Prolactin in the Afrotheria: characterization of genes encoding prolactin in elephant (*Loxodonta africana*), hyrax (*Procavia capensis*) and tenrec (*Echinops telfairi*)

Michael Wallis

Biochemistry Department, School of Life Sciences, University of Sussex, Falmer, Brighton, Sussex BN1 9QG, UK
(Correspondence should be addressed to M Wallis; Email: m.wallis@sussex.ac.uk)

Abstract

Pituitary prolactin shows an episodic pattern of molecular evolution, with occasional short bursts of rapid change imposed on a generally rather slow evolutionary rate. In mammals, episodes of rapid change occurred in the evolution of primates, cetartiodactyls, rodents and the elephant. The bursts of rapid evolution in cetartiodactyls and rodents were followed by duplications of the prolactin gene that gave rise to large families of prolactin-related proteins including placental lactogens, while in primates the burst was followed by corresponding duplications of the related GH gene. The position in elephant is less clear. Extensive data relating to the genomic sequences of elephant and two additional members of the group Afrotheria are now available, and have been used here to characterize the prolactin genes in these species and

explore whether additional prolactin-related genes are present. The results confirm the rapid evolution of elephant (*Loxodonta africana*) prolactin – the sequence of elephant prolactin is substantially different from that predicted for the ancestral placental mammal. Hyrax (*Procavia capensis*) prolactin is even more divergent but tenrec (*Echinops telfairi*) prolactin is strongly conserved. No evidence was obtained from searches of public databases for additional genes encoding prolactin-like proteins in any of these species. Detailed analysis of evolutionary rates, and other factors, indicates that the episode of rapid change in hyrax, and probably elephant, was adaptive, though the nature of the associated biological change(s) is not clear.

Journal of Endocrinology (2009) **200**, 233–240

Introduction

Prolactin is a protein hormone found in the pituitary gland of most vertebrates, which in mammals plays a central role in the regulation of mammary growth and lactation. In most mammals, prolactin is encoded by a single gene, but important exceptions occur in rodents (best characterized in rat and mouse) and ruminants; in each of these groups there is an extensive cluster of genes encoding pituitary prolactin and a number of closely related proteins including placental lactogens (Wallis 1992, Takahashi 2006, Ushizawa & Hashizume 2006, Soares *et al.* 2007). Previous studies on sequences of mammalian prolactins have shown it to be generally a strongly conserved protein, but in primates, cetartiodactyls (Cetacea plus Artiodactyla), rodents and elephant bursts of rapid evolution have led to prolactins that differ considerably from those of other mammals (Wallis 1981, 2000, 2001, Ohta 1993, Forsyth & Wallis 2002, Li *et al.* 2005a). Such a pattern of episodic evolution is also seen in various other protein hormones (Wallis 2001). The bursts of rapid change in prolactin seen in rodents and cetartiodactyls appear to have preceded the duplications of the prolactin gene that have occurred in these groups. There is no such

duplication of the prolactin gene in primates, but there has been duplication of the related GH gene (Chen *et al.* 1989, Wallis & Wallis 2002, 2006, Li *et al.* 2005b, Gonzalez-Alvarez *et al.* 2006), giving rise to placental lactogens, at least in Old World monkeys and apes, and this may relate to the changes seen in prolactin. In the case of elephant prolactin, the available data are limited to a study of the prolactin sequence at the protein level (Li *et al.* 1989); no information is available about the occurrence of placental lactogens in this species.

The recently available genome sequences for a variety of mammals allow the sequences of specific genes to be derived for various species. This approach has recently been used to derive the sequence of elephant GH, confirming that this is strongly conserved (Wallis 2008). Here, this comparative genomics approach is extended to the study of elephant prolactin and that of two other members of the mammalian group Afrotheria (to which elephant belongs) for which genomic data are available (hyrax, *Procavia capensis* and tenrec, *Echinops telfairi*). Molecular studies on mammalian phylogeny suggest that the Afrotheria diverged early during evolution of Eutheria (Murphy *et al.* 2004). The approach taken has established that: 1) these Afrotherian species do not appear to have families of prolactin-like genes – if such additional genes

do exist in any of these species they must be considerably more different from pituitary prolactin than is the case for the prolactin-like genes seen in rodents or ruminants; 2) tenrec prolactin is strongly conserved but the evolution of elephant and hyrax prolactins has been very rapid, probably driven by positive selection. The evolution of prolactin in Afrotheria thus shows a pattern of episodes of rapid change imposed on a slow basal evolutionary rate, as has previously been seen for other mammalian prolactins, and indeed many other mammalian protein hormones (Wallis 2000, 2001). However, unlike the episodes of accelerated evolution of prolactin seen in rodents, ruminants and primates, the bursts of rapid change seen in Afrotheria are not followed by duplications of the prolactin or GH genes leading to families of prolactin- or GH-like genes expressed in the placenta.

Materials and Methods

DNA sequence data

Genomic sequence data were obtained from the ensembl (release 50; <http://www.ensembl.org>; Hubbard *et al.* 2007) or ncbi (<http://www.ncbi.nlm.nih.gov/traces>) websites, separately for each of the species studied, by searching ensembl assemblies, and/or the ncbi trace archive, using the BLAST or BLAT search methods (Altschul *et al.* 1990). Table 1 summarizes the data extracted for Afrotherian species for which a substantial amount of genomic sequence is available, together with dog and opossum, used as out-groups. The prolactin gene sequences derived for the species studied here are given at the following website: http://www.lifesci.sussex.ac.uk/home/Mike_Wallis/Prolactin/Afro-PRL/.

Assessment of sequence quality

The reliability of the available sequence data was assessed in two ways, using the approach taken previously (Wallis 2008). First, examination of traces allowed direct identification of difficult regions; in a few cases, this led to changes being made

in the sequence derived from the ensembl database. In many cases, regions of sequence corresponding to the prolactin gene were covered by two or more sequence runs, often on both strands, but this was not always the case. In general, the traces indicated that most of the data are of high quality – exceptions are noted below.

The second criterion used to judge the reliability of the data was comparison of sequences obtained from the genomic projects with those available previously (usually protein or cDNA sequence). This was possible for elephant, dog and opossum prolactin sequences. Discrepancies were found in each of these cases and are discussed in detail below.

Sequence analysis

Prolactin gene and protein sequences were aligned with those available for other mammalian prolactins using the Clustalw programme (Higgins & Sharp 1988), followed by manual adjustment where necessary. A putative sequence for the prolactin of the ancestral placental mammal (a.p.m.) was derived from this (Wallis 2000). Since all orders of placental mammals arose at approximately the same point in evolutionary time, this was constructed as a consensus sequence, with each eutherian order being given equal weight, and the marsupial sequences being used as out-groups to resolve ambiguities. Nonsynonymous (d_N) and synonymous (d_S) substitution rates in coding sequences were determined using the codeml programme in the paml (phylogenetic analysis by maximum likelihood) package of Yang (2007) with a defined tree based on that in Murphy *et al.* (2004). Evolutionary rates for DNA sequences (corrected substitution rates) were determined using the baseml programme in the same package, with the HKY85 model. Phylogenetic trees were constructed using data from codeml and the programme TREEVIEW (Page 1996). 3D structures were assessed using the molecular modelling programme rasmol. Accessibility of residues within the 3D structure was determined using the programme NACCESS (Hubbard & Thornton 1993).

Table 1 Prolactin gene sequences derived from genomic data

Species	Common name	Extent ^a	Previous information	Ensembl id and notes
<i>Loxodonta africana</i>	African savannah elephant	Complete	Protein sequence; Li <i>et al.</i> (1989) ^b	Assembled from traces (only fragments found in genomic assembly)
<i>Procavia capensis</i>	Cape rock hyrax	Complete	None	Assembled from traces. Rather poor sequence at start of exon 5
<i>Echinops telfairi</i>	Small Madagascar hedgehog (tenrec)	Complete	None	Scaffold_294622: 9684-811; completed using traces
<i>Canis familiaris</i>	Dog	Complete	cDNA sequence; acc no AY741405 ^b	Chromosome 35: 23701102-23691308
<i>Monodelphis domestica</i>	Grey short-tailed opossum	Complete	cDNA sequence; acc no AF067726 ^b	Chromosome 3: 317151879-317161546

The ensembl ids refer to ensembl release 50 or traces obtained via the ncbi Trace Server. The two numbers following chromosome or scaffold id are the bases corresponding to the beginning and end of the available CDS plus introns, from the first base of the initiating ATG to the last base of the stop codon.
^aComplete means complete sequences for exons 1–5, plus at least some flanking (intron etc) sequence.
^bDiscrepancies between previously-described sequences and sequences derived from genomic data are discussed in the text.

Results and discussion

Prolactin gene sequences

The genomic data analysed are summarized in Table 1, and the prolactin gene sequences derived here are available on the website referred to above. Prolactin gene sequences were derived for two Afrotherian species for which no data were previously available: *P. capensis* (cape rock hyrax) and *E. telfairi* (tenrec; small Madagascar hedgehog). The prolactin gene sequence was also obtained for *Loxodonta africana* (African elephant), for which the protein sequence has been described previously (Li *et al.* 1989). The prolactin gene sequences of dog (*Canis familiaris*) and a marsupial, *Monodelphis domestica* (grey short-tailed opossum) were also obtained and used as out-groups. In each case, the gene sequence was assessed on the basis of the known organization of the human prolactin gene (Truong *et al.* 1984), and shown to accord with the five-exon structure of this gene. The alternative 5' untranslated exon found about 5800 nucleotides upstream of the pituitary start site in the human prolactin gene (DiMattia *et al.* 1990, Berwaer *et al.* 1994) was not detected in the species studied here.

The elephant, dog and opossum prolactin gene sequences can be checked against data previously available, providing a basis for assessing the reliability of the approach used.

The sequence of elephant prolactin derived from the genomic data differed significantly from the previously published protein sequence (Li *et al.* 1989). These differences are discussed in detail below; there seems to be little discrepancy in terms of the original data. A cDNA sequence for dog prolactin is available in the ncbi/embl/ddbj database (AY741405), but this sequence differs substantially from that derived from the genomic sequence. The prolactin gene sequence derived from the genome is of high quality, with all regions covered by multiple traces, covering both DNA strands. It has been used previously (Wallis *et al.* 2005, Alam *et al.* 2006); dog prolactin is very similar to pig prolactin. No indication was found of more than one prolactin-like gene in the dog genome, on the basis of extensive BLAT searches using dog, human and pig prolactin coding sequence. This does not completely exclude the possibility of additional prolactin-like genes, but does imply that if they are present they must differ very substantially from the identified gene. The approach taken provides no information about possible splice variants. A cDNA sequence for opossum prolactin is available in the ncbi/embl/ddbj database (AF067726). This differs from the sequence derived from the genome at 12 nt, giving seven differences at the protein level. The genome-derived sequence was of high quality, with all these sites covered by 8–12 traces, on both DNA strands. No indication was found of additional prolactin-like genes in the opossum genome. The derived coding sequence was similar to that reported for prolactin from two other marsupials – brushtail possum (*Trichosurus vulpecula*) and brown bandicoot (*Isodon macrourus*; Curlew *et al.* 1998, Veitch *et al.* 2006).

For each of the Afrotherian prolactin gene sequences, analysis of traces indicated some potential polymorphism, which is indicated on the sequences on the website. For elephant and tenrec, this would not affect the derived amino acid sequence, but for hyrax two potential polymorphisms could alter the protein sequence. Although polymorphism seems the most likely explanation for the variation observed, occurrence of more than one very similar gene following gene duplication cannot be ruled out. But for none of these species was there any evidence for additional genes encoding distinctly different prolactin-like proteins.

Novel Afrotherian prolactin sequences

The derived amino acid sequences of elephant, hyrax and tenrec prolactins are included in the alignment given in Fig. 1.

As indicated above, the amino acid sequence derived for elephant prolactin was substantially different (at 15 residues) from that described by Li *et al.* (1989). The latter was produced by determining amino acid compositions and partial sequences of tryptic peptides and aligning these using ovine prolactin as a template. Detailed comparison shows that the expected amino acid compositions for tryptic peptides based on the genome-derived sequence are identical to those given by Li *et al.* (1989) in all but one case. The discrepancies are due to the assumption that the ovine sequence provided a satisfactory basis for alignment and sequence deduction, which for these peptides (in retrospect) was not the case, probably because of the substantial divergence between the elephant and ovine sequences. A similar approach was used for elephant GH (Hulmes *et al.* 1989), and gave a sequence identical to that subsequently deduced from the genomic sequence (Wallis 2008), but here the difference from the template was much less.

There is thus little disagreement between the data provided by Li *et al.* (1989) and the elephant prolactin sequence derived from the genome. Detailed examination of the sequence traces indicates that the data for the elephant prolactin coding sequence (CDS) are of high quality, with several sequence runs covering each exon, on both strands, with no discrepancies. The genomic sequence also provides the signal peptide sequence. Despite the differences between the sequence described previously and that obtained from the genome, the conclusion that the elephant prolactin sequence diverges considerably from that derived for the ancestral placental mammal (a.p.m.) is unchanged – elephant and a.p.m. sequences differ at 49 residues (Fig. 1).

The sequence derived for hyrax prolactin is even more different from that of the a.p.m. than is that of elephant – differing at 94 residues, 47% of all sites. On the other hand, the sequence derived for tenrec prolactin is conserved, differing from that of the a.p.m. at only eight sites (Fig. 1). Thus, the very variable evolutionary rates previously identified for prolactin across the mammals (Wallis 1981, 2000) also apply within the Afrotheria.

ratio for each branch). The values obtained were used to construct the phylogenetic trees shown in Fig. 2. As expected, rates of evolution based on d_S are relatively uniform, but those based on d_N values are very variable, reflecting particularly the episodes of rapid change identified for protein sequences. The episodes of rapid change on the lineages leading to elephant and hyrax contrast with the slow evolution seen for tenrec. d_N/d_S ratios for branches leading to elephant (0.99) and hyrax (1.13) are elevated substantially compared with that on the branch leading to tenrec (0.046). However, use of model 2 in codeml (2 d_N/d_S ratios for branches) showed that the elevated ratio for hyrax was not significantly greater than 1.0.

The more powerful branch-sites method within codeml allows d_N/d_S ratio determination to be made with allowance for rate variation between sites (Yang & Nielsen 2002). When this was applied, using codeml model A, the d_N/d_S ratio for the branch leading to hyrax was 2.50, significantly greater than 1.0 ($P < 0.01$; likelihood ratio test). This provides strong evidence for adaptive evolution as the basis for the episode of rapid change on the branch leading to hyrax.

Varying evolutionary rates within Afrotherian prolactin genes

A disadvantage of the use of d_N/d_S ratios is that d_S may be correlated with d_N (Mouchiroud *et al.* 1995). In this case, the elevation of d_N may be accompanied by elevation of d_S , so that d_S is not really a suitable reference. An alternative approach is

to compare the rate of evolution of exons and introns within a gene. Normally, exons (constrained by purifying selection) evolve more slowly than introns (little if at all constrained); if the exon rate increases to a level above that of adjacent introns, then adaptive change is indicated. Accordingly, rates of evolution for exons and ~300 bases of adjacent introns (a limitation necessary because not all intron sequences were complete) were determined using the programme baseml (Yang 1994) and alignments of human, dog, tenrec, elephant and hyrax DNA (opossum could not be used as an out-group because of problems aligning introns). Results are shown in Fig. 3. The pattern for dog and tenrec prolactin genes is characteristic of fairly slowly evolving genes (Graur & Li 2000), with in every case the evolutionary rate for an exon being much less than that of the adjacent intron. For hyrax the situation is reversed; the mean rate for exons is $2.3\times$ that for introns ($P = 0.002$, Student's *t*-test), again providing strong support for adaptive evolution. For elephant the mean rate for exons is $1.49\times$ greater than that for introns, but the difference is not significant.

Distribution of substitutions in the 3D structure

The 3D structure of human prolactin has been determined using nmr spectroscopy (Keeler *et al.* 2003, Teilum *et al.* 2005). The distribution of the sites that change along the branches leading to hyrax, elephant and tenrec was explored

Figure 2 Phylograms illustrating the evolution of mammalian prolactin coding sequences, based on (a) synonymous sites and (b) nonsynonymous sites which were constructed using the codeml and TREEVIEW programmes, with a defined tree based on Murphy *et al.* (2004). In (b) thick branches indicate those which show accelerated evolution – with $d_N/d_S > 0.5$, except for some branches with very small numbers (< 5) of substitutions. The scale bars indicate 0.1 substitutions/nucleotide site. The alignment of coding sequences is given on the website referred to in the Methods section.

Figure 3 Evolutionary rates of exons and introns from prolactin sequences of dog, tenrec, elephant and hyrax. Rates were determined using the baseml programme, and are expressed as substitutions per nucleotide site. Grey columns, total CDS; open columns, introns; black columns, exons. Exon 1 is not included because its CDS is too short to be useful.

(Table 2). There is a tendency for residues in the central core (identified as residues with solvent accessibility less than 10%) and receptor binding sites (as given by Teilum *et al.* 2005) to be relatively conserved, but only in the case of elephant is this significant ($P=0.015$, Fisher’s exact test, Zhang *et al.* 1998, with application of Bonferroni correction). Overall, there is no strikingly non-random distribution of substitutions of the sort observed for GH (Wallis 2008).

Mechanisms underlying episodes of rapid evolution – absence of placental lactogens in Afrotherians

Previous studies have established that prolactin, like several other mammalian polypeptide hormones, shows an episodic pattern of evolution. This work has established that this is true also of three orders of Afrotherian mammals. The rapid evolution on the lineage leading to elephant has been confirmed, PRL of the tenrec is strongly conserved, but PRL of the hyrax has evolved at a remarkably accelerated rate.

Table 2 Distribution of substitutions within the 3D structure of prolactin

	Overall	Core	Binding site	Remainder ^a
Elephant	49/199 (24.6%) ^b	5/51 (9.8%)	3/21 (14.3%)	41/127 (32.3%)
Hyrax	94/199 (47.2%)	17/51 (33.3%)	10/21 (47.6%)	67/127 (52.8%)
Tenrec	8/199 (4.0%)	2/51 (3.9%)	0/21 (0%)	6/127 (4.7%)

^aAll residues not identified as core or binding sites.
^bNumber of substitutions/residues in that category (percentage).

In the case of the hyrax, study of d_N/d_S ratios, and the observation that exons are evolving faster than introns, strongly indicate that the accelerated evolution was driven by adaptive selection. In the case of the elephant the case is less clear cut, but the fact that elephant prolactin is produced in high yield in the pituitary gland, and retains biological activity (Li *et al.* 1987), militates against loss of function as the explanation of accelerated evolution.

If the bursts of rapid evolution seen in these hormones are driven by adaptive selection, what is the nature of the functional changes involved? Too little is known about the biological properties and functions of prolactin in Afrotheria to allow firm conclusions to be drawn, but it could result from the process of function switching (Wallis 1997, Forsyth & Wallis 2002). If prolactin acquired a second function, the importance of which fluctuated over time, each switch (fluctuation) would lead to adaptation and additional substitutions. Repeated fluctuations could lead to substantial sequence change with relatively little change in function. Another explanation for rapid change could be maternal-foetal competition (Haig 2008), although for this to be fully effective gene imprinting may be necessary, about which no information is available for Afrotherian prolactins.

In other groups where prolactin evolution shows marked acceleration, gene duplication is also found. In ruminants and rodents, multiple prolactin gene duplications have given rise to extensive families of prolactin-related genes, expressed mainly in the placenta, and with diverse functions (Wallis 1992, Takahashi 2006, Ushizawa & Hashizume 2006, Soares *et al.* 2007). In primates, the prolactin gene is not duplicated, but there is duplication of the GH gene, and GH in at least some primates possesses lactogenic activity. In the case of Afrotheria careful searching of the genomic data bases and Trace Archives has revealed no evidence for additional genes encoding prolactin-like proteins. Similarly, no evidence for duplication of the GH gene in elephant or hyrax was obtained (Wallis 2008). This may be because such genes are too diverged from prolactin (or GH) itself to allow detection, but if this is the case the divergence must be very substantial. Searching bovine, rat and mouse Trace Archives using homologous or heterologous prolactin CDS sequences revealed multiple prolactin-like genes, clearly reflecting the

additional prolactin-like genes in these species, so if additional prolactin-like genes do exist in tenrec, elephant or hyrax they must be more divergent than is the case for the ruminant or rodent prolactin gene families. It thus seems likely that the rapid evolution of prolactin in lineages leading to elephant and hyrax is not associated with duplication of prolactin or GH genes in these species. It should be noted, however, that the duplications that gave rise to families of prolactin-like genes in ruminants and rodents appear to have occurred after the episodes of rapid evolution in these groups. In Afrotheria, therefore, the mechanisms leading to rapid evolution of prolactin may have been similar to those occurring in ruminants and rodents, but without subsequent duplications of the prolactin gene, at least in the species considered here.

Declaration of interest

There is no conflict of interest that could be perceived as prejudicing the impartiality of the research reported.

Funding

This research did not receive any specific grant from any funding agency in the public, commercial or not-for-profit sector.

References

- Alam SMK, Ain R, Konno T, Ho-Chen JK & Soares MJ 2006 The rat prolactin gene family locus: species-specific gene family expansion. *Mammalian Genome* **17** 858–877.
- Altschul SF, Gish W, Miller W, Myers EW & Lipman DJ 1990 Basic local alignment search tool. *Journal of Molecular Biology* **215** 403–410.
- Berwaer M, Martial JA & Davis JRE 1994 Characterization of an up-stream promoter directing extrapituitary expression of the human prolactin gene. *Molecular Endocrinology* **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.
- Curlewis JD, Saunders MC, Kuang J, Harrison GA & Cooper DW 1998 Cloning and sequence analysis of a pituitary prolactin cDNA from the brushtail possum (*Trichosurus vulpecula*). *General and Comparative Endocrinology* **111** 61–67.
- 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. *Journal of Biological Chemistry* **265** 16412–16421.
- Forsyth IA & Wallis M 2002 Growth hormone and prolactin – molecular and functional evolution. *Journal of Mammary Gland Biology and Neoplasia* **7** 291–312.
- González Alvarez R, Revol de Mendoza A, Esquivel Escobedo D, Corrales Félix G, Rodríguez Sánchez I, González V, Dávila G, Cao Q, de Jong P, Fu Y-X *et al.* 2006 Growth hormone locus expands and diverges after the separation of new and old world monkeys. *Gene* **380** 38–45.
- Graur D & Li W-H 2000 *Fundamentals of Molecular Evolution*. edn 2, Sunderland, MA: Sinauer.
- Haig D 2008 Placental growth hormone-related proteins and prolactin-related proteins. *Placenta* **22** S36–S41.
- Higgins DG & Sharp PM 1988 CLUSTAL: a package for performing multiple sequence alignment on a microcomputer. *Gene* **73** 237–244.
- Hubbard SJ & Thornton JM 1993 'NACCESS', Computer Program, Department of Biochemistry and Molecular Biology, University College London.
- Hubbard TJP, Aken BL, Beal K, Ballester B, Caccamo M, Chen Y, Clarke L, Coates G, Cunningham F & Cutts T 2007 Ensembl 2007. *Nucleic Acids Research* **35** D610–D617.
- Hulmes JD, Miedel MC, Li CH & Pan YCE 1989 Primary structure of elephant growth hormone. *International Journal of Peptide and Protein Research* **33** 368–372.
- Keeler C, Dannies PS & Hodsdon ME 2003 The tertiary structure and backbone dynamics of human prolactin. *Journal of Molecular Biology* **328** 1105–1121.
- Li CH, Chung D, Bewley TA & Cabrera CM 1987 Elephant prolactin: isolation and characterization. *International Journal of Peptide and Protein Research* **29** 472–477.
- Li CH, Oosthuizen MMJ & Chung D 1989 Primary structure of elephant pituitary prolactin. *International Journal of Peptide and Protein Research* **33** 67–69.
- Li Y, Wallis M & Zhang Y-P 2005a Episodic evolution of prolactin receptor gene in mammals: coevolution with its ligand. *Journal of Molecular Endocrinology* **35** 411–419.
- Li Y, Ye C, Shi P, Zou X-J, Xiao R, Gong Y-Y & Zhang Y-P 2005b Independent origin of the growth hormone gene family in New World monkeys and Old World monkeys and hominoids. *Journal of Molecular Endocrinology* **35** 399–409.
- Mouchiroud D, Gautier C & Bernardi G 1995 Frequencies of synonymous substitutions in mammals are gene-specific and correlated with frequencies of nonsynonymous substitutions. *Journal of Molecular Evolution* **40** 107–113.
- Murphy WJ, Pevzner PA & O'Brien SJ 2004 Mammalian phylogenomics comes of age. *Trends in Genetics* **20** 631–639.
- Ohta T 1993 Pattern of nucleotide substitutions in growth hormone-prolactin gene family: a paradigm for evolution by gene duplication. *Genetics* **134** 1271–1276.
- Page RDM 1996 TREEVIEW: an application to display phylogenetic trees on personal computers. *Computer Applications in the Biosciences* **12** 357–358.
- Soares MJ, Konno T & Alam SMK 2007 The prolactin family: effectors of pregnancy-dependent adaptations. *Trends in Endocrinology and Metabolism* **18** 114–121.
- Takahashi T 2006 Biology of the prolactin family in bovine placenta. I. Bovine placental lactogen: expression, structure and proposed roles. *Animal Science Journal* **77** 10–17.
- Teilum K, Hoch JC, Goffin V, Kinet S, Martial JA & Kragelund BB 2005 Solution structure of human prolactin. *Journal of Molecular Biology* **351** 810–823.
- Truong AT, Duez C, Belayew A, Renard A, Pictet R, Bell GI & Martial JA 1984 Isolation and characterization of the human prolactin gene. *EMBO Journal* **3** 429–437.
- Ushizawa K & Hashizume K 2006 Biology of the prolactin family in bovine placenta. II. Bovine prolactin-related proteins: their expression, structure and proposed roles. *Animal Science Journal* **77** 18–27.
- Veitch C, Gemmell RT & Curlewis JD 2006 Cloning and sequence analysis of a pituitary prolactin cDNA from the northern brown bandicoot (*Isodon macrourus*). *General and Comparative Endocrinology* **146** 304–309.
- Wallis M 1981 The molecular evolution of pituitary growth hormone, prolactin and placental lactogen: a protein family showing variable rates of evolution. *Journal of Molecular Evolution* **17** 10–18.
- Wallis M 1992 The expanding growth hormone/prolactin family. *Journal of Molecular Endocrinology* **9** 185–188.
- Wallis M 1997 Function switching as a basis for bursts of rapid change during the evolution of pituitary growth hormone. *Journal of Molecular Evolution* **44** 348–350.
- Wallis M 2000 Episodic evolution of protein hormones: molecular evolution of pituitary prolactin. *Journal of Molecular Evolution* **50** 465–473.
- Wallis M 2001 Episodic evolution of protein hormones in mammals. *Journal of Molecular Evolution* **53** 10–18.
- Wallis M 2008 Mammalian genome projects reveal new growth hormone (GH) sequences. Characterization of the GH-encoding genes of armadillo (*Dasypus novemcinctus*), hedgehog (*Eriacus europaeus*), bat (*Myotis lucifugus*), hyrax (*Procavia capensis*), shrew (*Sorex araneus*), ground squirrel (*Spermophilus tridecemlineatus*), elephant (*Loxodonta africana*), cat (*Felis catus*) and opossum (*Monodelphis domestica*). *General and Comparative Endocrinology* **155** 271–279.

- Wallis OC & Wallis M 2002 Characterisation of the GH gene cluster in a new-world monkey, the marmoset (*Callithrix jacchus*). *Journal of Molecular Endocrinology* **29** 89–97.
- Wallis OC & Wallis M 2006 Evolution of growth hormone in primates: the GH gene clusters of the New World monkeys marmoset (*Callithrix jacchus*) and white-fronted capuchin (*Cebus albifrons*). *Journal of Molecular Evolution* **63** 591–601.
- Wallis OC, Mac-Kwashie AO, Makri G & Wallis M 2005 Molecular evolution of prolactin in primates. *Journal of Molecular Evolution* **60** 606–614.
- Yang Z 1994 Estimating the pattern of nucleotide substitution. *Journal of Molecular Evolution* **39** 105–111.
- Yang Z 2007 PAML4: phylogenetic analysis by maximum likelihood. *Molecular Biology and Evolution* **24** 1586–1591.
- Yang Z & Bielawski JP 2000 Statistical methods for detecting molecular adaptation. *Trends in Ecology and Evolution* **15** 496–503.
- Yang Z & Nielsen R 2002 Codon-substitution models for detecting molecular adaptation at individual sites along specific lineages. *Molecular Biology and Evolution* **19** 908–917.
- Zhang J, Rosenberg HF & Nei M 1998 Positive Darwinian selection after gene duplication in primate ribonuclease genes. *PNAS* **95** 3708–3713.

Received in final form 12 November 2008

Accepted 14 November 2008

Made available online as an Accepted Preprint

17 November 2008