that for GH. Studies on the same *C. albifrons* library have revealed the presence of at least 40 GH-related genes and pseudogenes, probably arranged in a cluster. In marmoset there is a cluster of eight GH-like genes and pseudogenes (Wallis and Wallis 2002).

Slow Loris and Marmoset Prolactin

Exons 2–5 encode most of the sequence of the signal peptide and the complete sequence of mature prolactin. The derived amino acid sequences of marmoset and slow loris prolactins are shown in Fig. 2, aligned with other available mammalian prolactins. The sequence of slow loris prolactin is similar to that of pig and other conserved mammalian prolactin sequences and differs from the prolactin sequence deduced for the ancestral pla-

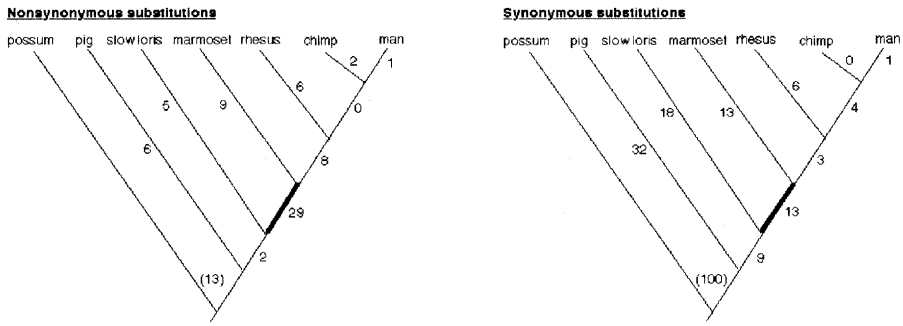

Fig. 3. Phylogenetic trees for primate prolactins, showing changes undergone by synonymous and nonsynonymous sites in the coding region for the mature protein. Pig and marsupial possum are used as outgroups.

cental mammal at only seven residues. The sequence of marmoset prolactin differs substantially (at 32 residues) from the ancestral sequence and is similar to that of human prolactin, suggesting that much, though not all, of the period of accelerated evolution occurred before divergence of lines leading to New World and Old World monkeys. This is discussed further below.

The substitutions accepted during the evolution of these primate prolactins can be assessed in the light of the 3D structure of human prolactin determined by Keeler et al. (2003), in which can be identified elements of secondary structure, hydrophobic core, and residues important for binding of receptor 1. Of the seven residues differing between slow loris prolactin and the ancestral sequence, none appears to be involved in binding site 1, and only two, Val-84 and Lys-115 (Ile-84 and Arg-115 in the ancestral sequence), are in helices. Val-84 appears to be important in the hydrophobic core of the molecule, but the Ile-to-Val substitution is also seen in camel and rabbit prolactins, suggesting that this conservative change may be of little functional significance. Lys-115 does not appear to contribute to the hydrophobic core, and the Arg-to-Lys substitution is seen in several other species.

Of the 32 residues in marmoset prolactin that differ from the ancestral sequence, again none is in binding site 1. Sixteen (~52%) are located in α-helices, but only two or three contribute substantially to the hydrophobic core. Most of these differences arose during the episode of rapid evolution and are discussed further below.

Episodic Molecular Evolution of Prolactin

Like several other polypeptide hormones, prolactin in mammals has evolved in an episodic manner, with bursts of rapid change interrupting a rather slow basal rate of evolution (Wallis 2000, 2001). The results presented above establish that the burst that occurred during primate evolution started after divergence of lines leading to slow loris and higher primates and was largely complete before divergence of lines leading to New World monkeys and Old World monkeys/apes (Fig. 3). In the period since their divergence, the branch leading to slow loris prolactin accumulated ~5 nonsynonymous substitutions, while that leading to simians and apes accumulated 38–43 nonsynonymous substitutions (Fig. 3). The rate acceleration for prolactin evolution in the line leading to simians is largely specific to nonsynonymous substitutions (those leading to amino acid changes). The line leading to slow loris prolactin accumulated about 18 synonymous substitutions, while that leading to simians and man accumulated 20–26 (Fig. 3). Therefore the ratio of nonsynonymous-to-synonymous substitutions (d_N/d_S), which has been widely used to assess the adaptive significance of periods of rapid evolution, increased to 0.87 on the branch preceding divergence of New and Old World monkeys. This compares with 0.11 on the branch leading to slow loris and 0.04–0.07 for the "basal" evolutionary periods seen for prolactin (Wallis 2000).

There are two possible explanations for this episode of accelerated evolution of prolactin in primates: (1) increased acceptance of neutral mutations associated with relaxation of purifying selection following loss of function and (2) positive selection associated with adaptive change in function. The second explanation is supported both by biological arguments and by the uneven nature of the changes seen within the prolactin molecule.

The increase in evolutionary rate is not associated with major loss of biological activity—prolactin in humans appears to have the same main physiological role in the regulation of lactation as in other mammals, with no notable difference in potency. Prolactin has a range of additional actions in mammals, the relative importance of which may vary from group to group, but there is no evidence to suggest overall loss of function in higher primates. It is also notable that after the episode of rapid change, the rate of evolution in primates appears to slow toward the basal rate. It is very difficult to see how a period of neutral

evolution following loss of functional constraints, with the acceptance of about 30 substitutions into a protein that was previously highly conserved, could lead to a sequence which regained function and was then conserved by purifying selection. More likely is that the period of rapid change corresponded to a period in which the relative importance of two or more physiological functions of prolactin varied, and slowing of the rate of evolution accompanied a stabilization of the biological function.

The value for d_N/d_S observed for the episode of rapid evolution (0.87) would be consistent with evolution driven by accumulation of neutral mutations in the almost-complete absence of functional constraint, but in that case one would expect the changes to be more or less evenly distributed through the prolactin structure. This does not appear to be the case. When the substitutions accepted were examined on the basis of the 3D structure of human prolactin (Keeler et al. 2003), it was observed that none of the 14 residues that have been identified as involved in binding of the hormone to the first receptor is changed during the episode of rapid evolution. Of 26 residues that can be identified as contributing to the hydrophobic core of prolactin, only 2 change during the burst of rapid change. Thus of 40 residues identified as playing a major part in the structure or function of the hormone, only 5% change during the episode of rapid change, compared with 14% for the rest of the molecule. These key residues appear to be conserved. An alternative approach to categorizing the structural significance of amino acid residues within the protein is to identify those which make many contacts within the structure. Eighty-five residues ''contact'' (are within 3.6 Å of) six or more nonhydrogen atoms in other residues (excluding close neighbors in the sequence—those less than 4 residues away) and 114 make five or fewer contacts. Surprisingly those that make many contacts are less conserved (d_N/d_S = 2.16) than those that make few (d_N/d_S = 0.50). Within the group of 85 residues making many contacts are included most of those identified as being in the hydrophobic core and some of those in binding site 1. If these are excluded, a subpopulation of 56 residues is identified that appear to make little contribution to the hydrophobic core or binding site 1 but that still make many internal contacts within the structure; this subpopulation included a high proportion of those residues substituted during the burst of rapid change, and gave a d_N/d_S ratio of 19.7. This is significantly higher than 1.0 (p = 0.034, Fisher's exact test). When 100 sets of 56 residues randomly selected from the prolactin sequence were examined, only 4 gave a $d_N/d_S > 19.7$, in accordance with this significance level. While the functional significance of this subset of rapidly evolving residues is not clear, their existence, and the associated evidence of a nonrandom distribution of substitutions, coupled with no major loss of biological function, does appear to support positive selection as the most probable explanation for the episode of rapid evolution.

Analysis of the data (complete mature prolactin sequences) by the codeml method (Model B [Yang and Nielsen 2002]) also showed that d_N/d_S is elevated during the episode of accelerated evolution, to a value of 2.21 (not significantly higher than 1.0; p = 0.07), and identified about 20 residues likely to be subject to positive selection.

An episode of rapid change is also seen during the evolution of primate GH, in approximately the same place/time as that seen for prolactin (Wallis 1994; Liu et al. 2001; Wallis et al. 2001). However, there are a number of differences. The burst for prolactin is less rapid than that for GH, although elsewhere in mammalian evolution episodes of rapid evolution are more marked for prolactin. The burst of change seen during the evolution of GH in primates is followed by duplication of the GH gene to give a family of GH-like proteins; this is not seen in the case of prolactin, which appears to occur as a single gene in all the primate species studied. The cases of GH and prolactin also differ in that the substitutions accepted during the episode of rapid change of GH include a high proportion of residues located in binding site 1, whereas for prolactin, as indicated above, binding site 1 is conserved. The hydrophobic core is conserved in both proteins, but in other respects there is no obvious similarity between the patterns of variation.

GC Content

Mammalian prolactin genes have a low G + C content, presumably because they are located in a GC-poor isochore (Bernardi 2000), contrasting with GH genes, which have a high GC content (Maniou et al. 2004). There is evidence that GC-rich regions are tending to disappear in primates and cetartiodactyls (Duret et al. 2002; Belle et al. 2004). We have shown previously that the episode of rapid evolution in cetartiodactyl GHs is not associated with a change in GC content (Maniou et al. 2004), and here we test whether this is also true of primate prolactins (Table 1).

Intron and 5′ flanking sequences of prolactin genes (where available) showed a low GC content for all of the prolactin genes examined, confirming that the genes are in a low GC isochore and contrasting with the situation for GHs (pig GH is included in Table 1 for comparison). The prolactin coding sequences show a higher GC content, though this is still substantially lower than that of GHs; interestingly the difference between GC contents of prolactin and GH

Table 1. Percentage of G + C in prolactin genes

				Codon position			
	5′	Intron X	CDS	1	2	3	Introns A–D
Pig			51.6	55.3	33.2	66.3	
Ox		(33.8)	50.4	55.3	37.7	58.3	37.7
Human	35.4	34.4	51.3	56.8	35.7	61.3	37.4
Chimp	34.8	34.4	51.1	56.3	35.7	61.3	(37.1)
Rhesus			50.6	55.8	36.2	59.8	
Marmoset			52.4	58.3	36.2	62.8	(35.7)
Slow loris			51.4	58.3	34.7	61.3	(37.3)
Pig GH	53.8		61.3	53.9	38.4	87.4	63.3

Note. Numbers in parentheses were derived from incomplete sequences.

Table 2. Comparison of human and chimpanzee genes

				Intron X			
	Total	Exons 1b–5	Introns A–D	5′ end	3′ end	5′ upstream	3′ downstream
Percentage identity		99.3	98.5	98.7	98.7	99.4	98.6
Ts/Tv ratio	2.0	0.25	2.6	3.3	0.88	1.7	1.8

coding sequences is mainly due to a much lower GC content at the third codon position in prolactin. There was very little variation in GC content between the prolactin genes studied, and in particular, the content for the slow loris prolactin gene is not noticeably different from that of higher primate prolactin genes. Thus, as for cetartiodactyl GHs, although the episode of rapid evolution in primate prolactin evolution involved substantial change in amino acid sequence, there was no associated change in GC content. Analysis of individual substitutions accepted during the course of primate prolactin evolution indicates that there is no significant difference between the number of A/T $\rightarrow$ G/C substitutions and the number of G/C $\rightarrow$ A/T substitutions, contrasting with the GC-rich GH genes, where G/C $\rightarrow$ A/T substitutions exceed A/T $\rightarrow$ G/C substitutions (Maniou et al. 2004).

The Chimpanzee Prolactin Gene

The episode of rapid change during primate prolactin evolution is followed by a slowing of the rate of evolution, though there is still some variation between prolactin sequences of higher primates. The availability of the chimpanzee (*Pan troglodytes*) prolactin gene sequence allows a detailed examination of a recent part of this period. A BLAST search of the draft sequence for the chimpanzee genome identified a single chimpanzee prolactin gene on scaffold_37460, extending over two contigs. This was compared in detail with the human prolactin gene. As expected, the overall structures of the two genes are similar (Table 2), with five coding exons and four introns, plus the additional noncoding exon (exon 1a), which is used in nonpituitary expression (Berwaer et al. 1994; DiMattia et al. 1990; Hiraoka et al. 1991).

Intron sequences were mostly very similar, with a few small indels in no obvious pattern. Chimpanzee intron B includes an unsequenced break corresponding to the gap between two contigs; the corresponding region in the human gene contains 597 bp, including a polypurine-rich region. Intron D includes an *Alu* sequence in both human and chimpanzee and a 64-bp insertion in the human sequence corresponding to an extended microsatellite-type repeat. Intron X (Fig. 1; spliced out between exon 1a and exon 1b in the transcript of the prolactin gene in extrapituitary tissues, but 5′ to the transcription start point used in the pituitary gland) is no more conserved than the other introns, although it includes control regions for expression in the pituitary (Peers et al. 1990).

The coding sequence of chimpanzee prolactin differs from that of human at five sites, two synonymous and three nonsynonymous substitutions. Two of the latter lead to changes in the mature protein: one (at residue 75; Met in human, Ile in chimpanzee) a conservative substitution also seen in some other primate prolactins, the other (at residue 16; Arg in human, Gly in chimpanzee) a radical change, close to binding site 1. The latter is not seen in any other mammalian prolactin and may be associated with a functional change, although confirmation of the sequence in chimpanzee is needed.

Although intron X is not more conserved than other introns, study of Ts/Tv ratios suggests that it may have evolved differently from them. For introns A–D combined, this ratio is 2.59 (96/37), reflecting the bias toward transitions usually seen in sequences evolving neutrally. Overall for intron X the ratio for this intron is 1.66, but for the 2500 bases closest to the 3′ end of the intron, which includes the main control regions for the gene expressed in the pituitary (Peers et al. 1990), the ratio is 0.88 (15/17). This is significantly different from the ratio for the rest of this intron (3.30; 33/10; $p = 0.008$) and from the ratio for introns A–D (see above; $p = 0.007$). This suggests that the substitutions seen in the 3′ part of intron X (the control region for pituitary expression, maybe including elements that inhibit expression at extrapituitary sites) reflect in part adaptive rather than neutral processes. For none of introns A–D is there a significant difference between the Ts/Tv ratio for the 5′ half and that for the 3′ half of the intron ($p > 0.4$ in all cases). For the region 5′ to exon 1a (the control region for extrapituitary expression), the sequence within 2.5 kbp of exon 1a has a slightly lower Ts/Tv ratio (1.4) than the sequence 2.5 kbp farther upstream (2.0), but the difference is not significant. More notable is the low overall substitution rate in the whole region 5 kbp upstream of exon 1a, which is 40% of that in introns A–D ($p < 0.001$); thus this region seems to be strongly conserved.

In conclusion, prolactin sequences in lower primates are strongly conserved, and similar to those of some nonprimate mammals such as pig. An episode of rapid, probably adaptive change occurred before divergence of Old and New World monkeys. The rate of evolution of prolactin then slowed again, but the prolactin gene may still have been subject to adaptive change, since detailed comparison of human and chimpanzee prolactin gene sequences suggests that differences seen between regions associated with controlling expression in the pituitary may have adaptive significance. Associated differences in the biological properties of prolactin have not been identified, but they may involve changes in the relative importance of the various actions of the hormone. A similar explanation (function switching) has been proposed as underlying the episode of rapid evolution in primate GHs (Wallis 1997), and given that there is some overlap in the actions of GH and prolactin in primates, the same fluctuating physiological requirements may have affected both hormones. Duplications of the GH gene may have led to the end of the rapid burst, by providing extra hormones to take over one or more of the physiological functions previously met by GH and prolactin (Forsyth and Wallis 2002).

Acknowledgments. We thank Dr. David Hunt, Dr. Ya-Ping Zhang, Dr. Jeremy Slightom, and Dr. Joseph Martial for providing DNA samples, genomic libraries, and plasmids.

References

Adkins RM, Nekrutenko A, Li W-H (2001) Bushbaby growth hormone is much more similar to nonprimate growth hormones than to rhesus monkey and human growth hormones. Mol Biol Evol 18:55–60

Belle EMS, Duret L, Galtier N, Eyre-Walker A (2004) The decline of isochores in mammals: an assessment of the GC content variation along the mammalian phylogeny. J Mol Evol 58:653–660

Ben-Jonathan N, Mershon JL, Alien DL, Steinmetz RW (1996) Extrapituitary prolactin: Distribution, regulation, functions, and clinical aspects. Endocr Rev 17:639–669

Bernardi G (2000) Isochores and the evolutionary genomics of vertebrates. Gene 241:3–17

Berwaer M, Martial JA, Davis JRE (1994) Characterization of an up-stream promoter directing extrapituitary expression of the human prolactin gene. Mol Endocrinol 8:635–642

Chen EY, Liao Y-C, Smith DH, Barrera-Saldana HA, Gelinas RE, Seeburg PH (1989) The human growth hormone locus: Nucleotide sequence, biology, and evolution. Genomics 4:479–497

Cooke NE, Coit D, Shine J, Baxter JD, Martial JA (1981) Human prolactin. cDNA structural analysis and evolutionary comparisons. J Biol Chem 256:4007–4016

DiMattia GE, Gellersen B, Duckworth ML, Friesen HG (1990) Human prolactin gene expression. The use of an alternative noncoding exon in decidua and the IM-9-P3 lymphoblast cell line. J Biol Chem 265:16412–16421

Duret L, Semon M, Piganeau G, Mouchiroud D, Galtier N (2002) Vanishing GC-rich isochores in mammalian genomes. Genetics 162:1837–1847

Edery M, Binart N, Bouchard B, Goffin V, Kelly PA (2001) Prolactin receptors. In: Horseman ND (ed) Prolactin. Kluwer Academic, Boston, pp 341–353

Felsenstein J (1978) Cases in which parsimony or compatability methods will be positively misleading. Syst Zool 27:401–410

Forsyth IA, Wallis M (2002) Growth hormone and prolactin—Molecular and functional evolution. J Mammary Gland Biol Neoplasia 7:291–312

Fuh G, Colosi P, Wood WI, Wells JA (1993) Mechanism-based design of prolactin receptor antagonists. J Biol Chem 268:5376–5381

Goffin V, Binart N, Touraine P, Kelly PA (2002) Prolactin: the new biology of an old hormone. Annu Rev Physiol 64:47–67

Goffin V, Norman M, Martial JA (1992) Alanine-scanning mutagenesis of human prolactin: importance of the 58–74 region for bioactivity. Mol Endocrinol 6:1381–1392

Goffin V, Shiverick KT, Kelly PA, Martial JA (1996) Sequence-function relationships within the expanding family of prolactin, growth hormone, placental lactogen, and related proteins in mammals. Endocr Rev 17:385–410

Higgins DG, Sharp PM (1988) CLUSTAL: A package for performing multiple sequence alignment on a microcomputer. Gene 73:237–244

Hiraoka Y, Tatsumi K, Shiozawa M, Aiso S, Fukasawa T, Yasuda K, Miyai K (1991) A placenta-specific 5′ non-coding exon of human prolactin. Mol Cell Endocrinol 75:71–80

Kawauchi H, Suzuki K, Yamazaki T, Moriyama S, Nozaki M, Yamaguchi K, Takahashi A, Youson J, Sower SA (2002) Identification of growth hormone in the sea lamprey, an extant representative of a group of the most ancient vertebrates. Endocrinology 143:4916–4921

Keeler C, Dannies PS, Hodsdon ME (2003) The tertiary structure and backbone dynamics of human prolactin. J Mol Biol 328:1105–1121

Kossiakoff AA, De Vos AM (1999) Structural basis for cytokine hormone-receptor recognition and receptor activation. Adv Protein Chem 52:67–108

Kumar S, Hedges SB (1998) A molecular timescale for vertebrate evolution. Nature 392:917–920

Liu J-C, Makova KD, Adkins RM, Gibson S, Li W-H (2001) Episodic evolution of growth hormone in primates and emergence of the species specificity of human growth hormone receptor. Mol Biol Evol 18:945–953

Maddison WP, Maddison DR (1992) MacClade: Analysis of phylogeny and character evolution, version 3.0 Sinauer, Sunderland, MA

Maniou Z, Wallis OC, Wallis M (2004) Episodic molecular evolution of pituitary growth hormone in Cetartiodactyla. J Mol Evol 58:743–753

Nei M, Gojobori T (1986) Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. Mol Biol Evol 3:418–426

Nicoll CS (1993) Role of prolactin and placental lactogens in vertebrate growth and development. In: Schreibman MP, Scanes CG, Pang PKT (eds) The endocrinology of growth, development and metabolism in vertebrates. Academic Press, San Diego, pp 183–196

Peers B, Voz ML, Monget P, Mathy-Hartert M, Berwaer M, Belayew A, Martial JA (1990) Regulatory elements controlling pituitary-specific expression of the human prolactin gene. Mol Cell Biol 10:4690–4700

Revol de Mendoza A. Esquivel Escobedo D, Martinez Davila I, Barrera Saldana H (2004) Expansion and divergence of the GH locus between spider monkey and chimpanzee. Gene 336:185–193

Sayle R, Milner-White EJ (1995) RasMol: Biomolecular graphics for all. Trends Biochem Sci 20:374–376

Schulz-Aellen M-F, Schmid E, Movva RN (1989) Nucleotide sequence of porcine preprolactin cDNA. Nucl Acids Res 17:3295

Swofford DL (1998) PAUP*. Phylogenetic analysis using parsimony (* and other methods), version 4 Sinauer, Sunderland, MA

Truong AT, Duez C, Belayew A, Renard A, Pictet R, Bell GI, Martial JA (1984) Isolation and characterization of the human prolactin gene. EMBO J 3:429–437

Varma S, Kwok S, Ebner KE (1989) Cloning and nucleotide sequence of ovine prolactin cDNA. Gene 77:349–359

Wallis M (1981) The molecular evolution of pituitary growth hormone, prolactin and placental lactogen: A protein family showing variable rates of evolution. J Mol Evol 17:10–18

Wallis M (1992) The expanding growth hormone/prolactin family. J Mol Endocr 9:185–188

Wallis M (1994) Variable evolutionary rates in the molecular evolution of mammalian growth hormones. J Mol Evol 38:619–627

Wallis M (1997) Function switching as a basis for bursts of rapid change during the evolution of pituitary growth hormone. J Mol Evol 44:348–350

Wallis M (2000) Episodic evolution of protein hormones: Molecular evolution of pituitary prolactin. J Mol Evol 50:465–473

Wallis M (2001) Episodic evolution of protein hormones in mammals. J Mol Evol 53:10–18

Wallis OC, Wallis M (2002) Characterisation of the GH gene cluster in a new-world monkey, the marmoset (*Callithrix jacchus*). J Mol Endocrinol 29:89–97

Wallis OC, Zhang Y-P, Wallis M (2001) Molecular evolution of GH in primates: characterization of the GH genes from slow loris and marmoset defines an episode of rapid evolutionary change. J Mol Endocrinol 26:249–258

Wiemers DO, Shao L-J, Ain R, Dai G, Soares MJ (2003) The mouse prolactin gene family locus. Endocrinology 144:313–325

Yang Z, Nielsen R (2002) Codon-substitution models for detecting molecular adaptation at individual sites along specific lineages. Mol Biol Evol 19:908–917

Zhang J, Nei M (2000) Positive selection in the evolution of mammalian interleukin-2 genes. Mol Biol Evol 17:1413–1416

Zhang J, Rosenberg HF, Nei M (1998) Positive Darwinian selection after gene duplication in primate ribonuclease genes. Proc Natl Acad Sci USA 95:3708–3713