Do some viruses use growth hormone, prolactin and their receptors to facilitate entry into cells?

Article (Accepted Version)

Wallis, Michael (2021) Do some viruses use growth hormone, prolactin and their receptors to facilitate entry into cells? BioEssays, 43 (4). a2000268. ISSN 0265-9247

This version is available from Sussex Research Online: <http://sro.sussex.ac.uk/id/eprint/98025/>

This document is made available in accordance with publisher policies and may differ from the published version or from the version of record. If you wish to cite this item you are advised to consult the publisher's version. Please see the URL above for details on accessing the published version.

Copyright and reuse:

Sussex Research Online is a digital repository of the research output of the University.

Copyright and all moral rights to the version of the paper presented here belong to the individual author(s) and/or other copyright owners. To the extent reasonable and practicable, the material made available in SRO has been checked for eligibility before being made available.

Copies of full text items generally can be reproduced, displayed or performed and given to third parties in any format or medium for personal research or study, educational, or not-for-profit purposes without prior permission or charge, provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.

Do some viruses use growth hormone, prolactin and their receptors to facilitate entry into cells?

Episodic evolution of hormones and receptors suggests host-virus arms races; related placental lactogens may provide protective viral decoys.

Michael Wallis

Department of Biochemistry and Biomedicine

School of Life Sciences

University of Sussex

Brighton BN1 9QG

UK

m.wallis@sussex.ac.uk

Keywords: growth hormone, prolactin, placental lactogen, virus, arms race, episodic evolution, viral decoy

Abbreviations

dN non-synonymous substitution rate

dS synonymous substitution rate

ecd extracellular domain

EGFR epidermal growth factor receptor

GH growth hormone

GH-N pituitary (normal) growth hormone

GH-V placental (variant) growth hormone

GHR GH receptor

hPL human placental lactogen

IGF-I insulin-like growth factor I

IGF1R IGF-I receptor

LH luteinizing hormone

NWM New World monkey

OWM Old World monkey

PL placental lactogen

PRLR prolactin receptor

VILP viral insulin-like peptide

RSV respiratory syncytial virus

TfR transferrin receptor

Summary

The molecular evolution of pituitary growth hormone and prolactin in mammals shows two unusual features: episodes of markedly accelerated evolution and, in some species, complex families of related proteins expressed in placenta and resulting from multiple gene duplications. Explanations of these phenomena in terms of physiological adaptations seem unconvincing. Here I propose an alternative explanation, namely that these evolutionary features reflect the use of the hormones (and their receptors) as viral receptors. Episodes of rapid evolution can then be explained as due to “arms races” in which changes in the hormone lead to reduced interaction with the virus, and subsequent changes in the virus counteract this. Placental paralogues of the hormones could provide decoys that bind viruses, and protect the foetus against infection. The hypothesis implies that the extensive changes introduced into growth hormone, prolactin and their receptors during the course of mammalian evolution reflect viral interactions, not endocrine adaptations.

1. Introduction

Entry of a virus into a target cell requires attachment to one or more receptors on the cell surface. A range of host molecules are utilized as such receptors, including sialylated glycans and integral membrane proteins, including receptors for growth factors and hormones [1-4]. Viral interaction with a host protein can potentially lead to a host-virus "arms race" in which recurring rounds of mutation in host and virus, functioning respectively to restrict and enhance interaction, lead to very rapid evolution of both receptor and virus [5-8]. Here I propose that episodes of rapid evolution that are seen in the molecular evolution of mammalian growth hormone (GH) and prolactin and their receptors are a consequence of such arms races, in which interaction between virus (or possibly other pathogen) and hormone often plays a central role. I also propose that the GH- and prolactin-like proteins, placental lactogens (PLs), that are expressed at high levels in the placenta of some mammals, including humans, act as viral decoys, providing foetal protection.

GH and prolactin are related protein hormones produced in the pituitary gland, which regulate growth, lactation and metabolism via interactions with receptors in a wide range of tissues. They arose as a consequence of a gene duplication early in vertebrate evolution, followed by divergent evolution, and acquisition of specific, distinct functions, although they also show some overlap in biological actions [9,10]. They bind to structurally distinct but related type I cytokine receptors, which are integral membrane glycoproteins normally forming homodimers. Each monomer comprises an extracellular domain (ecd), largely consisting of two fibronectin III-like domains, a single helical transmembrane domain, and an intracellular domain that may be largely

unstructured. The ligand binds to each subunit via two binding sites that are quite distinct on the hormone, but similar on the receptor subunits (**Figure 1**), leading to activation of protein tyrosine kinase JAK2, by a mechanism particularly well-defined in the case of GH [11-13].

In the following I consider the main observations underlying the hypothesis proposed here: the nature and basis of the episodic molecular evolution of GH and prolactin, the families of GH- and prolactin-related genes that are expressed in the placenta in some mammals, the nature of host-virus arms races seen for other genes, and the role of hormones and growth factors in infectivity and growth of many viruses.

2. GH, prolactin and their receptors show an unusual pattern of molecular evolution

2.1 Evolution of the hormones is markedly episodic

The molecular evolution of GH and prolactin in mammals shows unusual features. In both a slow basal evolutionary rate (near stasis) is interrupted by occasional episodes of rapid change [9,14-19]. In several cases the end of the episode of rapid evolution apparently coincided with additional gene duplication(s), leading to clusters of closely related genes expressed mainly in the placenta [9, 17].

Figure 2a illustrates evolutionary trees for mature (lacking signal peptide) mammalian GH based on non-synonymous (dN) and synonymous substitutions (dS) (**Box 1**). For most of the dN tree (equivalent to a tree based on protein sequence [20]) the sequence is strongly conserved, reflecting a low basal rate of evolution, seen, for example, in

pig, rabbit and all Carnivora. However, the rate of evolution increases markedly on several occasions. Thus, on the lineage to Simiiformes (monkeys and apes) there is a burst of change with incorporation of ~80 substitutions (increasing the evolutionary rate at least 10-20-fold), after divergence from prosimians. This rapid evolution ceased before the divergence of New World monkeys (NWM) and Old World monkeys (OWM)/apes [9,14,15]. A similar episode of rapid GH evolution occurred in cetartiodactyls, on the lineage leading to ruminants, and in bats, Xenarthra and some Afrotheria [18]. The episodic evolution is not seen in the dS tree, indicating that it reflects only those substitutions resulting in change in protein sequence.

Figure 3a shows equivalent evolutionary trees for mammalian prolactin. The pattern is similar, but not identical, to that for GH. Here too there are bursts of rapid evolution in primates and cetartiodactyls, at about the same time periods as for GH [9,17]. Prolactin evolution in some bats is high with, in vespertilionid bats, an insertion of variable length not seen in any other vertebrate [21]. However, unlike GH, rapid evolution of prolactin is also seen on the lineages leading to elephants [22] and some rodents.

Clearly GH and prolactin in mammals do not follow the general rule that the rate of evolution for any given protein is rather constant, though between proteins rates vary extensively [23].

2.2. The receptors also show episodic evolution

Like the hormones, the receptors for GH and prolactin (GHR, PRLR) show episodic patterns of evolution, which are seen most clearly in the ecds [24] (**Figures 2b and 3b**). Thus, for primates and cetartiodactyls both GHR and PRLR show accelerated evolution during the same evolutionary period as noted for GH and prolactin. On the other hand, for armadillo GHR there is no burst of rapid change corresponding to that seen for GH, and for elephant PRLR there is no rapid evolution equivalent to that seen for prolactin. Notably in these last two cases the substitutions seen in the hormone ligands are mostly located away from the receptor binding sites [18,22; see **Section 2.4**].

2.3. What are the causes of the episodic evolution?

The remarkable episodic evolution shown by GH and prolactin has two potential explanations. Accelerated evolution may reflect loss of function, leading to relaxation of purifying selection and hence rapid change (see **Box 1**). Alternatively it may reflect positive selection (adaptive evolution) associated with changed function. However, there is no clear evidence for substantial changes in physiological function associated with the large sequence changes seen, for example, in human or bovine GH or prolactin. The growth-promoting activity of human or bovine GH in the rat is similar to that of pig or rat GH, and there is no major change in other biological activities; the same is true for prolactin. Major loss of activity during the episodes of rapid evolution seen in primate and cetartiodactyl GHs and prolactins is also unlikely because the rate of evolution eventually falls back towards basal (**Figures 1 and 2**); relaxed purifying selection with incorporation of many 'random' substitutions, would

be expected to give a much-changed protein, unlikely to then regain functionality and a structure subsequently strongly maintained by purifying selection.

The case for adaptive evolution can be supported by a dN/dS ratio greater than 1.0 (**Box 1**), and this is the case for many of the branches showing accelerated evolution (**Figures 1 and 2**). Likewise, a rate of change of coding sequence greater than that of adjacent introns, supports adaptive evolution in some cases [22]. But adaptive evolution implies changes in biological activity, and few such changes are associated with the episodes of rapid change. A significant change has occurred in species specificity, associated with the human GH/GHR pair [25,26]; non-primate GHs are inactive in humans, even though, in the rat, human GH has similar activity to other GHs. This is of experimental and medical interest, but probably not of significance for GH physiology. Human GH also shows greater lactogenic activity than GHs of some mammals, but the physiological significance of this is unclear given that humans have a fully functioning lactogenic hormone (prolactin). A mechanism (function switching) by which relatively small functional changes in GH could give rise to substantial change in sequence has been proposed [27]; this might apply in some cases but seems unlikely to explain the many instances of accelerated evolution that have now been identified in GH and prolactin and their receptors.

The alternative explanation proposed here is that the bursts of rapid evolution reflect attempts by the host to limit infection by a virus using GH, prolactin, GHR and/or PRLR as a receptor to gain entry to target cells. This too would involve positive selection. Changes in the host would lead to counter changes in the virus, leading to a host-virus arms race, with a sustained period of rapid evolution, but little change in

function. The end of the accelerated evolution/arms race seen in primates and cetartiodactyls could reflect the development of an alternative defence mechanism by the host, using placentally-expressed genes derived from duplications of *GH* or prolactin genes, as discussed further below.

2.4. How are substitutions distributed in the 3D structure?

GH and prolactin have 3D conformations comprising a 4-helix bundle with up-up-down-down topology and N- and C-terminal residues at the same (proximal) end of the bundle [28,29]. The distribution in the structure of substitutions introduced during the episodes of rapid change varies depending on lineage. Thus for the rapid change of both GH and prolactin on the lineage leading to Simiiformes the substitutions are widely distributed over the protein, including the receptor binding sites (**Figure 4**). Here many substitutions are also seen in the receptors (**Figures 2 and 3**). For the rapid evolution leading to armadillo GH, most changes are on the top surface of the protein [18] (**Figure 4**), away from the binding sites; on the other hand, for the lineage leading to elephant prolactin they are mostly at the proximal end of the molecule [22]. In these last two cases few changes are seen in the receptors. In vespertilionid bats, prolactin evolution is rapid and there is an insertion of 18-60 residues; this insertion, and most substitutions, are located at the distal end of the protein [21].

3. Duplications of GH and prolactin genes give rise to complex gene clusters

3.1. Many such duplicate genes are expressed in the placenta

A second remarkable feature of GH and prolactin evolution in mammals is the occurrence, on several lineages, of gene duplications giving rise to complex families of proteins expressed largely in the placenta. Thus, in primates, duplication of the *GH* gene gave rise to a cluster of genes encoding placental lactogens and other placental proteins: 5 such genes in human (including pituitary GH itself) [30], 4-9 in other apes and OWM [31-33], 8 in marmoset and over 20 in some other NWM [34] (**Figure 5**). Independent duplications appear to have given rise to the placentally expressed genes in NWM and OWM/apes [34-37]. A more-recent duplication of the *GH* gene is seen in sheep and goats [38,39].

Duplications of the prolactin gene occurred in cetartiodactyls and rodents, here giving rise to large and complex families of genes expressed largely in the placenta [40-43]. In the mouse genome, these occur as a cluster of about 26 prolactin-related genes. The gene duplications in rodents and cetartiodactyls occurred independently.

Notably, all of these duplications, of *GH* and prolactin genes, occurred after episodes of rapid evolution of the hormones. In primates and cetartiodactyls they appear to have occurred at about the time that the accelerated evolution stopped [9]. Equivalent gene duplications do not appear to have occurred in other cases of accelerated GH and prolactin evolution referred to above (bats, Afrotheria including elephant, Xenarthra including armadillo). In some of these cases the rapid evolution may not have ceased, though lack of sequences from extant species limits analysis. There is no clear evidence for duplication of the corresponding receptor genes, *GHR* or *PRLR*, in any mammalian group.

3.2. What are the functions of placental lactogens and consequences of loss?

The physiological functions of the GH- and prolactin-like placental hormones are unclear. The human gene cluster is relatively simple and well-studied [30,44], comprising five genes (**Figure 5**), *GH-N* (*GHI*, expressed in the pituitary), *PL-A* and *PL-B* (*CSH1*, *CSH2*, encoding identical placental lactogens, hPL), *GH-V* (*GH2*, encoding placental GH, GH-V) and *PL-L* (*CSHL1*, probably a pseudogene). hPL has 85% sequence identity with human GH, low growth-promoting activity but high activity in some lactogenic assays. It reaches very high plasma concentration (5-7µg/ml) in the second half of pregnancy, at least 100-fold greater than mean GH level in non-pregnant human [44]. Two genes encoding the same protein contribute to this high concentration. GH-V has 93% sequence identity to GH-N; it has similar growth-promoting activity to GH, but no lactogenic activity. It replaces GH-N during pregnancy, circulating at relatively low levels (~25ng/ml) [45].

The role of hPL in human pregnancy is unclear. Its main biological activity is lactogenic, but it disappears from the circulation at parturition, as lactation is initiated. A role in mammary gland development during pregnancy is possible, but prolactin plays this role in other mammals, and is present in humans (plasma level 150-180 ng/ml at term [46]). hPL may have other actions during pregnancy, including diabetogenic, but their physiological significance is not clear. The biggest question mark over the role of hPL comes from rare cases in which the *hPL* genes are deleted.

About 10 cases have been reported in which hPL concentration during pregnancy was very low or undetected [46], including some cases with a deletion of much of the *GH*

locus, including all genes except *GH-N* [47]. In several cases, including some where *PL* and *GH-V* genes were deleted, gestation was normal, and babies had normal birth weight, despite absence of circulating hPL. In others birth weight was rather low, while in one the baby showed severe growth retardation [48]. In the absence of *GH-V* and hPL, the normal suppression of *GH-N* during pregnancy was not seen. These observations suggest that hPL lacks an important physiological role in pregnancy, but are consistent with an antiviral function as decoy or blocker of the PRLR. Normal pregnancy would reflect the case in which absence of viral infection would not require the functions of PL, but severe growth retardation could result from absence of such functions if viral infection occurred. Low birth weight is known to be a consequence of some viral infections, including rubella and cytomegalovirus [49].

The clusters of prolactin-like genes in mouse and rat are more complex, including about 25 genes, with the gene for pituitary prolactin at the 5' end of the cluster [40-42]. The other genes are expressed mainly in the placenta, mostly in foetal (trophoblast) tissue, sometimes in maternal (decidual) tissue. 3-5 of the genes encode proteins (placental lactogens) that bind to PRLR. These are found in plasma at very high levels [50] and could act as viral decoys. The other genes are expressed at lower levels, mostly do not bind to PRLR, and may play a role in defence against environmental stresses during pregnancy [40]. Ruminants also have a complex cluster of prolactin-like genes (at least 12 genes in cattle), including genes for prolactin, placental lactogen, and many prolactin-related proteins, mostly lacking lactogenic activity [43]. All but prolactin are expressed primarily in the placenta. Non-ruminant cetartiodactyls lack such a gene cluster, having no paralogues of the prolactin gene.

4. A host-virus arm race can lead to rapid evolutionary change

A host-virus arms race can occur when a viral protein utilises a host protein in the course of infection or replication [5-8,51,52]. Substitutions occur in the gene encoding the host protein, decreasing the interaction, and subsequent substitutions occur in the virus that restore it. This can repeat, potentially without end, leading to rapid evolutionary change with little apparent significant functional change in either protein (Red Queen dynamics [53]). The substitutions in the host protein are limited by the need to maintain its normal function, though this may be compromised by the urgency of rejecting the virus. Such limitations may also apply to the viral protein, though here the primary need will be to maintain interaction with the receptor, since failure to do so will lead to inability to infect the host and potential extinction. An arms race can also arise (perhaps more frequently) from interaction between a defensive host protein and a viral protein; in this case the initial mutational change will be in the virus.

Host-virus arms races have been widely studied in the context of interactions with components of the immune system [51,54]. However, interactions with receptors are also clear candidates [4,5,8,55,56]. The transferrin receptor (TfR) provides an interesting example [55]. This protein serves as a receptor for several different viruses, including New World arenaviruses that infect several rodent species. Evolutionary analysis of TfR in these species showed that it has undergone rapid adaptive evolution at sites within the interface that interacts with viruses. Corresponding changes in the viruses were also identified. The substitutions had little effect on the physiological activity of TfR.

The episodes of rapid change seen in the evolution of GH, prolactin and their receptors resemble the changes seen in the host protein in such an arms race, in which many substitutions are introduced, but little change in physiological action occurs. If this is actually the explanation, corresponding changes in one or more viruses would be expected.

5. Viruses can exploit hormones, growth factors and their receptors

The potential role played by growth factors and hormones and their receptors in viral infection and growth is becoming increasingly recognized.

Receptors for hormones, growth factors or cytokines are clearly potential targets as viral receptors in that they are already adapted to bind to specific circulating ligands, and subsequently promote downstream processes that facilitate viral entry and survival, such as receptor-mediated endocytosis [57,58]. Receptors known to be used by viruses include epidermal growth factor receptor (EGFR), platelet-derived growth factor receptor (PDGFR) and insulin-like growth factor 1 receptor (IGF1R). EGFR is used by many viruses [59], including influenza A virus, cellular uptake of which has been shown to depend on binding to this receptor, followed by downstream signalling and endocytosis [60]. Human cytomegalovirus utilizes several different receptors, depending on virus strain and host target cell, including PDGF and possibly EGFR [61]. Recently IGF1R has been shown to act as a receptor for respiratory syncytial virus (RSV) [62]. RSV glycoprotein F binds to the IGF1R, triggering activation of

protein kinase C zeta, which recruits nucleolin to the cell surface where it probably facilitates entry of RSV into the cell.

Recent studies also showed that some viruses produce proteins structurally similar to polypeptide hormones [63,64]. Viral insulin-like peptides (VILPs) from members of the family *Iridoviridae* were well characterised; they showed clear sequence similarity to human insulin and IGF-I, including retention of all 6 cysteine residues, suggesting similar 3D structures. These VILPs showed insulin-like activity *in vitro* and *in vivo*, and could potentially alter cellular metabolism in ways favouring viral survival and growth. It has also been suggested that such molecules could be utilized by the virus as ligands for hormone receptors on host cells [64,65], a mechanism that would resemble that proposed here for GH and prolactin. Thus, in some cases (e.g. armadillo GH) it is suggested here that virus binds directly to the hormone, which in turn binds to its receptor and facilitates attachment and entry of the associated virus. In other cases binding of virus to both hormone and receptor would be involved.

6. GH, prolactin, GHR, PRLR as potential viral receptors; receptor-mediated endocytosis

In order to infect a host cell a virus has first to bind to one or more receptors on the host cell surface [1-4]. Receptors for GH and prolactin clearly have potential as viral receptors. They are widely distributed in many tissues and are subject to ligand-promoted receptor endocytosis [66,67]. However, activation of the receptors for GH and prolactin is a complex process, requiring binding to both subunits in the homodimeric receptor, followed by subtle conformational changes [11,28]. This may

not be easily achieved by a virus opportunistically associating with the receptor directly. Interaction of the virus with the ligand, GH or prolactin, which subsequently interacts with the receptor (with or without additional virus-receptor interaction) could provide a mechanism for associating virus with receptor, followed by activation and endocytosis. Such a mechanism is speculative, and has not been previously described for other hormones or growth factors.

The pattern of substitutions seen in GH and prolactin may provide a guide to the nature of any putative viral interactions. As indicated above (**Section 2.4**) for armadillo GH and elephant and bat prolactin, substitutions on the hormone occurred at sites distal from the receptor binding sites, and accelerated evolution of ligand is not matched by concurrent evolution of receptor ecd. This would accord with a virus simply interacting with the ligand in a fashion that allows the latter to bind and activate its receptor as normal. On the other hand, for primates and cetartiodactyls, accelerated evolution of GH and prolactin is accompanied by accelerated evolution of GHR and PRLR, suggesting that the virus may interact with both components, perhaps enhancing ligand-receptor binding, or possibly binding to receptor independently of ligand. It is possible that in some cases interaction is primarily between virus and receptor, and that accelerated evolution of the hormone occurs as a consequence of the changes in the receptor. This would be more similar to the roles of EGFR or IGF1R as viral receptors (see **Section 5**). This might apply particularly in the cases of rodent and hyrax GHRs, where accelerated evolution of the receptors is not accompanied by corresponding rapid change of GH (**Figure 2**).

7. Placental lactogens as potential decoys; defence against infection of foetus

Decoy receptors are used by both hosts and pathogens to modulate the mammalian immune response [68]. Thus, the actions of many cytokines are controlled by the host by production of truncated, soluble forms of cytokine receptors that bind the cytokine but do not activate signalling. Similarly, viruses may produce proteins equivalent to truncated host receptors that bind cytokines such as interleukin-1 and interfere with their antiviral actions [68].

Decoy receptors produced by the host that bind to viruses and interfere with their ability to infect cells have also been described. For example, soluble glycans in human milk can bind to rotaviruses, potentially providing protection against gastroenteritis [69]. Sulphated glycosaminoglycans act as decoy receptors for human adenovirus type 37, reducing infection of corneal epithelial cells [70]. I suggest here that PLs act as decoy receptors, protecting a foetus from the damaging effects of a virus.

Duplications of the genes encoding GH or prolactin to give complex gene clusters, many of which are expressed in the placenta, have occurred at least four times in mammalian evolution. In each case the appearance of the gene duplications follows an episode of rapid evolution of GH and/or prolactin. If a host-virus arms race is the basis of the accelerated evolution, then the gene duplication(s) seem to be associated with relaxation of this arms race, possibly by providing an alternative mechanism to limit viral damage. This could occur if the placental protein(s) acted as decoys, binding virus without providing a mechanism for cell entry. The placental gene cluster in human is relatively simple and best understood. hPL is produced in very

large amounts, and has very low affinity for the GHR [44]. It can bind to PRLR, but whether it activates this receptor is not clear - in some *in vitro* assays it shows marked species specificity: for example, activation of human PRLR is much lower than that of rabbit or rat PRLR [71]. The very high concentrations of hPL could in fact inhibit, rather than activate, the receptor, as has been shown for high concentrations of GH [72] and prolactin [73] for their respective receptors. If hPL retained the ability to bind viruses it could act as a decoy, blocking viral attachment and preventing entry and infection. Furthermore, pituitary GH expression is suppressed during pregnancy and replaced by placental GH-V, which retains growth-promoting but not lactogenic activity [45]. If GH-V also lacked ability to bind virus, it could allow basic endocrine function to be maintained during pregnancy without risk of viral infection.

This mechanism could allow for the defence against viral infection during pregnancy, guarding against specific, as yet unidentified viruses. In particular it would protect the foetus from infection. It would not act as a defence in non-pregnant individuals. The implication is that the virus is a serious threat to the unborn child, but less so to adults or children. Such a situation is seen in many human viruses, for example rubella and Zika [49, 74] though it is not suggested that GH or prolactin may act as receptors for these. Infection by rubella virus also increases the spontaneous abortion rate [49]. An important point is that the PLs and GH-V are expressed in and secreted from the foetal part of the placenta - they are foetal rather than maternal genes [75,76].

Although protection of the foetus from infection is important for both mother and offspring, the balance is not even. For the mother loss or serious damage to the foetus would represent a significant loss of fitness, but this would be limited if the ability to

produce further offspring was unaffected. For the foetus viral infection would potentially lead to a 100% loss of fitness due to death or serious disability.

8. Episodes of rapid evolution occur in other protein hormones and growth factors.

Receptors for peptide hormones and growth factors clearly provide a potential route for virus to enter the host cell. The use of EGFR and IGF1R in this way has been discussed above

Episodes of rapid evolution as seen in GH and prolactin are also seen in several other polypeptide hormones [77], for example luteinizing hormone (LH) and insulin. LH, one of the gonadotropins that control mammalian reproduction, comprises two subunits, α and β , both of which show an episode of rapid change during primate evolution. Interestingly, the end of this period appears to coincide with the appearance of a placental protein, human chorionic gonadotropin, derived from duplication of the LH β -subunit gene and expressed from a cluster of 6 genes (in addition to *LH*) all encoding identical proteins [78-81]. The similarity to human GH and PL is striking. Insulin is a small protein that regulates carbohydrate metabolism by binding to a receptor tyrosine kinase. In mammals it is generally strongly conserved, except in hystricomorph rodents (guinea pig and relatives) where insulin evolution is very rapid [82,83]. A significant, though less marked, period of rapid evolution of insulin and IGF-I also occurs in NWM [84,85]. The pattern of evolution in these protein hormones suggests that here too host-virus arms races might be involved.

9. Testing the hypothesis

The case is made here that the pattern of evolution of GH and prolactin in mammals reflects host-pathogen arms races. The crucial missing component is of course the virus(es), or possibly other pathogen(s). In humans the implication is that the virus still occurs, or has done until recently, possibly showing minimal symptoms except in the foetus - this is implied because PLs continue to be strongly conserved and produced in large amounts. Investigation of interactions between GH, PL or prolactin and viroid material in individual serum samples might be revealing. It is known that both GH and prolactin can occur in human serum as high molecular weight forms ("big-" and "big-big-" GH and prolactin; macroprolactin), and that the relative amounts of these vary between individuals [86,87]. The "big" forms are probably dimers, while the "big-big-" forms include oligomers (especially GH [86]) and complexes with autoantibodies (especially prolactin [87]), together with uncharacterized material. The possible inclusion of viroid material in the last category would be worth investigation. A further avenue of investigation in humans could be the rare cases in which PLs are deleted (see above); in most cases healthy offspring are produced, but less commonly severe growth retardation is observed - consideration of whether viral infection is contributing to the latter would be appropriate.

Of the various potential non-human models, bats might be instructive. Some bats show very rapid evolution of GH and, particularly, prolactin, which appears to be ongoing. Bats are infected by many viruses [88,89] and the possibility of interaction between GH, prolactin, GHR and/or PRLR and viruses in bats would be worth

investigation. Sheep also provide a further potentially useful animal model, because of a recent duplication of the GH gene occurring as a polymorphism in some breeds; one of the duplicates is expressed in the placenta [38,39]. The possibility that the placentally expressed gene may give some viral protection would be worth investigation.

10. Conclusions and outlook

The hypothesis proposed here arises from the observations of markedly episodic evolution in GH, prolactin and their receptors, and the occurrence in some mammals of PLs, resulting from duplications of the *GH* and prolactin genes and expressed at very high levels during pregnancy. Convincing explanations of these phenomena in terms of physiological function are not yet available. The explanation that they reflect host-virus arms races and mechanisms to protect the foetus against infection seems a plausible alternative, though it is clearly speculative in the absence of evidence about the viruses involved, and does not necessarily exclude alternative explanations for some of the evolutionary change observed.

The hypothesis does not just provide an explanation for a puzzling evolutionary curiosity - it also has medical implications, particularly regarding the role of hPL. hPL is generally considered a hormone, but is remarkable in that during pregnancy it reaches very high, rather constant, concentration, about 100-fold greater than normal circulating level for most polypeptide hormones. High, unfluctuating concentration does not accord with an endocrine function and possible role in endocrine related

disease, but would fit with a role as a decoy in defending the foetus against pathogens, particularly given that PLs are produced by foetal tissue.

An implication of the proposal made here is that much of the change introduced into pituitary GH, prolactin and their receptors during the course of mammalian evolution has been driven by viral (or other pathogen) infections. This applies both to the episodes of rapid change and to duplication of hormone genes. The consequences may go further than defence against pathogens, however, allowing more extensive exploration of evolutionary space, and facilitating production of gene copies that can be adapted to additional physiological roles. This would be another example of the pervasive role that viruses have played in mammalian evolution [90,91].

Conflict of Interest

The author declares no conflict of interest.

Acknowledgements

I thank the editor and three anonymous reviewers for helpful comments.

References

- [1] **Grove J, Marsh M.** 2011. The cell biology of receptor-mediated virus entry. *J Cell Biol* **195**: 1071-1082.
- [2] **Backovic M, Rey FA.** 2012. Virus entry: old viruses, new receptors. *Curr Op Virol* **2**: 4-13.
- [3] **Helenius A.** 2018. Virus entry: looking back and moving forward. *J Mol Biol* **430**: 1853-1862.
- [4] **Maginnis MS.** 2018. Virus-receptor interactions: the key to cellular invasion. *J Mol Biol* **430**: 2590-2611.
- [5] **Coffin JM.** 2013. Virions at the gates: receptors and the host–virus arms race. *PLoS Biol* **11**: e1001574.
- [6] **Daugherty MD, Malik HS.** 2012. Rules of engagement: molecular insights from host-virus arms races. *Annu Rev Genet* **46**: 677-700.
- [7] **Sironi M, Cagliani R, Forni D, Clerici M.** 2015. Evolutionary insights into host–pathogen interactions from mammalian sequence data. *Nat Rev Genet* **16**:224-236.
- [8] **Wang W, Zhao H, Han G-Z.** 2020. Host-virus arms races drive elevated adaptive evolution in viral receptors. *J Virol* **94**: e00684-20.
- [9] **Forsyth IA, Wallis M.** 2002. Growth hormone and prolactin-molecular and functional evolution. *J Mam Gland Biol Neoplasia* **7**: 291-312.
- [10] **Ocampo Daza D, Larhammar D.** 2018. Evolution of the growth hormone, prolactin, prolactin 2 and somatolactin family. *Gen Comp Endocrinol* **264**: 94-112.

- [11] **Brooks AJ, Dai W, O'Mara ML, Abankwa D, Chhabra Y, Pelekanos RA,...Waters MJ.** 2014. Mechanism of activation of protein kinase JAK2 by the growth hormone receptor. *Science* **344**: 1249783.
- [12] **Dehkhoda F, Lee CMM, Medina J, Brooks AJ.** 2018. The growth hormone receptor: mechanism of receptor activation, cell signaling , and physiological aspects. *Front Endocrinol* **9**: 35.
- [13] **Waters MJ.** 2016. The growth hormone receptor. *GH & IGF Res* **28**: 6-10.
- [14] **Wallis M.** 1981. The molecular evolution of pituitary growth hormone, prolactin and placental lactogen: a protein family showing very variable rates of evolution. *J Mol Evol* **17**: 10-18.
- [15] **Wallis M.** 1994. Variable evolutionary rates in the molecular evolution of mammalian growth hormones. *J Mol Evol* **38**: 619-627.
- [16] **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.
- [17] **Wallis M.** 2000. Episodic evolution of protein hormones: Molecular evolution of pituitary prolactin. *J Mol Evol* **50**: 465-473.
- [18] **Wallis M.** 2008. Mammalian genome projects reveal new growth hormone (GH) sequences. Characterization of the GH-encoding genes of armadillo (*Dasypus novemcinctus*), hedgehog (*Erinaceus europaeus*), bat (*Myotis lucifugus*), hyrax (*Procavia capensis*), shrew (*Sorex araneus*), ground squirrel (*Spermophilus tridecemlineatus*), elephant (*Loxodonta africana*), cat (*Felis catus*) and opossum (*Monodelphis domestica*). *Gen Comp Endocrinol* **155**: 271-279.
- [19] **Papper Z, Jameson NM, Romero R, Weckle AL, Mittal P, Benirschke K,...Wildman DE.** 2009. Ancient origin of placental expression in the growth

hormone genes of anthropoid primates. *Proc Natl Acad Sci USA* **106**: 17083-17088.

- [20] **Sydykova DK, Wilke CO.** 2017. Calculating site-specific evolutionary rates at the amino-acid or codon level yields similar rate estimates. *PeerJ* **5**: e3391.
- [21] **Wallis M.** 2018. Molecular evolution of prolactin in Chiroptera: Accelerated evolution and a large insertion in vespertilionid bats. *Gen Comp Endocrinol* **269**: 102-111.
- [22] **Wallis M.** 2009. Prolactin in the Afrotheria: characterization of genes encoding prolactin in elephant (*Loxodonta africana*), hyrax (*Procavia capensis*) and tenrec (*Echinops telfairi*). *J Endocrinol* **200**: 233-240.
- [23] **Graur D.** 2016. *Molecular and genome evolution*. Sinauer Associates Inc.
- [24] **Li Y, Wallis M, Zhang Y-P.** 2005. Episodic evolution of prolactin receptor gene in mammals: coevolution with its ligand. *J Mol Endocrinol* **35**: 411-419.
- [25] **Liu J-C, Makova KD, Adkins RM, Gibson S, Li WH.** 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.
- [26] **Yi S, Bernat B, Pál G, Kossiakoff A, Li W-H.** 2002. Functional promiscuity of squirrel monkey growth hormone receptor toward both primate and nonprimate growth hormones. *Mol Biol Evol* **19**:1083-1092.
- [27] **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.
- [28] **de Vos AM, Ultsch M, Kossiakoff A.** 1992. Human growth hormone and extracellular domain of its receptor: crystal structure of the complex. *Science* **255**: 306-312.

- [29] **Broutin I, Jomain J-B, Tallet E, van Agthoven J, Raynal B, Hoos S,...Goffin V.** 2010. Crystal structure of an affinity-matured prolactin complexed to its dimerized receptor reveals the topology of hormone binding site 2. *J Biol Chem* **285**: 8422-8433.
- [30] **Chen EY, Liao Y-C, Smith DH, Barrera-Saldaña HA, Gelinas RE, Seeburg PH.** 1989. The human growth hormone locus: nucleotide sequence, biology, and evolution. *Genomics* **4**: 479-497.
- [31] **Pérez-Maya AA, Wallis M, Barrera-Saldaña HA.** 2016. Structure and evolution of the gorilla and orangutan *growth hormone* loci. *Mamm Genome* **27**: 511-523.
- [32] **Revol de Mendoza A, Esquivel Escobedo D, Martínez Dávila I, Barrera Saldaña H.** 2004. Expansion and divergence of the GH locus between spider monkey and chimpanzee. *Gene* **336**:185-193.
- [33] **Ye C, Li Y, Shi P, Zhang Y-P.** 2005. Molecular evolution of growth hormone gene family in old world monkeys and hominoids. *Gene* **350**: 183-192.
- [34] **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*). *J Mol Evol* **63**: 591-601.
- [35] **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.
- [36] **Li Y, Ye C, Shi P, Zou X-J, Xiao R, Gong Y-Y, Zhang Y-P.** 2005. Independent origin of the growth hormone gene family in New World monkeys and Old World monkeys/hominoids. *J Mol Endocrinol* **35**: 399-409.

- [37] **Revol de Mendoza A, Esquivel Escobedo D, Santiago Alarcón D, Barrera Saldaña H.** 2001. Independent duplication of the growth hormone gene in three anthropoidean lineages. *J Endocr Genet* **2**:151–159.
- [38] **Valinsky A, Shani M, Gootwine E.** 1990. Restriction fragment length polymorphism in sheep at the growth hormone locus is the result of variation in gene number. *Animal Biotechnol* **1**:135-144.
- [39] **Wallis M, Lioupis A, Wallis OC.** 1998. Duplicate growth hormone genes in sheep and goat. *J Mol Endocrinol* **21**:1-5.
- [40] **Soares MJ, Konno T, Alam SMK.** 2007. The prolactin family: effectors of pregnancy-dependent adaptations. *Trends Endocrinol Metab* **18**: 114-121.
- [41] **Wiemers DO, Shao L-J, Ain R, Dai G, Soares MJ.** 2003. The mouse prolactin gene family locus. *Endocrinology* **144**: 313-325.
- [42] **Alam SMK, Ain R, Konno T, Ho-Chen JK, Soares MJ.** 2006. The rat prolactin gene family locus: species-specific gene family expansion. *Mamm Genome* **17**:859-877.
- [43] **Takahashi T, Hayashi K-G, Hosoe M.** 2013. Biology of the placental proteins in domestic ruminants: expression, proposed roles and practical applications. *Japan Agric Res Quart* **47**: 43-51.
- [44] **Handwerger S, Freemark M.** 2000. The roles of placental growth hormone and placental lactogen in the regulation of human fetal growth and development. *J Pediatr Endocrinol Metab.* **14**: 343-356.
- [45] **Liao S, Vickers MH, Stanley JL, Baker PN, Perry JK.** 2018. Human placental growth hormone variant in pathological pregnancies. *Endocrinology* **159**: 2186-2198.

- [46] **Newbern D, Freemark M.** 2011. Placental hormones and the control of maternal metabolism and fetal growth. *Curr Opin Endocrinol Diabetes Obes* **18**: 409-416.
- [47] **Wurzel JM, Parks JS, Herd JE, Nielsen PV.** 1982. A gene deletion is responsible for absence of human chorionic somatomammotropin. *DNA* **1**: 251-257.
- [48] **Rygaard K, Revol A, Esquivel-Escobedo D, Beck BL, Barrera-Saldaña HA.** 1998. Absence of human placental lactogen and placental growth hormone (HGH-V) during pregnancy: PCR analysis of the deletion. *Hum Genet* **102**: 87-92.
- [49] **Pereira L.** 2018. Congenital viral infection: Traversing the uterine-placental interface. *Annu Rev Virol* **5**: 273-299.
- [50] **Klindt J, Robertson MC, Friesen HG.** 1982. Temporal concentrations of rat placental lactogen in the peripheral circulation. *Biol Repr* **27**: 407-410.
- [51] **Duggal NK, Emerman M.** 2012 Evolutionary conflicts between viruses and restriction factors shape immunity. *Nat Rev Immunol* **12**: 687-695.
- [52] **Kok K-H, Jin D-Y.** 2013. Balance of power in host-virus arms races. *Cell Host & Microbe* **14**: 5-6.
- [53] **Ebert D, Fields PD.** 2020. Host-parasite co-evolution and its genomic signature. *Nat Rev Genet* **21**:754-768..
- [54] **Hoffmann H-H, Schneider WM, Rice CM.** 2015 Interferons and viruses: an evolutionary arms race of molecular interactions. *Trends Immunol* **36**: 124-138.
- [55] **Demogines A, Abraham J, Choe H, Farzan M, Sawyer SL.** 2013 Dual host-virus arms races shape an essential housekeeping protein. *PLoS Biol* **11**: e1001571.

- [56] **Carrillo-Bustamante P, Kesmir C, de Boer RJ.** 2015 A coevolutionary arms race between hosts and viruses drives polymorphism and polygenicity of NK cell receptors. *Mol Biol Evol* **32**: 2149-2160.
- [57] **Staring J, Raaben M, Brummelkamp TR.** 2018. Viral escape from endosomes and host detection at a glance. *J Cell Sci* **131**: jcs216259.
- [58] **Spriggs CC, Harwood MC, Tsai B.** 2019. How non-enveloped viruses hijack host machineries to cause infection. *Adv Virus Res* **104**: 97-122.
- [59] **Zheng K, Kitazato K, Wang Y.** 2014. Viruses exploit the function of epidermal growth factor receptor. *Rev Med Virol* **24**: 274-286.
- [60] **Eierhoff T, Hrincius ER, Rescher U, Ludwig S, Ehrhardt C.** 2010. The epidermal growth factor receptor (EGFR) promotes uptake of influenza A viruses (IAV) into host cells. *PLoS Pathog* **6**: e1001099.
- [61] **Nguyen CC, Kamil JP.** 2018. Pathogen at the gates: human cytomegalovirus entry and cell tropism. *Viruses* **10**: 704.
- [62] **Griffiths CD, Bilawchuk LM, McDonough JE, Jamieson KC, Elawar F, Cen Y,...Marchant DJ.** 2020. IGF1R is an entry receptor for respiratory syncytial virus. *Nature* **583**: 615-619.
- [63] **Altindis E, Cai W, Sakaguchi M, Zhang F, GuoXiao W, Liu F,...Kahn CR.** 2018. Viral insulin-like peptides activate human insulin and IGF-1 receptor signaling: A paradigm shift for host–microbe interactions. *Proc Natl Acad Sci USA* **115**: 2461-2466.
- [64] **Huang Q, Kahn CR, Altindis E.** 2019. Viral hormones: expanding dimensions in endocrinology. *Endocrinology* **160**: 2165-2179.
- [65] **Irwin DM.** 2019. Viral hormones: do they impact human endocrinology? *Endocrinology* **160**: 2326-2327.

- [66] **Deng L, He K, Wang X, Yang N, Thangavel C, Jiang J,...Frank SJ.** 2007. Determinants of growth hormone receptor down-regulation. *Mol Endocrinol* **21**: 1537-1551.
- [67] **Swaminathan G, Varghese B, Thangavel C, Carbone CJ, Plotnikov A, Kumar KGS,...Fuchs SY.** 2008. Prolactin stimulates ubiquitination, initial internalization, and degradation of its receptor via catalytic activation of Janus kinase 2. *J Endocrinol* **196**: R1-R7.
- [68] **Felix J, Savvides SN.** 2017. Mechanisms of immunomodulation by mammalian and viral decoy receptors: insights from structures. *Nat Rev Immunol* **17**: 112-129.
- [69] **Yu Y, Lasanajak Y, Song X, Hu L, Ramani S, Mickum ML,...Smith DF.** 2014. Human milk contains novel glycans that are potential decoy receptors for neonatal rotaviruses. *Mol Cell Proteomics* **13**: 2944-2960.
- [70] **Chandra N, Liu Y, Liu J-X, Frångsmyr L, Wu N, Silva LM,...Arnberg N.** 2019. Sulfated glycosaminoglycans as viral decoy receptors for human adenovirus type 37. *Viruses* **11**: 247.
- [71] **Oclon E, Lesniak-Walentyn A, Solomon G, Shpilman M, Hrabia A, Gertler A.** 2017. Comparison of *in vitro* bioactivity of chicken prolactin and mammalian lactogenic hormones. *Gen Comp Endocrinol* **240**: 27-34.
- [72] **Fuh G, Cunningham BC, Fukunaga R, Nagata S, Goeddel DV, Wells JA.** 1992. Rational design of potent antagonists to the human growth hormone receptor. *Science* **256**: 1677-1680.
- [73] **Fuh G, Wells, JA.** 1995. Prolactin receptor antagonists that inhibit the growth of breast cancer cell lines. *J Biol Chem* **270**: 13133-13137.

- [74] **Sirohi D, Kuhn RJ.** 2017. Zika virus structure, maturation, and receptors. *J Infect Dis* **216**: (Suppl 10) S935-S944.
- [75] **Haig D.** 1996. Placental hormones, genomic imprinting, and maternal-fetal communication. *J Evol Biol* **9**: 357-380.
- [76] **Haig D.** 2008. Placental growth hormone-related proteins and prolactin-related proteins. *Placenta* **22**: S36-S41.
- [77] **Wallis M.** 2001. Episodic evolution of protein hormones in mammals. *J Mol Evol* **53**: 10-18.
- [78] **Talmadge K, Vamvakopoulos NC, Fiddes JC.** 1984. Evolution of the genes for the β subunits of human chorionic gonadotropin and luteinizing hormone. *Nature* **307**: 37-40.
- [79] **Hallast P, Nagirnaja L, Margus T, Laan M.** 2005. Segmental duplications and gene conversion: Human luteinizing hormone/chorionic gonadotropin β gene cluster. *Genome Res* **15**: 1535-1546.
- [80] **Maston GA, Ruvolo M.** 2002. Chorionic gonadotropin has a recent origin within primates and an evolutionary history of selection. *Mol Biol Evol* **19**: 320-335.
- [81] **Casarini L, Santi D, Brigante G, Simoni M.** 2018. Two hormones for one receptor: evolution, biochemistry, actions, and pathophysiology of LH and hCG. *Endocr Rev* **39**: 549-592.
- [82] **Beintema JJ, Campagne RN.** 1987. Molecular evolution of rodent insulins. *Mol Biol Evol* **4**: 10-18.
- [83] **Irwin, DM.** 2020. Variation in the rates of evolution of the insulin and glucagon hormone and receptor genes in rodents. *Gene* **728**: 144296.

- [84] **Wallis M.** 2009. New insulin-like growth factor (IGF)-precursor sequences from mammalian genomes: the molecular evolution of IGFs and associated peptides in primates. *GH & IGF Res* **19**: 12-23.
- [85] **Wallis M.** 2015. Coevolution of insulin-like growth factors, insulin and their receptors and binding proteins in New World Monkeys. *GH & IGF Res* **25**: 158-167.
- [86] **Baumann GP.** 2009. Growth hormone isoforms. *GH & IGF Res* **19**: 333-340.
- [87] **Gibney J, Smith TP, McKenna TJ.** 2005. Clinical relevance of macroprolactin. *Clin Endocrinol* **62**: 633-643.
- [88] **Hayman DTS.** 2016. Bats as viral reservoirs. *Annu Rev Virol* **3**: 77-99..
- [89] **Schountz T.** 2014. Immunology of bats and their viruses: challenges and opportunities. *Viruses* **6**: 4880-4901.
- [90] **Enard D, Cai L, Gwennap C, Petrov DA.** 2016. Viruses are a dominant driver of protein adaptation in mammals. *eLife* **5**: e12469.
- [91] **Slodkowicz G, Goldman N.** 2020. Integrated structural and evolutionary analysis reveals common mechanisms underlying adaptive evolution in mammals. *Proc Natl Acad Sci USA* **117**: 5977-5986.
- [92] **Wallis M.** 2014. Molecular evolution of growth hormone. *The Biochemist* **36**: 4-8.
- [93] **Yang Z.** 2007. PAML 4: Phylogenetic analysis by maximum likelihood. *Mol Biol Evol* **24**: 1586-1591.
- [94] **Upham NS, Esselstyn JA, Jetz W.** 2019. Inferring the mammal tree: Species-level sets of phylogenies for questions in ecology, evolution, and conservation. *PLoS Biol* **17**: e3000494.

- [95] **Zhang J, Nielsen R, Yang Z.** 2005. Evaluation of an improved branch-site likelihood method for detecting positive selection at the molecular level. *Mol Biol Evol* **2**: 2472-2479.
- [96] **Yates AD, Achuthan P, Akanni, W, Allen J, Allen J, Alvarez-Jarreta J,...Flicek P.** 2020. Ensembl 2020. *Nucl Acids Res* **48**: D682-D688.
- [97] **Yang Z, Bielawski JP.** 2000. Statistical methods for detecting molecular adaptation. *Trends Ecol Evo* **15**:496-503.

BOX 1

Synonymous and Non-synonymous substitutions; evolutionary trees

Nucleotide substitutions occurring in a coding sequence during the course of evolution can be of two types, synonymous (changes that do not alter the amino acid sequence) and nonsynonymous (which lead to change in amino acid sequence). It is generally considered that synonymous substitutions are mostly neutral with regard to gene function, whereas nonsynonymous changes may be neutral, deleterious or adaptive. The ratio between the rate of change of nonsynonymous substitutions and that of synonymous substitutions (dN/dS) is frequently used to assess the nature of evolutionary change at the molecular level [97]. For most protein evolution, dN/dS is low, reflecting the rejection of most nonsynonymous changes by purifying selection. If a gene loses its function dN/dS may increase to 1.0. If a gene is subject to adaptive/positive selection dN/dS is expected to increase. A dN/dS value significantly greater than 1.0 is a clear indication of positive selection, but a value lower than this does not necessarily exclude it.

For a coding sequence, evolutionary trees can be constructed using either dN or dS . The programme *codeml* in the PAML package [93] was used to do this for the trees shown in **Figures 2 and 3**, using branch model 1. To construct these, a defined tree, based on generally accepted mammalian phylogeny [94] was used to determine the topology; branch lengths are determined by the individual sequences used. The long branch lengths seen in some cases in the dN trees reflect episodes of rapid evolution, and are generally not matched by long branches in the corresponding dS tree. For all

these cases dN/dS increases markedly, but in only a few cases is $dN/dS > 1$. This probably partly reflects the low basal dN/dS values for the trees. The branch-site model A in codeml was also used, allowing dN/dS to vary among sites and branches [93,95]. Foreground dN/dS values for positively selected residues are included in **Figures 2 and 3**. Whether these were significantly greater than 1 was tested using likelihood ratios and the branch-site test of positive selection.

Figure legends

Figure 1. Growth hormone bound to the two subunits of its homodimeric receptor.

Two distinct sites on the hormone (GH, 1 and 2) bind to similar sites on the two receptor subunits. Specific tyrosine residues (Y) on the intracellular domains are subject to phosphorylation by tyrosine kinase JAK2.

Figure 2. Phylogenetic trees for mammalian GH and GHR. Trees were constructed using Synonymous or Nonsynonymous substitutions, as described previously [18, 92], using alignments of coding (nucleotide) sequences and the codeml programme in the PAML package [93]. Tree topology was determined using a defined tree based on established mammalian phylogeny [94]; the codeml programme determines branch lengths. Episodes of accelerated evolution (branches on which the number of nonsynonymous substitutions exceeded the number of synonymous substitutions and/or the substitution rate (dN) was 5-fold greater than that of an appropriate sister branch) are shown as thick blue lines. The numbers on these lines are dN/dS values for site class 2, determined using the branch-site method in codeml [93,95; see **Box 1**]. Values significantly greater than 1.0 (likelihood ratio test) are shown as * (P<0.05) and ** (P<0.01).

Figure 3. Phylogenetic trees for mammalian prolactin and PRLR. Trees were constructed using Synonymous or Nonsynonymous substitutions, as described previously [21,22]. Other details are as given in the legend of Fig. 2.

Figure 4. Binding of GH to the extracellular domain of its receptor. The structure of human GH bound to two molecules of receptor ecd [28] is shown in each panel, with GH in space-filling format (blue and yellow), and the two chains of the receptor in cartoon style (purple). Three views are given, (a) sideways on (membrane at the bottom; orientation as in **Figure 1**), (b) from the top and (c) from the bottom (looking up from the membrane). Residues changing during the episodes of rapid evolution leading to GH of Simiiformes and armadillo GH are shown in yellow. For GH of Simiiformes (top) most substitutions occurred on the side close to the membrane, a substantial proportion (~30%) of these being within 5Å of the receptor-binding site. For armadillo GH (bottom) most of the substitutions occurred in the region of the molecule away from the receptor. Constructed using PyMOL [The PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC] and pdb (RCSB protein database) entry 3hhr (human GH bound to the extracellular domain of its receptor). Revised from [18,92].

Figure 5. *GH* gene loci in three primates (human, marmoset and bushbaby). Based on [30] (human), [35] (marmoset) and genome assembly in Ensembl [96] (bushbaby). *CD79B* and *SMARCD2* are genes immediately upstream and downstream, respectively, of the GH locus.

Figure 1. Growth hormone bound to the two subunits of its homodimeric receptor.

Figure 2. Phylogenetic trees for mammalian GH and GHR.

Figure 3. Phylogenetic trees for mammalian prolactin and PRLR.

Figure 4. Binding of GH to the extracellular domain of its receptor.

Figure 5. *GH* gene loci in three primates (human, marmoset and bushbaby).