

# **Mini-Review: New insights into the evolution of chorionic gonadotrophin**

Alexander Henke, Jörg Gromoll

### **To cite this version:**

Alexander Henke, Jörg Gromoll. Mini-Review: New insights into the evolution of chorionic gonadotrophin. Molecular and Cellular Endocrinology, 2008, 291 (1-2), pp.11. 10.1016/j.mce.2008.05.009 hal-00532027

# **HAL Id: hal-00532027 <https://hal.science/hal-00532027v1>**

Submitted on 4 Nov 2010

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

### Accepted Manuscript

Title: Mini-Review: New insights into the evolution of chorionic gonadotrophin

Authors: Alexander Henke, Jörg Gromoll





Please cite this article as: Henke, A., Gromoll, J., Mini-Review: New insights into the evolution of chorionic gonadotrophin, *Molecular and Cellular Endocrinology* (2007), doi:[10.1016/j.mce.2008.05.009](dx.doi.org/10.1016/j.mce.2008.05.009)

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

#### **NUSCRIPT** U, CCEP D 말



# **NUSCR**

#### Revised version of manuscript MCE-D-08-00055

Accepted Manuscript

### **Abstract**

 The glycoprotein hormones luteinizing hormone (LH) and chorionic gonadotrophin (CG) are crucial for reproduction, as LH induces sex hormone production and ovulation, and CG is essential for the establishment of pregnancy and fetal male sexual differentiation. Both consist 6 of two heterodimeric peptides of which the  $\alpha$ -subunit is common to both hormones whereas the β-subunit is hormone-specific. The CGB gene was derived from LHB by gene duplication and frame shift mutation that led to a read-through into the formerly 3'-untranslated region, giving rise to the carboxyl-terminal peptide. Owing to nucleotide changes within the 5'-region of CGB, a new transcriptional start site and regulatory region was gained. These changes led to the specific expression of CGB in the placenta and its decrease in the pituitary. Recent findings on gonadotrophins led to an extended model for the sequence of events in the evolution of the CGB gene in primates and its tissue-specific expression.

#### Revised version of manuscript MCE-D-08-00055

#### **Gonadotrophin physiology**

 Vertebrate reproduction is controlled by hormones of the hypothalamic-pituitary-gonadal axis. The adenohypophysis synthesizes and releases the gonadotrophins luteinizing hormone (LH) and follicle stimulating hormone (FSH), which act on the gonads and induce sex steroid synthesis and gamete production and maturation in both sexes. FSH is necessary for follicle maturation in the ovary and the stimulation of testicular Sertoli cells to support spermatogenesis. LH triggers sex steroid production in the ovary, whereas LH peak levels induce ovulation. In the testis, LH induces testosterone biosynthesis in Leydig cells.

and follicle stimulating hormone (FSH), which act on the gonads and induce sex steroid<br>sis and gamete production and maturation in both sexes. FSH is necessary for follicle<br>tion in the ovary and the stimulation of testicul The third gonadotrophin, chorionic gonadotrophin (CG), is the first biochemical signal from the embryo to the maternal environment. It is produced by extraembryonic trophoblasts at the beginning of a pregnancy and stimulates the corpus luteum to maintain progesterone synthesis that represses further follicle maturation, thereby ensuring embryo implantation and continuity of the pregnancy (Pierce and Parsons 1981). During pregnancy, CG is produced by the fetal part of the placenta and is also crucial for male fetal sexual differentiation, as CG stimulates the fetal testosterone synthesis in the testicular Leydig cells (Abramovich et al. 1974; Clements et al. 1976; Huhtaniemi et al. 1977; Tapanainen et al. 1981; Gromoll et al. 2000). Further biochemical and physiological details on LH and CG are compiled in Table 1. Both LH and CG bind to the same receptor, the LH-/CG-receptor (LHR). Reports show that the expression of LHR is not restricted to the gonads, but is also expressed in the endometrium and other reproduction-associated tissues (Ziecik et al. 2007). It was therefore speculated that the special structure of the primate placenta favoured CG function in facilitating embryo implantation (Licht et al. 2007). Reports of CG's promotion of angiogenesis support the hypothesis that the embryo fosters maternal blood vessel growth via CG for better supply of nutrients and oxygen and easier release of CG and other factors (Zygmunt et al. 2002; Herr et al. 2007). Furthermore, it was reported that CG is necessary for the invasion of cytotrophoblasts into the endometrium during embryo implantation (Islami et

# TED

Revised version of manuscript MCE-D-08-00055

 al. 2001; Carver et al. 2003). One study demonstrated the association between recurrent miscarriage and low serum CG values in patients; it was concluded that low serum CG hinders embryo implantation (Rull and Laan 2005).

 Biochemical studies have revealed that LH, FSH and CG consist of two subunits, of which the α-subunit is identical in all hormones and the β-subunit confers hormone-specificity (Bousfield and Ward 2006). Together with thyroid stimulating hormone (TSH) they form a family of glycoprotein hormones and the resolution of their three-dimensional structure classified them as part of the superfamily of cystine knot growth factors such as PDGF, NGF, VEGF or TGFβ (Lapthorn et al. 1994; Wu et al. 1994).

#### **Prepubertal gonadotrophin expression**

emical studies have revealed that LH, FSH and CG consist of two subunits, of which<br>subunit is identical in all hormones and the  $\beta$ -subunit confers hormone-specificity<br>field and Ward 2006). Together with thyroid stimulati LH and FSH in serum are not only detectable in adults but also at very low serum concentrations in prepubertal and pubertal children (Apter et al. 1989). In 5-6 year-old girls, a pulsatile and diurnal LH pattern is present and levels rise clearly before the onset of puberty (Dunkel et al. 1990a; Dunkel et al. 1990b; Dunkel et al. 1992; Demir et al. 1995; Clark et al. 1997; Mitamura et al. 2000). As found in man, a rise of LH and testosterone levels from the prepubertal to the adult could be shown in the rhesus macaque (Meeran et al. 2003). Despite these findings, the function, if any, of childhood gonadotrophins remains to be established.

#### **Human LH variants**

 There are several reports on common variant LH (V-LH), which escaped detection by conventional LH-antibodies (Pettersson et al. 1992). The investigation of the V-LHB gene identified two new SNPs in protein-coding parts that result in two amino acid changes: Trp8 => Arg8 and Ile15 => Thr15 (Furui et al. 1994; Okuda et al. 1994; Pettersson et al. 1994). Eight additional SNPs in the promoter of V-LHB lead to higher promoter activity compared

## PTED

Revised version of manuscript MCE-D-08-00055

 with the normal LHB promoter (Jiang et al. 1999). A study of a homozygous individual showed a higher activity of V-LH compared with normal LH but also a shorter half-life: 26 min in males and 44 min in females for V-LH versus 48 min in males and 53 min in females for WT-LT (Haavisto et al. 1995).

 V-LH is present in nearly all populations worldwide, posing an important genetic LHB 6 variant: among 30 different ethnicities, the carrier frequency had a mean  $\pm$  SD of 16.5  $\pm$  10.8 %, ranging from 0 % to 54% (Nilsson et al. 1997; Nilsson et al. 1998; Ramanujam et al. 1998; Elter et al. 1999; Ramanujam et al. 1999; Starka et al. 1999; Lamminen and Huhtaniemi 2001).

#### **Ectopic human CG**

T-LT (Haavisto et al. 1995).<br>
is present in nearly all populations worldwide, posing an important genetic LHB<br>
t: among 30 different ethnicities, the carrier frequency had a mean  $\pm$  SD of 16.5  $\pm$  10.8<br>
iging from 0 % Additional expression of CG was also found in the pituitary and at very low levels in the serum of normal individuals, in whom CG levels were slightly elevated in elderly individuals >50 years (Borkowski et al. 1984; Stenman et al. 1987; Hoermann et al. 1990; Odell et al. 1990; Alfthan et al. 1992; Birken et al. 1996; Dirnhofer et al. 1996). Furthermore, it was found that the CG levels were responsive to positive feedback via GnRH *in vivo* and *in vitro* and to negative feedback via progesterone and oestrogen *in vivo* (Stenman et al. 1987; Henke et al. 2007). Despite all these findings, so far no physiological function could be attributed to the CG expression in the pituitary and resulting serum CG in non-pregnant individuals.

#### **A common evolutionary origin of gonadotrophin subunits**

 The peptide sequences of the β-subunits of the gonadotrophins revealed a homology among them and therefore a common evolutionary origin (Dayhoff 1976). Sequence comparison of the human LHB and CGB mRNAs and genes displays high homology and a duplication of CGB from LHB, confirming the concept of a common evolutionary origin (Talmadge et al. 1984). The LHB and CGB transcripts encompass 1.1 and 1.5 kilobases, respectively, and are

#### CEPTED **MANUSCR**

Revised version of manuscript MCE-D-08-00055

- therefore relatively small transcripts. The detailed structures and lengths of the LHB/CGB
- genes and transcripts are given in Figure 1 and Table 2, respectively.
- 

#### **The structure of the LHB and CGB genes**

dructure of the LHB and CGB genes<br>duplication, copy diversification and the acquisition of new function are important<br>misms of gene evolution (Zhang 2003). In the case of primate gonadotrophin subunit<br>LHB underwent a gene Gene duplication, copy diversification and the acquisition of new function are important mechanisms of gene evolution (Zhang 2003). In the case of primate gonadotrophin subunit genes, LHB underwent a gene duplication, in which one copy was conserved as LHB, while the other copy developed into CGB by a single nucleotide deletion in codon 114; this led to a frame shift mutation, thereby elongating the open reading frame (ORF) into the previously 3' untranslated region (Table 2). In addition, two nucleotides inserted in codon 138 for human CGB led to a further frame-shift in the ORF. These changes resulted in eight different and 24 additional codons for CGB in comparison with LHB (Table 2).

#### **Primate LHB/CGB genes are organised in clusters**

 It was found that human CGB is a multicopy gene, as six CGB gene copies and one LHB gene copy form a cluster (Figure 1A) (Boorstein et al. 1982; Policastro et al. 1983; Talmadge et al. 1983; Fiddes and Talmadge 1984). The cluster is located on chromosome 19 at 19q13.32 (Naylor et al. 1983; Brook et al. 1984; Julier et al. 1984). However, the exact map of the cluster itself remained elusive until it was cloned completely as a 58 kb DNA fragment by Policastro *et al.*, (Policastro et al. 1986) and confirmed by others (Jameson and Lindell 1988). The single genes are adjacent to each other in tandem and inverted repeats within the gene cluster. A recent re-sequencing of this cluster in humans of different ethnicities revealed a high rate of SNPs and a low linkage disequilibrium due to excessive gene conversions so that this gene cluster represents a hot spot of genomic diversity (Hallast et al. 2005). CGB1 and CGB2 underwent further differentiation from canonical CGB owing to a DNA insertion (736 bp for CGB1, 724 bp for CGB2) and a deletion of 52 bp plus 356 bp in the region

#### Revised version of manuscript MCE-D-08-00055

 serving as 5'UTR in the paralogous CGB3/5/7/8 genes. These rearrangements led to a novel exon 1 (58 bp) and constitutive skipping of canonical exon 1, in which the donor splice site is disrupted. The two known exons 2 and 3 are included in the ORF but are frame-shifted, thereby leading to a hypothetical ORF of 132 codons with a new stop codon and elongated 3'UTR (Figure 1 and Table 2). Despite extensive descriptions of their mRNAs, no CGB1/2 protein product has been demonstrated yet (Bo and Boime 1992; Hollenberg et al. 1994; Dirnhofer et al. 1996; Hallast et al. 2007).

y leading to a hypothetical ORF of 132 codons with a new stop codon and elongated<br>
(Figure 1 and Table 2). Despite extensive descriptions of their mRNAs, no CGB1/2<br>
n product has been demonstrated yet (Bo and Boime 1992; H A BLAST search for CGB gene copies in the chimpanzee genome (Chimpanzee Sequencing and Analysis Consortium 2005) resulted in several hits within a contig allocated to chromosome 19 (NW\_001228244.1) but the hits were adjacent to gaps and did not allow an allocation of human CGB homologues to the chimpanzee counterparts except for LHB. We believe that a correct contig is not available yet.

 A BLAST search in the rhesus macaque genome (Rhesus Macaque Genome Sequencing and Analysis Consortium et al. 2007) resulted in four hits in the BAC CH250-325I15, available under accession number AC202849.6. These four hits were identified as one LHB and three CGB genes, which spanned approximately 30 kb and are shown as schematic cluster in Figure 1 B.

#### **Transcription of human CGB genes**

 In the human, the placental syncytiotrophoblasts are the major expression site of CGB during the first trimester of pregnancy. Gene copies CGB3/5/8 are expressed strongest whereas CGB1/2/7 genes are expressed at very low levels. CGB7 was also detected in normal breast tissue with an as yet unknown function (Talmadge et al. 1984; Bo and Boime 1992; Rull and Laan 2005). Interestingly, human CGB isoforms 3,5,8 were also detected, albeit at low expression, in the pituitary at the mRNA and protein levels (Hoermann et al. 1990; Odell et al. 1990; Birken et al. 1996; Dirnhofer et al. 1996). It follows that among the different human CGB gene copies there are differences in the strength of their promoters, dependent on the

# PTED

#### Revised version of manuscript MCE-D-08-00055

 tissue (Rull et al. 2008). It is also well known that epithelial tumours of the urogenital system, breast tumours and testis tumours can express *CGB* ectopically. Real time PCR of the mRNA isoforms in breast cancer revealed an up-regulation of *CGB* isoforms 3, 5 and 8 but not 1, 2 and7 (Bellet et al. 1997; Giovangrandi et al. 2001). However, the ultimate reason for ectopic CGB expression in some tumours is unknown.

#### **New World monkey LH/CG**

Bellet et al. 1997; Giovangrandi et al. 2001). However, the ultimate reason for ectopic<br>expression in some tumours is unknown.<br> **Norld monkey LH/CG**<br> **Norld monkey LH/CG**<br> **Norld monkey LH/CG**<br> **Norld monkey LH/CG**<br> **Norld**  A surprising plasticity of the LH/CG system was recently discovered in the common marmoset (*Callithrix jacchus*), a New World monkey. It appears that LHB became a pseudogene and its loss was compensated for by CGB (Muller et al. 2004; Scammell et al. 2008). The marmoset LHB and CGB genes have exon/intron structures homologous to the human situation, however marmoset CGB transcripts of the pituitary have a remarkably short 5'UTR of only 7 bp, which is comparable with other mammalian LHB 5'UTRs that are 9 bp in length ( Jameson et al. 1986; Muller et al. 2004; Henke et al. 2007).

#### **Gonadotropins in equid species**

 In the equine species horse (*Equus ferus caballus*), donkey (*Equus asinus*) and zebra (*Equus burchelli*), a single copy gene gives rise to an LHB/CGB-like gonadotrophin peptide that also harbours a CTP (Leigh and Stewart 1990; Sherman et al. 1992; Chopineau et al. 1995; Chopineau et al. 1999). The horse CTP was derived from a 10 bp deletion in exon 3, which was different from the single nucleotide deletion in the primate CGB gene. The same 22 promoter region of the horse LHB/CGB gene serves as promoter in the pituitary and placenta. The main difference from other mammalian LHB promoters is a small insertion (50 bp) that gave rise to a second TATA-box, leaving the remaining promoter largely unchanged (Sherman et al. 1992). Owing to these characteristics it was concluded that this single copy

#### Revised version of manuscript MCE-D-08-00055

eral studies at the protein level (Bousfield et al. 1987; Sugino et al. 1987).<br>
g gestation, equine CG is synthesised and secreted by the endometrial cups that arise<br>
the placental trophoblast chorionic girdle cells. In se gene evolved independently from the primate CGB, occurs only in equine species and therefore demonstrates convergent evolution at the molecular level (Sherman et al. 1992). This gene is expressed in both the pituitary and placenta of equine species and was confirmed in several studies at the protein level (Bousfield et al. 1987; Sugino et al. 1987). During gestation, equine CG is synthesised and secreted by the endometrial cups that arise from the placental trophoblast chorionic girdle cells. In serum, CG appears from day 40 of gestation onwards, increases until it reaches its peak level between days 55-70, decreases until day 130 and then becomes undetectable (Moore et al. 1975; Spincemaille et al. 1975; Wooding et al. 2001). In contrast, in humans, CG is detectable earlier in the pregnancy than in horses and drops after a peak to a low, but still detectable, level, until parturition. There are no reports on whether pregnant mare CG contributes to pregnancy recognition, since in horses the conceptus signals its presence to the maternal environment *via* prostaglandin F2α (PGF2α) in a paracrine manner. PGF2α normally induces luteolysis of the *corpus luteum* and results in pregnancy loss, when embryos secrete too much PGF2α. (Allen 2001). This mechanism contrasts not only the primate CG mechanism of pregnancy recognition but also the pregnancy recognition in ruminants via interferon tau (Roberts 2007; Roberts et al. 2008). Unfortunately, there are no studies on the mechanism of pregnancy recognition in prosimians and tarsier primate species that lack the CGB gene; otherwise it might have been possible to track the switch from prostaglandin/interferon-tau to CG-mediated pregnancy recognition.

#### **Glycosylation patterns and the role of the carboxy-terminal peptide (CTP) in CG**

 The human LHB and CGB peptides are N-glycosylated at Asn13 and Asn 30. In addition, the CTP of CGB prolongs the β-subunit peptide and enables extensive glycosylation of four novel O-glycoslation sites at Ser121, Ser 127, Ser 132 and Ser 138 (Birken and Canfield 1977; Keutmann and Williams 1977). CTP glycosylation leads to a prolonged half-life of 5-6 hours

#### Revised version of manuscript MCE-D-08-00055

1998c). The glycosylated CTP also results in an altered route of hormone secretion: LH<br>reted baso-laterally from gonadotrophic storage granules after GnRH-stimulation,<br>as CG is released constantly and apically by trophobla for CG compared with 40-60 minutes for LH and an increased molecular weight from 31 kDa for LH holohormone to 36 kDa for CG holohormone (Wehmann and Nisula 1981; Storring et al. 1988; Haavisto et al. 1995; le Cotonnec et al. 1998a; le Cotonnec et al. 1998b; le Cotonnec et al. 1998c). The glycosylated CTP also results in an altered route of hormone secretion: LH is secreted baso-laterally from gonadotrophic storage granules after GnRH-stimulation, whereas CG is released constantly and apically by trophoblasts towards the maternal blood circulation system (Table 2) (Childs et al. 1983; Handwerger et al. 1987; Lloyd and Childs 1988; Currie and McNeilly 1995; Jablonka-Shariff et al. 2002). *In vitro* experiments with polarised Madin-Darby canine kidney cells demonstrated that monomeric CGB subunit is released 30% baso-laterally and 70% apically, whereas the LHB subunit is released 30% apically and 70% baso-laterally. When the CTP is removed from the CGB peptide, it is released like LHB and thereby proving that the CTP determines the route of secretion (Jablonka-Shariff et al. 2002). As in equine species a previously cryptic CTP became expressed, and Nakav *et al.* (Nakav et al. 2005) investigated whether other species also possess a cryptic CTP in the 3' region of their LHB genes. *In vitro*, bovine LHB was fused to either bovine or human CTP. In both cases efficient glycosylation failed and therefore the protein was secreted baso-laterally as found for LH. The absence of apical routing, which is necessary for efficient placental release of CG into the maternal blood stream, provides an explanation of why CGB did not emerge in ruminants (Nakav et al. 2005). Furthermore, predominantly oviparous vertebrate taxa such as fishes, amphibians, reptiles and birds, do not show *in silico* potential CTP sequences. Only mammals do so, and even marsupials (e.g., kangaroo and opossum) show a putative CTP, although its sequence deviates significantly from CTP sequences found in the taxon Placentalia. Thus, genuine CTP-like sequences, either encrypted or expressed, correlate well with the presence of a placenta (Nakav et al. 2005). However, the glycosylation pattern of a potential CTP would not be beneficial in all species,

Revised version of manuscript MCE-D-08-00055

with respect to the mode of secretion, and we speculate that this is one reason why CG is

restricted to primate and equine species.

oset CGB possesses one N-glycosylation and two O-glycosylation sites, and therefore<br>the differences both with human CGB and LHB (see also Table1). Thus, mannoset CGB<br>ents a peptide like hCGB but with a glycosylation patter The marmoset differs from the human with regard to the glycosylation pattern of CGB, as marmoset CGB possesses one N-glycosylation and two O-glycosylation sites, and therefore there are differences both with human CGB and LHB (see also Table1). Thus, marmoset CGB represents a peptide like hCGB but with a glycosylation pattern intermediate between hLHB and hCGB (Simula et al. 1995; Amato et al. 1998). These peculiar biochemical properties lead to a disruption of clear CG pulses emanating from the pituitary, as the half-life of marmoset CG appears to be intermediate between that of human CG and human LH (Tannenbaum et al. 2007a; Tannenbaum et al. 2007b). The half-life of marmoset CG has not been investigated yet, nor has its placental glycosylation pattern, which might differ from that in the pituitary, since the exact composition of glycosylation depends upon the tissue-specific enzymes expressed. For instance, there is a hyperglycosylated form of human CG (hCG-H) that has a molecular weight of approximately 40 kDa. This form is produced in the first trimester of pregnancy (O'Connor et al. 1998). Its biological role is the cytotrophoblast invasion into the endometrium. Ectopic expression of hCG-H occurs in pregnancies with Down-syndrome and in some malignant tumours. The more malignant the cells were, the more CG was hyperglycosylated (Valmu et al. 2006; Cole and Khanlian 2007). Although the reason for ectopic CGB expression in tumours is still unclear, it could be demonstrated that CG can drive prostate carcinoma malignancy (Wu and Walker 2006).

#### **An extended model of primate CGB evolution**

 The recent findings on gonadotrophins led us to propose a revised and extended version of primate CGB gene evolution, covering the sequence of events such as duplication, mutation and acquisition of new functions (Fiddes and Talmadge 1984).

#### **ANUSCRI** CEPTED

#### Revised version of manuscript MCE-D-08-00055

#### **1. Duplication of the LHB gene**

 Approximately 55-50 million years ago (mya) the split of the tarsiers from the lineage of Simiiformes occurred. The anthropoid line encompasses New World monkeys (NWM), Old World monkeys (OWM), small apes, great apes and humans. The latter three groups form the taxon Hominoidea. As there is no CGB or second LHB gene in prosimians and tarsiers, the LHB gene duplication must have occurred before the NWM diverged from the OWM, approximately 55-35 mya (see also Fig. 2) (Bailey et al. 1991; Schrago and Russo 2003; Opazo et al. 2006; Schrago 2007). Because LHB and CGB have not only highly homologous coding sequences, but also highly homologous promoters, we assume that, after the LHB gene duplicated, both LHB copies were functionally expressed in the pituitary.

#### **2. The appearance of CGB**

monkeys (OWM), small apes, great apes and humans. The latter three groups form the<br>
Hominoidea. As there is no CGB or second LHB gene in prosimians and tarsiers, the<br>
gene duplication must have occurred before the NWM div 13 In one LHB gene copy a single nucleotide was deleted in the  $114<sup>th</sup>$  codon and led to a frame shift mutation. This slipped the last seven codons out of frame compared with LHB, and led to a read-through into the previous 3'UTR and an incorporation of this sequence into the 16 coding region. An insertion of the two nucleotides C and G in the  $138<sup>th</sup>$  codon further elongated the ORF, expanding the CGB subunit for 24 additional codons that encode the CTP. These frame shifts resulted in a new stop codon, TAA, in the previous 3'UTR. It is part of the polyadenylation signal and yields a very short 3'UTR of only 16 nucleotides (Talmadge et al. 1984). However, it is not possible to elucidate which event occurred first; the single nucleotide deletion or the insertion of the two bases.

 This concept of events for the mechanism of CGB gene evolution by LHB gene duplication and nucleotide modifications that led to the CTP, has already been proposed (Dayhoff 1976;

Fiddes and Goodman 1980; Fiddes and Talmadge 1984; Talmadge et al. 1984).

#### Revised version of manuscript MCE-D-08-00055

 Since CGB derived from LHB, including the promoter, and because CGB is still expressed at very low levels in human pituitaries, we speculate that CGB was still expressed in the pituitary in ancient primates at this stage of evolution.

#### **3. CG gains expression in the placenta**

gains expression in the placenta<br>
es in the promoter region of CGB led to a shift in expression from the pituitary to the<br>
tat (Hollenberg et al. 1994). The hCGB transcriptional start site relocated to 365 bp<br>
am of ATG, s Changes in the promoter region of CGB led to a shift in expression from the pituitary to the placenta (Hollenberg et al. 1994). The hCGB transcriptional start site relocated to 365 bp upstream of ATG, so that the 5' UTR of the mRNA comprises the region which serves as core promoter and transcriptional start site for LHB transcription. Human CGB transcription is mainly regulated by a TATA-less promoter, which is regulated by the transcription factors (TF) cAMP element binding protein (CREB), v-ETS erythroblastosis virus E26 oncogene homolog 2 (ETS2), activating protein 2 (AP2) and selective promoter factor 1 (SP1). Although each of these factors is expressed in several different tissues, they form together a trophoblast-specific combinatorial code for CGB gene transactivation (Jameson and Hollenberg 1993; Hollenberg et al. 1994; Pestell et al. 1994; Johnson and Jameson 1999; Johnson and Jameson 2000; Ghosh et al. 2003). At the blastula stage, CG is synthesised by the trophectoderm but not by the embryonic stem cells in the inner cell mass. The pluripotency marker Oct3/4 is expressed by the embryonic stem cells and acts as a transcription factor. Oct3/4 is able to bind to the promoters of the glycoprotein hormone alpha gene (GPHA) and CGB and thereby repress their transcription (Liu and Roberts 1996; Liu et al. 1997). Thus, in the placenta, CGB expression is controlled tissue-specifically by a completely different regulatory region and set of TFs compared with pituitary-expressed LHB.

 The glycoprotein α-subunit, which is necessary for full functional and biological activity of the gonadotrophins, also gained expression in the placenta, otherwise no functional CG holohormone could have formed. However, it remains to be clarified when exactly this happened and which of the two subunit genes was expressed first in the placenta.

#### Revised version of manuscript MCE-D-08-00055

#### **4. Differences of CG expression in pituitaries depend on taxonomic group**

 The LHB core promoter for mouse, rat, cattle and man comprises two transcription factor binding sites for steroidogenic factor 1 (SF1), two for early response growth protein 1 (EGR1), one for pituitary homeobox factor 1 (PITX1) and a TATA box. SF1 is involved in steroid synthesis in a wide spectrum of tissues (e.g. gonads and adrenals) and EGRr1 is expressed mainly during embryogenesis. PITX1 is a pituitary-specific transcription factor and all three factors together form a gonadotrophe-specific, tripartite regulatory complex for LHB activation (Kaiser et al. 2000; Jorgensen et al. 2004).

g sites for steroidogenic factor 1 (SF1), two for early response growth protein 1<br>
1), one for pituitary homeobox factor 1 (PITX1) and a TATA box. SF1 is involved in<br>
1 synthesis in a wide spectrum of tissues (e.g. gonads It was demonstrated that in human CGB only three nucleotide changes in transcription factor binding sites (TFBS) were sufficient to cause CGB down-regulation in the pituitary and reduced GnRH-responsiveness in comparison with LHB (Henke et al. 2007). In one TFBS for EGR1, a C-T transition and a G-C transversion occurred at positions -119 and -112, respectively, in relation to the start codon ATG. An A-G transition at -106 in the TFBS for PITX1 decreased its protein binding ability, and all three nucleotide changes together demonstrated synergistic effects in pituitary CGB down-regulation (Henke et al. 2007).

#### **5. CGB gene duplications in OWM and Hominoidea and a new LH/CG system in NWM**

 In the lineages of OWM and Hominoidea, the CGB gene itself underwent several duplications, leading to different copy numbers varying from three to six, as indicated in Table 3 and Fig. 2. This was reported first for humans and later investigated systematically for the order of primates (Boorstein et al. 1982; Talmadge et al. 1983; Maston and Ruvolo 2002). A recent study demonstrated that CGB1 is the most recent gene copy, as it evolved in the last common ancestor of African great apes (gorilla, chimpanzee and human), after the split from gibbons and orang utan, approximately 15 to 8 mya (Hallast et al. 2007).

Revised version of manuscript MCE-D-08-00055

tra (Muller et al. 2004).<br>
Solid for the complete branch of NWM because of the conserved CGB promoter<br>
Solid for the complete branch of NWM families (unpublished data from our laboratory), and the<br>
f LHB and presence of CG Originally both genes, LHB and CGB, must have been present in NWM. However, for yet unknown reasons, LHB became non-functional in the common marmoset and its loss was compensated for by CGB, which is a single copy gene, expressed in both pituitary and placenta (Muller et al. 2004). This is valid for the complete branch of NWM because of the conserved CGB promoter architecture across all four NWM families (unpublished data from our laboratory), and the lack of LHB and presence of CGB in pituitaries in other NWM species (Scammell et al. 2008). Further striking evidence comes from the similar exon architecture of the LHR type 2 (Gromoll et al. 2003; Gromoll et al. 2007). The LHR types 1 and 2 are G protein-coupled, seven transmembrane receptors. In OWM and Hominoidea, the LHR type 1 is able to bind and transduce signals from LH and CG equally well. However, in NWM, the LHR type 2 lacks exon 10 of 11 exons at the mRNA and protein levels, although it is present at the genomic level (Gromoll et al. 2003; Gromoll et al. 2007). Exon 10 encodes the extra-cellular hinge region of the receptor and its absence in LHR type 2 impairs LH signal transduction but not CG signalling (Muller et al. 2003). Thus, the LHR type 2 of the NWM is able to discriminate between LH and CG and transduces CG signalling only. Therefore Fig. 2 illustrates not only the CGB evolution in primates but also includes the LHR type 2 co-evolution and important dates of primate speciation.

#### **Summary and Conclusion**

 In this review we have discussed the evolutionary events that led to the rise of CGB in primates and were able to approximate the time lines during which each event took place. In brief, we propose the following sequence of events:

24 LHB gene duplication, emergence of the CTP, co-expression of CGB  $\&$  GPH $\alpha$  in the placenta and gain-of-function of CG in the placenta occurred approximately 55-35 mya, a relatively short time in evolutionary terms.

Revised version of manuscript MCE-D-08-00055

ation, which is specific for gorilla, chimpanzee and humans, occurred 15-8 mya.<br>
GB gain-of-function, its gene copy number variance in OWM and Hominoidea, and the<br>
pituitary CG expression in the NWMs demonstrate a remarkab The events specific for NWM, such as emergence of LHR type 2 and LHB's becoming a pseudogene, most likely happened 35-20 mya. In the lineage leading to OWM, small apes, great apes and the human, CGB was duplicated in the period 35-25 mya. A further 4 duplication, which is specific for gorilla, chimpanzee and humans, occurred 15-8 mya. The CGB gain-of-function, its gene copy number variance in OWM and Hominoidea, and the strong pituitary CG expression in the NWMs demonstrate a remarkable plasticity of the LH/CG hormone system not previously anticipated. Thus, the CGB gene cluster represents a hallmark of primate evolution and is an excellent example of how a few changes in duplicated DNA sequences can lead to a novel gene, characterised by a changed cellular expression, peptide extension as well as a novel physiological role. Therefore this gene cluster gives insight into currently ongoing evolution in primates, including the human.

#### **Acknowledgment**

 The authors thank Dr. J. Wistuba and Dr. T.G. Cooper for useful comments on the manuscript and language editing. Financial support was received from the German Research council (DFG), grant Gr 1541-8-1 and the Innovative Medical Research Fund of the University of Münster, grant SI 520601.

#### SCRII P CCEPTED U,

#### Revised version of manuscript MCE-D-08-00055

### 1 **Table 1**: **Main differences between human LH and CG.**

Ccepted Manuscript



2 The data are based on Pierce and Parsons 1981; Wehmann and Nisula 1981; Storring et al.

3 1988; le Cotonnec et al. 1998a, b, c ; Jablonka-Shariff et al. 2002; Bousfield and Ward 2006.

4 INSL3 = insulin-like growth factor 3.

#### **CCEPTED** SCRI P U, W Æ IJ

### Revised version of manuscript MCE-D-08-00055

### 1 **Table 2: Comparison of human gonadotrophin β-subunit transcripts**.

2



3

4 \* length of UTR sequence is given separately and therefore excluded from given exon size

 $A$ 

#### B CCEPTED R U, 0

### Revised version of manuscript MCE-D-08-00055

### 1 **Table 3: CGB copy numbers in different primate species.**



2

3 Data are based on (Maston and Ruvolo 2002; Hallast et al. 2007).

Accepted Manuscript

Revised version of manuscript MCE-D-08-00055

### **Figure 1: The structure of LH/CGB gene clusters and their genes.**

2 A schematic overview of the (A) organisation of the human LHB/CGB gene cluster on chromosome 19.3 with indications of the orientation of the genes and the intergenic distances, (B) the organisation of the rhesus macaque's LHB/CGB gene cluster and (C) details of the human genes themselves. The numbering of rhesus macaque CGB genes is in numerical order and does not imply special homology to the accordingly numbered human CGB genes. Black boxes indicate UTR, white boxes the exons (Ex) and solid vertical lines introns (Int) and promoter regions. The original CGB structure for CGB1/2 before deletions and insertions is indicated by grey shading.

Tendores Manuscript

Revised version of manuscript MCE-D-08-00055

### **Figure 2: Schematic evolution of CGB within the primates.**

Ccepted Manuscript

 This is a schematic overview of CGB evolution in primates. Prosimians and tarsiers possess only LHB but not CGB. Solid arrows indicate when an event happened, based on published time points. Dashed arrows indicate events for which the point of time is unknown. The period in which the event most likely happened is explained in the text. When more than one CGB copy number is present in a lineage, it is indicated by (n+2) or (n+3). Mya, million years ago; NWM, New World monkeys; OWM, Old World monkeys. Branching time points are according to Bailey et al. 1991; Kumar and Hedges 1998; Schrago and Russo 2003; Poux et al. 2006; Rhesus Macaque Genome Sequencing and Analysis Consortium et al. 2007 and Schrago 2007.

- 
- 

#### Revised version of manuscript MCE-D-08-00055

#### **References**

- Abramovich D. R., T. G. Baker, and P. Neal. 1974. Effect of human chorionic gonadotrophin on testosterone secretion by the foetal human testis in organ culture. J. Endocrinol. 60:179-185.
- Alfthan H., C. Haglund, J. Dabek, and U. Stenman. 1992. Concentrations of human choriogonadotropin, its beta-subunit, and the core fragment of the beta-subunit in serum and urine of men and nonpregnant women. Clin Chem 38:1981-1987.
- Allen W. 2001. Fetomaternal interactions and influences during equine pregnancy. Reproduction 121:513-527.
- Amato F., A. P. Simula, L. J. Gameau, and R. J. Norman. 1998. Expression, characterisation and immunoassay of recombinant marmoset chorionic gonadotrophin dimer and beta-subunit. J. Endocrinol. 159:141-151.
- Apter D., B. Cacciatore, H. Alfthan, and U. H. Stenman. 1989. Serum luteinizing hormone concentrations increase 100-fold in females from 7 years to adulthood, as measured by time-resolved immunofluorometric assay. J. Clin. Endocrinol. Metab. 68:53-57.
- Bailey W. J., D. H. Fitch, D. A. Tagle, J. Czelusniak, J. L. Slightom, and M. Goodman. 1991. Molecular evolution of the psi eta-globin gene locus: gibbon phylogeny and the hominoid slowdown. Mol. Biol. Evol. 8:155-184.
- n H., C. Haglund, J. Dabek, and U. Stenman. 1992. Concentrations of human oriogonadotropin, its beta-subunit, and the core fragment of the beta-subunit in serum of univer of the oriogonadotropin, its beta-subunit, and the Bellet D., V. Lazar, I. Bieche, V. Paradis, Y. Giovangrandi, P. Paterlini, R. Lidereau, P. Bedossa, J. M. Bidart, and M. Vidaud. 1997. Malignant transformation of nontrophoblastic cells is associated with the expression of chorionic gonadotropin beta genes normally transcribed in trophoblastic cells. Cancer Res. 57:516-523.
- Birken S., and R. E. Canfield. 1977. Isolation and amino acid sequence of COOH-terminal fragments from the beta subunit of human choriogonadotropin. J. Biol. Chem. 252:5386- 5392.
- Birken S., Y. Maydelman, M. A. Gawinowicz, A. Pound, Y. Liu, and A. S. Hartree. 1996. Isolation and characterization of human pituitary chorionic gonadotropin. Endocrinology 137:1402-1411.
- Bo M., and I. Boime. 1992. Identification of the transcriptionally active genes of the chorionic gonadotropin beta gene cluster in vivo. J. Biol. Chem. 267:3179-3184.
- Boorstein W. R., N. C. Vamvakopoulos, and J. C. Fiddes. 1982. Human chorionic gonadotropin beta-subunit is encoded by at least eight genes arranged in tandem and inverted pairs. Nature 300:419-422.
- Borkowski A., V. Puttaert, M. Gyling, C. Muquardt, and J. J. Body. 1984. Human chorionic gonadotropin-like substance in plasma of normal nonpregnant subjects and women with breast cancer. J. Clin. Endocrinol. Metab. 58:1171-1178.
- Bousfield G. R., W. K. Liu, H. Sugino, and D. N. Ward. 1987. Structural studies on equine glycoprotein hormones. Amino acid sequence of equine lutropin beta-subunit. J. Biol. Chem. 262:8610-8620.













#### **ANUSCRI** CCEPTED

- Policastro P., S. Daniels-McQueen, G. Carle, and I. Boime. 1986. A map of the hCG beta-LH beta gene cluster. J. Biol. Chem. 261:5907-5916.
- Poux C., P. Chevret, D. Huchon, W. W. de Jong, and E. J. Douzery. 2006. Arrival and diversification of caviomorph rodents and platyrrhine primates in South America. Syst. Biol. 55:228-244.
- Ramanujam L., W. X. Liao, A. C. Roy, S. C. Ng, and S. S. Ratnam. 1998. Molecular variants of luteinizing hormone in three populations of Southeast Asia. Hum. Hered. 48:232-234.
- 8 Ramanujam L. N., W. X. Liao, A. C. Roy, A. Loganath, H. H. Goh, and S. C. Ng. 1999. Association of molecular variants of luteinizing hormone with menstrual disorders. Clin. Endocrinol. (Oxf) 51:243-246.
- nijam I., W. X. Liao, A. C. Roy, S. C. Ng, and S. S. Ramam. 1998. Molecular variants<br>
l'uteinizing hormone in three populations of Southeast Asia. Ilum. Ilered. 48:232-234.<br>
nujam I., N., V. X. Liao, A. C. Roy, A. Loganath Rhesus Macaque Genome Sequencing and Analysis Consortium, R. A. Gibbs, J. Rogers, M. G. Katze, R. Bumgarner, G. M. Weinstock, E. R. Mardis, K. A. Remington, R. L. Strausberg, J. C. Venter, R. K. Wilson, M. A. Batzer, C. D. Bustamante, E. E. Eichler, M. W. Hahn, R. C. Hardison, K. D. Makova, W. Miller, A. Milosavljevic, R. E. Palermo, A. Siepel, J. M. Sikela, T. Attaway, S. Bell, K. E. Bernard, C. J. Buhay, M. N. Chandrabose, M. Dao, C. Davis, K. D. Delehaunty, Y. Ding, H. H. Dinh, S. Dugan-Rocha, L. A. Fulton, R. A. Gabisi, T. T. Garner, J. Godfrey, A. C. Hawes, J. Hernandez, S. Hines, M. Holder, J. Hume, S. N. Jhangiani, V. Joshi, Z. M. Khan, E. F. Kirkness, A. Cree, R. G. Fowler, S. Lee, L. R. Lewis, Z. Li, Y. S. Liu, S. M. Moore, D. Muzny, L. V. Nazareth, D. N. Ngo, G. O. Okwuonu, G. Pai, D. Parker, H. A. Paul, C. Pfannkoch, C. S. Pohl, Y. H. Rogers, S. J. Ruiz, A. Sabo, J. Santibanez, B. W. Schneider, S. M. Smith, E. Sodergren, A. F. Svatek, T. R. Utterback, S. Vattathil, W. Warren, C. S. White, A. T. Chinwalla, Y. Feng, A. L. Halpern, L. W. Hillier, X. Huang, P. Minx, J. O. Nelson, K. H. Pepin, X. Qin, G. G. Sutton, E. Venter, B. P. Walenz, J. W. Wallis, K. C. Worley, S. P. Yang, S. M. Jones, M. A. Marra, M. Rocchi, J. E. Schein, R. Baertsch, L. Clarke, M. Csuros, J. Glasscock, R. A. Harris, P. Havlak, A. R. Jackson, H. Jiang, Y. Liu, D. N. Messina, Y. Shen, H. X. Song, T. Wylie, L. Zhang, E. Birney, K. Han, M. K. Konkel, J. Lee, A. F. Smit, B. Ullmer, H. Wang, J. Xing, R. Burhans, Z. Cheng, J. E. Karro, J. Ma, B. Raney, X. She, M. J. Cox, J. P. Demuth, L. J. Dumas, S. G. Han, J. Hopkins, A. Karimpour-Fard, Y. H. Kim, J. R. Pollack, T. Vinar, C. Addo-Quaye, J. Degenhardt, A. Denby, M. J. Hubisz, A. Indap, C. Kosiol, B. T. Lahn, H. A. Lawson, A. Marklein, R. Nielsen, E. J. Vallender, A. G. Clark, B. Ferguson, R. D. Hernandez, K. Hirani, H. Kehrer-Sawatzki, J. Kolb, S. Patil, L. L. Pu, Y. Ren, D. G. Smith, D. A. Wheeler, I. Schenck, E. V. Ball, R. Chen, D. N. Cooper, B. Giardine, F. Hsu, W. J. Kent, A. Lesk, D. L. Nelson, W. E. O'brien, K. Prufer, P. D. Stenson, J. C. Wallace, H. Ke, X. M. Liu, P. Wang, A. P. Xiang, F. Yang, G. P. Barber, D. Haussler, D. Karolchik, A. D. Kern, R. M. Kuhn, K. E. Smith, and A. S. Zwieg. 2007. Evolutionary and biomedical insights from the rhesus macaque genome. Science 316:222-234. Roberts R. M. 2007. Interferon-tau, a Type 1 interferon involved in maternal recognition of
- pregnancy. Cytokine Growth Factor Rev. 18:403-408.
- Roberts R. M., Y. Chen, T. Ezashi, and A. M. Walker. 2008. Interferons and the maternal-conceptus dialog in mammals. Semin. Cell Dev. Biol. 19:170-177.
- Rull K., and M. Laan. 2005. Expression of beta-subunit of HCG genes during normal and failed pregnancy. Hum. Reprod. 20:3360-3368.

# COEPTED

#### Revised version of manuscript MCE-D-08-00055

nd chorionic gonadotropin  $\beta$ -subunits from New World squirrel monkey and owl<br>onkey. Gen. Comp. Endocrinol. 155:534-541.<br>So C. G. 2007. On the time scale of New World primate diversification. Am. J. Phys.<br>go C. G., and Rull K., P. Hallast, L. Uuskula, J. Jackson, M. Punab, A. Salumets, R. K. Campbell, and M. Laan. 2008. Fine-scale quantification of HCG beta gene transcription in human trophoblastic and non-malignant non-trophoblastic tissues. Mol. Hum. Reprod. 14:23-31. Scammell J. G., J. D. Funkhouser, F. S. Moyer, S. V. Gibson, and D. L. Willis. 2008. 5 Molecular cloning of pituitary glycoprotein  $α$ -subunit and follicle stimulating hormone and chorionic gonadotropin β-subunits from New World squirrel monkey and owl monkey. Gen. Comp. Endocrinol. 155:534-541. Schrago C. G. 2007. On the time scale of New World primate diversification. Am. J. Phys. Anthropol. 132:344-354. Schrago C. G., and C. A. M. Russo. 2003. Timing the Origin of New World Monkeys. Mol Biol Evol 20:1620-1625. Sherman G. B., M. W. Wolfe, T. A. Farmerie, C. M. Clay, D. S. Threadgill, D. C. Sharp, and J. H. Nilson. 1992. A single gene encodes the beta-subunits of equine luteinizing 14 hormone and chorionic gonadotropin. Mol. Endocrinol. 6:951-959. Simula A. P., F. Amato, R. Faast, A. Lopata, J. Berka, and R. J. Norman. 1995. Luteinizing hormone/chorionic gonadotropin bioactivity in the common marmoset (Callithrix jacchus) is due to a chorionic gonadotropin molecule with a structure intermediate between human chorionic gonadotropin and human luteinizing hormone. Biol. Reprod. 53:380-389. Spincemaille J., R. Bouters, M. Vandeplassche, and P. Bonte. 1975. Some aspects of endometrial cup formation and PMSG production. J Reprod Fert 23, Supplement 415- 418. Starka L., M. Hill, R. Hampl, and I. T. Huhtaniemi. 1999. Genetic variant of luteinizing hormone in Czech Republic. Endocr. Regul. 33:103-108. Stenman U. H., H. Alfthan, T. Ranta, E. Vartiainen, J. Jalkanen, and M. Seppala. 1987. Serum levels of human chorionic gonadotropin in nonpregnant women and men are modulated by gonadotropin-releasing hormone and sex steroids. J. Clin. Endocrinol. Metab. 64:730- 736. Storring P. L., S. A. Khan, Y. G. Mistry, and E. Diczfalusy. 1988. Plasma clearance in the rat of the LH bioactivity of two human LH standards of differing molecular composition. J. Endocrinol. 119:327-324. Sugino H., G. R. Bousfield, W. T. Moore Jr, and D. N. Ward. 1987. Structural studies on equine glycoprotein hormones. Amino acid sequence of equine chorionic gonadotropin beta-subunit. J. Biol. Chem. 262:8603-8609. Talmadge K., N. C. Vamvakopoulos, and J. C. Fiddes. 1984. Evolution of the genes for the beta subunits of human chorionic gonadotropin and luteinizing hormone. Nature 307:37- 40. Talmadge K., W. R. Boorstein, and J. C. Fiddes. 1983. The human genome contains seven genes for the beta-subunit of chorionic gonadotropin but only one gene for the beta-subunit of luteinizing hormone. DNA 2:281-289.

## Revised version of manuscript MCE-D-08-00055





Page 32 of 33

