Cloning and Characterisation of Orthologous Genes Encoding Growth Hormone in Egyptian Holstein Crossbred and Ossimi Sheep

Waleid Mohamed El-Sayed Shakweer*  
Animal Production Dept.  
National Research Centre  
Dokki, Giza, Egypt  
waleid_mohamed2000@yahoo.com

Yassein Mohamed Hafez  
Animal Production Dept.,  
Cairo University Research Park(CURP)  
Faculty of Agriculture, Cairo University,  
Giza, Egypt,  
yasseinhafez@agr.cu.edu.eg

Marwa Naguib Mohamed Esmaail Sanad  
Genetic and Cytology Dept.  
National Research Centre  
Dokki, Giza, Egypt  
drmarwanaguib24@gmail.com

Ashraf Abd El-Halim El-Sayed  
Animal Production Dept.,  
Cairo University Research Park(CURP)  
Faculty of Agriculture  
Cairo University, Giza, Egypt  
ashrafah99@agr.cu.edu.eg

Ibrahim Mohamed Awadalla  
Animal Production Dept.  
National Research Centre  
Dokki, Giza, Egypt  
ibrahimawadalla@hotmail.com

Mamdouh Ibrahim Mohamed  
Animal Production Dept.  
National Research Centre  
Dokki, Giza, Egypt  
shalaby1956@yahoo.com

Abstract: This study aims to investigate the compatibility between growth hormone (GH) cDNA sequence in Egyptian cows and sheep as an introduction to produce bovine GH-Transgenic sheep using sperm mediated gene transfer (SMGT) technique to improve their productivity. Our previous in vivo study indicates a significant increase in the concentration of GH mediator (IGF-1) in blood, and daily growth rate, when Egyptian rams were injected with bovine somatotropin. The GH gene of Egyptian Holstein crossbred (EH_GH) and Ossimi sheep breed (Os_GH) has been characterized in this study. The mRNA of pituitary gland has been isolated and then the cDNA was cloned into pTZ57R/T cloning vector. The complete coding domain sequences (CDSs) of EH_GH (AC: KP221576) AND Os_GH (Ac: KP221575) were detected and annotated comprehensively. The sequenced GH genes varied slightly along gene length; EH_GH was 740 bp instead of 690 bp in Os_GH, but with similar gene structure (5 exons and 4 introns). High homology (98%) in GH protein length and structure was observed; the predicted open reading frames (ORF) length was 217 aa and 216 aa of EH_GH and Os_GH, respectively, which included only four different amino acids in the predicted ORFs. A gene synteny was developed of the homologous bovine and ovine GH1 genes Ensemble annotated genome database. The results showed strong conserved structure and function, indicated that to GH gene of Egyptian Holstein crossbred and Ossimi breed are orthologous genes. We assume in this study that the novel genetic information about the bovine and ovine GH in Egyptian breeds can support; further success possibilities of producing the bovine GH-transgenic sheep to improve the growth performance and milk production.

Keywords: Growth hormone (GH), Orthologous genes, Egyptian Holstein crossbred, Ossimi breed.

1. INTRODUCTION

Sheep is one of the most important domestic animals reared in Egypt. There are almost 5.5 million head that contributed about 6% of the total red meat that produced in Egypt [1]. The cows' population in Egypt is estimated about 4.53 million heads that produce 2.90 and 0.41 million tons of milk and meat, respectively [2]. In the last three decades, Egyptian farms adopted crosses between the Holstein breed and the Egyptian Baladi breeds to improve their productivity by producing Egyptian Holstein crossbred adapted under the Egyptian environment. Due to a the
world food gap in the world, a lot of scientists have been turned on to gene transfer techniques as a quick tool to increase livestock production. This study aims to investigate the compatibility between the growth hormone cDNA sequence in Egyptian cows and sheep as an introduction to produce bovine GH-Transgenic sheep using sperm mediated gene transfer (SMGT) to improve their productivity. The candidate gene is growth hormone that spans approximately 1.8 kilobases and consists of five exons interrupted by four introns [3]. Growth Hormone gene is located at chromosome 19 in *Bos taurus* genome, while at chromosome 11 in *Ovis aries* genome [4] and [5]. The growth hormone is a single-chain polypeptide synthesised and secreted by the pituitary gland under hypothalamic control and regulates the body growth, cell reproduction and regeneration in cattle and sheep [6]. Growth hormone affects the partitioning of nutrients among tissues in sheep and cattle, increasing bone growth and milk production and decreasing fatness [7]. It belongs to the same hormonal family as prolactin (PRL) and placental lactogen (PL), all of them sharing certain biological, immunological and structural features; they are considered to be derived from duplication of a single ancestral gene [8] and [9]. Plasma concentration of GH is increased in cows genetically selected for high milk production [10] and in sheep selected for low levels of fatness [11], indicating that GH is genetically correlated with these traits.

In our previous study, recombinant bovine somatotropin (100 mg rbST/14 days) had been injected in Egyptian Rahmani lambs. The average daily gain has been increased significantly (8.43%, P<0.05) in rbST treated rams comparing to the control. The feed conversion did not differ significantly in rbST treated rams compared to the control that may be gives us an indicator that the treated group utilizes the feed intake in a good way to increase the average daily gain [12]. Based on that aforementioned, our aim was cloning and characterization of growth hormone gene in both Egyptian Holstein crossbred and Ossimi sheep breed. The functional bioinformatic analysis can guide the research accurately to predict for the success possibilities to introduce and express bovine GH gene into transgenic sheep for enhancing their productivity.

2. MATERIALS AND METHODS

2.1. cDNA Synthesis

Total RNA has been isolated from pituitary gland (50 mg) of Egyptian Holstein crossbred and Ossimi breed using Booze reagent kit (Bioflux®). Total RNA was reverse transcribed with Oligo dT18 using RevertAid First Strand cDNA Synthesis Kit (Fermentas, Canada). The EH_GH primer was designed using custom primers-OligoPerfect™ designer tool (https://tools. lifetechnologies.com/content.cfm?pageid=9716&icid=fr-oligo-6?CID=fl-oligoperfect) based on NCBI database (Ac: M27325.1, 783bp). The EH_GH primer sequence was EH_GH-F 5'-GAGCTCCAGGGTCTCGTGGACAGC-3' and EH_GH-R 5'-CCGCGGTCGCATGCAATTTCCTCAT-3'. While, the Os_GH primer was published on NCBI (Ac: EU935861.1), Os_GH-F 5'-GCTCACCAGCTATGATGGCTGACGCTG-3' and Os_GH-R 5'-TGCCAATAGGAAGGCCGACT-3'.

2.2. Gene Cloning

The PCR reaction mixture (10µl) was prepared for each gene as follow; [1µl of 10X buffer containing MgCl2 (25mM), 1µl of dNTBase mixture (200 µM), 1µl of F-primer, 1µl of R-primer, 0.2 µl of Taq polymerase (1U/µl), 1µl of both EH_GH/Os_GH cDNA and water nuclease-free up to 10µl] [13]. The amplification reaction was carried out in a Thermocycler PCT-100 (MJ Research, USA). The PCR conditions of EH_GH-cDNA was denaturation for 1 min at 94°C, annealing for 2 min at 57°C, and extension for 3 min at 72°C for 29 cycles, and final extension at 72°C for 7 min, while for Os_GH-cDNA, denaturation for 1 min at 94°C, annealing for 2 min at 63°C, and extension for 3 min at 72°C for 29 cycles. A final extension was 72°C for 7 min. The PCR products were visualized on a 1% agarose gel stained with Ethidium bromide in 1x TAE buffer at 100 v for 60 min with 10 Kb BioLabs ladder, (New England, UK). The image of the gel band was taken using Bio-Rad Gel Doc XR. The detected EH_GH and Os_GH amplicons were extracted from the gel using GeneElute™ Gel Extraction Kit (Fermentas, USA), then cloned into pTZ57R/T vector (Fig.1). The vector construction was designed according to the cloning kit procedure (Fermentas, USA) as follow; 3µl of pTZ57R/T vector, 2.5µl of both EH_GH/Os_GH cDNA, 6µl of 10x ligase buffer, 1µl T4 DNA ligase and 17.5µl sterile water in total 30 µl reaction mixture, incubated at 22°C/1h for ligation. The ligated product was transformed into DH5α cells (Fermentas, USA). The constructed plasmids (EH_GH and Os_GH-pTZ57R/T) were extracted
Cloning and Characterisation of Orthologous Genes Encoding Growth Hormone in Egyptian Holstein Crossbred and Ossimi Sheep

using GeneJET plasmid Miniprep kit, and sequenced. The efficiency of transformation was determined using the polymerase chain reaction.

Fig1. Map of the pTZ57R/T cloning vector.

2.3. Bioinformatics Analysis

The consensus sequence of cloned EH_GH and Os_GH cDNA were created using CodonCode Aligner V4.2.7 software. The SNPs among EH_GH and Os_GH cDNA sequence have been done using the pairwise alignment via EMBOSS Needle tool (http://www.ebi.ac.uk/Tools/psa/emboss_needle/). The transcription promoter has been predicted using Neural Network Promoter Prediction via http://www.fruitfly.org/seq_tools/promoter.html [14]. The open reading frames (ORF) of EH_GH and Os_GH cDNA sequences were predicted using http://web.expasy.org/translate/ [15]. Predicted protein codons relative to the nucleotide sequence were done using translated tool via http://www.fr33.net/translator.php. The predicted cleavage site in signal peptide sequence was done using SignalP 4.1 Server tool http://www.cbs.dtu.dk/services/SignalP/ [16]. The conserved domain of EH_GH and Os_GH cDNA sequences were detected using NCBI conserved domains tool via http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi [17]. The motif of EH_GH and Os_GH cDNA was predicted using search motif library tool at http://www genie.jp/tools/motif/. The cysteine state and disulfide bonds in the EH_GH and Os_GH protein sequences were carried out using cysteine and disulfide bond prediction tool at http://clavius.bc.edu/~clotelab/DiANNA/ [18]. Pairwise alignment was done between EH_GH and Os_GH protein sequences using EMBOSS Needle tool (http://www.ebi.ac.uk/Tools/psa/emboss_needle/). Also, the multiple alignments of EH_GH and Os_GH protein sequences vs. NCBI protein database have been carried out using Bioedit 7.2.5 software [19]. Phylogenetic tree of EH_GH and Os_GH protein sequences was conducted using MEGA version 6 [20]. The synteny location of GH in ovine chromosome (Chr11) vs. bovine chromosome (Chr19) has been achieved using annotation and prediction tool available at http://www.ensembl.org.

3. RESULTS AND DISCUSSION

3.1. Gene Cloning of GH

The cDNA and translated protein sequences of EH_GH and Os_GH had been published at Genebank-NCBI database with accession number KP221576 and KP221575, respectively. The electrophoretic band of EH_GH was 740 bp and was 690 bp for Os_GH (Fig. 2). In Fig. 3, five uncut plasmid forms were showed on one sample of EH_GH-TZ57R/T, whereas the Os_GH-TZ57R/T had three to four forms. About five forms of uncut plasmid on the electrophoretic gel based on its form and molecular weight are; 1) Nicked open-circular 2) Relaxed circular 3)
Linear, 4) Supercoiled, 5) Supercoiled denature [21]. The electrophoresis results of EH_GH-TZ57R/T and Os_GH-TZ57R/T are shown in Fig. 4.

**Fig2.** Gel electrophoresis of cDNA band of EH_GH and Os_GH.

**Fig3.** Photograph of Holstein and Ossimi GH-TZ57R/T uncut plasmid electrophoresis.

**Fig4.** Photograph of EH_GH-TZ57R/T and Os_GH-TZ57R/T PCR plasmid electrophoresis.

**Fig5.** Photograph of EH_GH-TZ57R/T and Os_GH-TZ57R/T cut plasmid.
3.2. Growth Hormone cDNA SNPs

The pairwise alignment (Fig. 8) has been shown that there are ten nonsynonymous alleles substitutions (small letter) which yield a new codon that encodes a different amino acids in EH\_GH cDNA sequence vs. Os\_GH cDNA sequence (Table 1); gct for Alanine (Ala) at 34, 35 and 36, ccc for Proline (Pro) at 43, 44 and 45, Cgc for Arginine (Arg) at 650 and 651, gcc for Alanine (Ala) at 688, 689 and 690, atC for Isoleucine (Ile) at 691 and 692, t Gt for cysteine (Cys) at 694 and 696, tGt for Cysteine (Cys) at 697 and 698, Ttg for Leucine (Leu) at 701 and 702, ccc for Proline (Pro) at 703, 704 and 705 and ccc for Proline (Pro) at 709, 710 and 711 in EH\_GH cDNA sequence vs. aag for Lysine (Lys) at 12, 13 and 14, aCC for Threonine (Thr) at 21, Cag for Glutamine (Glu) at 628 and 629, aaa for Lysine (Lys) at 666, 667 and 668, cgC for Arginine (Arg) at 669 and 670, cGg for Arginine (Arg) at 772 and 774, tT for Leucine (Leu) at 675 and 676, Tgc for Cysteine (Cys) at 780, tgT for Cysteine (Cys) at 697 and 698, Ttg for Leucine (Leu) at 701 and 702, ccc for Proline (Pro) at 703, 704 and 705 and ccc for Proline (Pro) at 709, 710 and 711 in Os\_GH cDNA sequence.

**Fig6.** The growth hormone cDNA sequence and the encoded amino acid residues in Egyptian Holstein crossbred. 1 Signal protein sequence (Red, 1...27 aa), 2 Conserved domain of predicted residues sequence (Yellow: 36...215 aa), 3 First motif Somatotropin_1 (Green: 79...112, 192 related sequences and 22 related structures), 4 Second motif Somatotropin_2 (Green: 190...207, 194 related sequences and 22 related structures), 5 Predicted promoter (Purple, start 603 bp...end 653 bp and score 0.85), The conserved Cysteine residues are illustrated by boxes.
The growth hormone cDNA sequence, predicted amino acids sequences and protein features in Ossimi breed.  

**Fig7.** The growth hormone cDNA sequence, predicted amino acids sequences and protein features in Ossimi breed.  

<table>
<thead>
<tr>
<th>Signal protein sequence (Red: 1...26 aa)</th>
<th>Conserved domin of predicted residues sequence (Yellow: 35...214 aa)</th>
<th>First motif Somatotropin_1 (Green: 78...111, 192 related sequences and 22 related structures)</th>
<th>Second motif Somatotropin_2 (Green: 189...206, 194 related sequences and 22 related structures)</th>
<th>Predicted promoter (Purple: start 581bp... end 631bp and score 0.85). The conserved Cysteine residues are illustrated by boxes.</th>
</tr>
</thead>
</table>

The conserved residues are illustrated by boxes.

**Fig8.** The pair wise alignment of EH and Os_GH cDNA sequences.

On the other hand, there are five synonymous allel’s substitutions (small letter) which yield a new codon that encodes the same amino acids TGt for Cysteine (Cys) at 672, TTc for Phenylalanine (Phe) at 678, TAg Stop codon at 681, CCa for Proline (Pro) at 687 and cTc Leucine (Leu) at 706 and 708, in EH_GH cDNA sequence vs. TGc for Cysteine (Cys) at 650, TTt for Phenylalanine (Phe) at 656, TAA Stop codon at 659, CCc for Proline (Pro) at 665 and tTg Leucine (Leu) at 684 and 686 in Os_GH cDNA sequence. That single nucleotide polymorphism can be used to develop SNPs markers for each breed.

### 3.3. The Prediction of Growth Hormone Promoter

The promoter is the most important regulatory region that regulates the very first step of gene expression [22]. The promoter sequence of EH_GH cDNA was AAGGACCTGCATAAGACGGAGACGTACCTGAGGGTTGGC (50bp), contains start codon (ATG) and spanning between 603bp and 653bp with prediction score 0.85 (Fig. 6). Whereas the promoter of Os_GH was GAAGGCAATCCCTGAGGCTGCTAGGTGCCGCC (38bp), contains start codon (ATG) and spanned between 581bp to 631bp with 0.85 prediction score (Fig. 7). The identity score between both promoters was 57.89 considering length differences.

### 3.4. Protein Structure

#### 3.4.1. Signal Peptide Sequence

The signal peptide of EH_GH sequence ranged from 1 to 27 residues and from 1 to 26 residues in Os_GH cDNA (Fig. 6 and 7, respectively). The predicted cleavage site was located between 26 and 27 amino acids in EH_GH protein and between 25 and 26 amino acids in Os_GH protein (Fig. 9). The signal peptide has been defined as a sequence generally found at the N-terminal end of most secreted and cell-surface protein precursors [23]. The signal sequence that bound at the amino terminus of the emerging protein “flag”, which transport the mechanism of the cell to prompt them as to where the emerging protein should go. The flags recognize by signal recognition particle (SRP) locate freely in the cytoplasm or attached to the ribosome, the SRP interacts with the signal sequence and leads to the transient arrest of translation. That may indicate that the isolated GH cDNA of both breed could be expressed in transfected cell [24].
Cloning and Characterisation of Orthologous Genes Encoding Growth Hormone in Egyptian Holstein Crossbred and Ossimi Sheep

Table 1. The modified nucleotides (SNPs) and encoded amino acids in EH_GH cDNA vs. Os_GH cDNA sequences.

<table>
<thead>
<tr>
<th>Nucleotides</th>
<th>Location</th>
<th>Residue</th>
<th>Nucleotides</th>
<th>Location</th>
<th>Residue</th>
</tr>
</thead>
<tbody>
<tr>
<td>gct</td>
<td>34, 35 and 36</td>
<td>Alanine</td>
<td>aag</td>
<td>12, 13 and 14</td>
<td>Lysine</td>
</tr>
<tr>
<td>ccc</td>
<td>43, 44 and 45</td>
<td>Proline</td>
<td>aCC</td>
<td>21</td>
<td>Threonine</td>
</tr>
<tr>
<td>Cgc</td>
<td>650 and 651</td>
<td>Arginine</td>
<td>Cag</td>
<td>628 and 629</td>
<td>Glutamine</td>
</tr>
<tr>
<td>TGt</td>
<td>672</td>
<td>Cysteine</td>
<td>TGc</td>
<td>650</td>
<td>Cysteine</td>
</tr>
<tr>
<td>TTc</td>
<td>678</td>
<td>Phenylalanine</td>
<td>TTt</td>
<td>656</td>
<td>Phenylalanine</td>
</tr>
<tr>
<td>TAg</td>
<td>681</td>
<td>Stop</td>
<td>TAa</td>
<td>659</td>
<td>Stop</td>
</tr>
<tr>
<td>CCa</td>
<td>687</td>
<td>Proline</td>
<td>CCc</td>
<td>665</td>
<td>Proline</td>
</tr>
<tr>
<td>gcc</td>
<td>668, 689 and 690</td>
<td>Alanine</td>
<td>aaa</td>
<td>666, 667 and 668</td>
<td>Lysine</td>
</tr>
<tr>
<td>atC</td>
<td>678 and 692</td>
<td>Isoleucine</td>
<td>cGC</td>
<td>669 and 670</td>
<td>Arginine</td>
</tr>
<tr>
<td>tGt</td>
<td>694 and 696</td>
<td>Cysteine</td>
<td>cGg</td>
<td>672 and 674</td>
<td>Arginine</td>
</tr>
<tr>
<td>tgT</td>
<td>697 and 698</td>
<td>Cysteine</td>
<td>cTt</td>
<td>675 and 676</td>
<td>Leucine</td>
</tr>
<tr>
<td>Ttg</td>
<td>701 and 702</td>
<td>Leucine</td>
<td>tgg</td>
<td>679 and 680</td>
<td>Cysteine</td>
</tr>
<tr>
<td>ccc</td>
<td>703, 7024 and 705</td>
<td>Proline</td>
<td>tgg</td>
<td>681, 682 and 683</td>
<td>Tryptophan</td>
</tr>
<tr>
<td>cTc</td>
<td>706 and 708</td>
<td>Leucine</td>
<td>tTg</td>
<td>684 and 686</td>
<td>Leucine</td>
</tr>
<tr>
<td>ccc</td>
<td>709, 710 and 711</td>
<td>Proline</td>
<td>egg</td>
<td>687, 688 and 689</td>
<td>Glycine</td>
</tr>
<tr>
<td>C……..C</td>
<td>1 to 22</td>
<td>No complement nucleotide</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G………A</td>
<td>713 to 740</td>
<td>R……..C</td>
<td>No complement nucleotide</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* Lower case letters have been indicated for SNPs, ** The same amino acids with different encoding sequence

3.4.2. Conserved Domain and Motifs

Fig 9. The predicted cleavage site of signal peptide in Egyptian Holstein crossbred (A) and Ossimi breed (B). C-score: Signal peptide cleavage sites. S-score: Signal peptide positions within signal peptides from positions in the mature part of the proteins and from proteins without signal peptides. Y-score: Better cleavage site prediction than the raw C-score alone. Mean S: The average S-score of the possible reticulum the SRP detaches and this leads to the continuation of protein translation, and eventually to the signal peptide (from position 1 to the position immediately before the maximal Y-score). D-score: Discriminate signal peptides from non-signal peptides. For non-secretory proteins all the scores represented in the SignalP output should ideally be very low (close to the negative target value of 0.1) [15].

Highly conserved domain of predicted growth hormone protein has been detected in EH_GH and Os_GH (Fig. 10). Detecting the significant conserved protein domains is often required for basic cellular function, stability or reproduction. Conservation of protein structures is indicated by the presence of functionally equivalent, though not necessarily identical, amino acid residues and structures between analogous parts of proteins [25]. The predicted conserved domain hit somatotropin like hormone and specific family with growth hormone like superfamily; Interval 35 to 214 aa, 1.40e-104 E-value for Os_GH protein and Interval 36 to 215 aa, 5.30e-106 E-value for EH_GH protein. Protein motifs are signatures of protein families and can often be used as tools...
for the prediction of protein function [26]. We detected two common motifs in both EH_GH and Os_GH cDNA but with different spanning (Fig. 6 and 7); Somatotropin_1 (CFSETIPAPTGEAQQKSDLLELRISLLLIQSW) from 79 to 112 aa for EH_GH cDNA and from 78 to 111 aa for Os_GH cDNA. Somatotropin_2 (CFRKDLHKTETYL RVM KC) from 190 to 207 aa for EH_GH cDNA and from 189 to 206 aa for Os_GH cDNA which reflects same function and characteristics.

A comprehensive summary of EH_GH and Os_GH annotation has been covered in Table (2).

![Conserved domain of EH_GH and Os_GH protein sequences](image.png)

**Table 2.** Comprehensive summary of EH_GH and Os_GH sequences annotation.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>EH_GH cDNA</th>
<th>Os_GH cDNA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chromosome location</td>
<td>19</td>
<td>11</td>
</tr>
<tr>
<td>Exon number</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Exon protein position</td>
<td>1..31..150..175..205</td>
<td>1..30..149..175..204</td>
</tr>
<tr>
<td>Intron number</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Length of mRNA</td>
<td>740 bp</td>
<td>690 bp</td>
</tr>
<tr>
<td>Predicted promoter</td>
<td>Start 603, End 653, Score 0.85</td>
<td>Start 581, End 631, Score 0.85</td>
</tr>
<tr>
<td>Motifs</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Somatotropin_1</td>
<td>From 79 to 112</td>
<td>From 78 to 111</td>
</tr>
<tr>
<td>Somatotropin_2</td>
<td>From 190 to 207</td>
<td>From 189 to 206</td>
</tr>
<tr>
<td>Molecular weight</td>
<td>24558 Da</td>
<td>24460 Da</td>
</tr>
<tr>
<td>Number of residues</td>
<td>217</td>
<td>216</td>
</tr>
<tr>
<td>Disulfide bonds</td>
<td>79 - 190</td>
<td>78 - 189</td>
</tr>
<tr>
<td></td>
<td>207 - 215</td>
<td>206 - 214</td>
</tr>
<tr>
<td>Cysteine position</td>
<td>18</td>
<td>17</td>
</tr>
<tr>
<td></td>
<td>79</td>
<td>78</td>
</tr>
<tr>
<td></td>
<td>190</td>
<td>189</td>
</tr>
<tr>
<td></td>
<td>207</td>
<td>206</td>
</tr>
<tr>
<td></td>
<td>215</td>
<td>214</td>
</tr>
<tr>
<td>Physical characteristics of growth hormone protein</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Length of protein</td>
<td>217 aa</td>
<td>216 aa</td>
</tr>
<tr>
<td>Chain peptide</td>
<td>27-217 (190 aa)</td>
<td>26-216 (190 aa)</td>
</tr>
<tr>
<td>Open reading frame</td>
<td>1-217</td>
<td>1-216</td>
</tr>
<tr>
<td>Signal peptide</td>
<td>1-27</td>
<td>1-26</td>
</tr>
<tr>
<td>Conserved domain</td>
<td>36-215</td>
<td>35-214</td>
</tr>
</tbody>
</table>

**3.4.3. Cysteine Bridge and Disulfide Bounds**

Five conserved cysteine (Cys) are founded in both EH_GH and Os_GH protein sequence at the following positions; (Cys\textsuperscript{18}, Cys\textsuperscript{79}, Cys\textsuperscript{190}, Cys\textsuperscript{207} and Cys\textsuperscript{215}) and (Cys\textsuperscript{17}, Cys\textsuperscript{78}, Cys\textsuperscript{189}, Cys\textsuperscript{206} and Cys\textsuperscript{204}), respectively (Fig 6 and 7). Two sulfide bounds are predicted between 79 and 190 aa
Cloning and Characterisation of Orthologous Genes Encoding Growth Hormone in Egyptian Holstein Crossbred and Ossimi Sheep

(99%) and between 207 and 215 aa (98%) in EH_GH protein and between at 78 and 189 aa (99%) and between 206 and 214 aa (93%) in Os_GH protein. The disulfide bonds play an important role in the folding and stability of some proteins, usually proteins secreted to the extracellular medium [27]. Four cysteine residues from five have been participated in the formation of two disulfide bonds between the thiol groups of cysteine residues by the process of oxidative folding. The disulfide bond stabilizes the folded form of a protein in several ways; 1) It holds two portions of the protein together, biasing the protein towards the folded topology. 2) The disulfide bond may form the nucleus of a hydrophobic core of the folded protein [28]. One possible role of this fifth Cysteine residue might be the formation of oligomeric complexes, which might affects the appropriate re-folding of growth hormone protein [29]. In the intracellular environment, cysteines can still play a key structural role. Their sulphhydryl side chain is excellent for binding to metals, such as zinc, meaning that cysteines are very common in metal binding motifs such as zinc fingers [30].

3.4.4. Pair wise alignment of Open Reading Frame

The pairwise alignment of EH_GH protein vs. Os_GH protein (Fig. 11) has been indicated that the protein sequences in both breeds are highly identical residues (98%) except in four positions (Table 3). The EH_GH protein sequence was started with two start codon (MM), while the Os_GH protein sequence was started with unusual single start codon (M). The modified residues of EH_GH protein sequence were Ala3, Pro6 (signal peptide) and Arg208 vs. Lys2, Thr2 (signal peptide) and Gln207 in Os_GH protein sequence. The differences in residues substitution in signal peptide did not affect translocation efficiency and cleavage site position and this conclusion is confirmed based on the results of highly values of C, Y, S and D in signalP chart (Fig. 9). But the difference in the N-terminus Methionine can be higher growth hormone transcription in EH_GH, and lower expression in Os_GH, refer to the active and inactive forms of GH [31].

![Fig11. The pairwise alignment between EH_GH and Os_GH protein sequences](image)

The Ala and Lys can substitute each other because both belong to the same amino acid group (small) as well as Arg vs. Gln both are polar residues. Although, the Pro is a small residue and Thr is a polar residue, in addition to a common role of Pro and Thr in facilitating the intracellular signal transduction but they can substitute without negative effect on protein function [15], [32] and [33]. It may worth to look into the unique residue in Os_GH (Lys) in exon2, which was associated with a mutation (AAG) vs. (Ala, GCT), which is a common codon among all the aligned proteins, and see if the mutation affect the receptor efficiency especially with the missing initial Met in Ossimi. In respect to the important role of Lys in structure and locate on the outside of the protein, while Ala rarely involved in protein function directly but can play a role in substrate recognition or specificity [33]. There is about 97.5% homology in the coding regions...
between bovine GH and ovine GH genes [34]. The GH homology can supports and explain the assumption the expression possibility of bovine growth hormone in injected sheep breed as mentioned previously in rams [12], [35], [36], [37] and [38].

3.4.5. Multiple Alignment

Gene families arisen by gene duplication and natural selection [39]. The multiple alignments are an essential study to understand the evolution among the bovine and ovine growth hormone precursors. The multiple alignment of EH_GH and Os_GH against other GH complete CDS (Genebank- NCBI database), and the main published GH1 of bovine and ovine (Ensemble genomes database) are shown in Fig. 12. There are three pakistanian ovine breeds (Latti, Lohi and Afghani) and four bovine breeds (Indian Zeubo, Iranian cattle, Pakistani Sahiwal cattle and Chinese cattle) have partial sequences (four exons) while in Zebu cattle contains five exons but the first exon is very short contains only four residues. In most mammals, GH sequence is strongly conserved, but differences in the biological and receptor-binding properties are due to the species-specificity of receptor-binding [40]. Interestingly, the results of multiple alignment have been illustrated that there is a distinct mutation in growth hormone protein sequence where the conserved Gly\textsuperscript{156} has been substituted with Val\textsuperscript{156} in following breeds; Ovis aries ref synteny, Pakistani Kamori cattle, Pakistani Latti ovine, Chines Tibetan ovine, Pakistani lohi ovine, Chines Kazakh ovine, Swiss Saanen goat, Portugal Serra da Estrela ovine, Chines Lezhi Black goat, Chines Tibet goat and Chines ovine compared to the other breeds (Table 3). Normally, if the conserved Glycine has been changed to any other amino acid, this change could have a drastic impact on protein function [33]. Therefore, it could be concluded that it is possible to substitute conserved Gly and get a functional growth hormone protein, this mutation may be considered as a specific genetic marker related to specific breed in specific region especially in China and Pakistan based on the investigated breeds. The Ossimi breed contains distinct modified residues; Lys\textsuperscript{2} and Thr\textsuperscript{5} vs. Ala\textsuperscript{3} and Pro\textsuperscript{5} in all investigated breeds, respectively (Table 3). The substitution of Pro\textsuperscript{5} to Thr and Ala\textsuperscript{5} to Lys\textsuperscript{2} are located out open reading frame therefore did not affect the protein function [32], [33] and [41]. In the present study, the Gly\textsuperscript{53} residue was presented in all investigated breeds vs. Ser\textsuperscript{55} residue in Afghani sheep breed. Because the both Gly and Ser are tiny and small residues and the Gly is not in the conserved domain sequence; therefore the substitution is functional [32]. The substitution of Phe\textsuperscript{123}, 217 and Leu\textsuperscript{188} in all tested breeds with Leu\textsuperscript{123}, 217 and Phe\textsuperscript{188} in Lezhi Black goat, Dhanni x Friesian cattle crossbred and Iranian cattle breed, respectively are acceptable and may use this substitution as a distinct genetic marker in Lezhi Black goats and in Iranian cattle breed. In the present study, the Glu\textsuperscript{154} residue is presence in all breeds vs. Gly\textsuperscript{154} in Pakistani Afghani ovine breed. This mutation may causes reduction in growth hormone function in this breed resulting in the differences between them in physio-

### Table 3. Substituted residues of growth hormone protein of Egyptian Holstein crossbred and Ossimi breed vs. investigated NCBI breeds.

<table>
<thead>
<tr>
<th>Studied breeds</th>
<th>Modified residues</th>
<th>Investigated breeds</th>
</tr>
</thead>
<tbody>
<tr>
<td>Egyptian Ossimi breed</td>
<td>Residue</td>
<td>Position</td>
</tr>
<tr>
<td></td>
<td>Glucose</td>
<td>34</td>
</tr>
<tr>
<td></td>
<td>Leucine</td>
<td>106</td>
</tr>
<tr>
<td></td>
<td>Glutamine</td>
<td>109</td>
</tr>
<tr>
<td></td>
<td>Phenylalanine</td>
<td>122</td>
</tr>
<tr>
<td></td>
<td>Glutamic acid</td>
<td>152</td>
</tr>
<tr>
<td></td>
<td>Glucose</td>
<td>155</td>
</tr>
<tr>
<td></td>
<td>Tyrosine</td>
<td>167</td>
</tr>
<tr>
<td></td>
<td>Glutamine</td>
<td>207</td>
</tr>
<tr>
<td>Egyptian Holstein crossbred</td>
<td>Leucine</td>
<td>107</td>
</tr>
<tr>
<td></td>
<td>Glutamine</td>
<td>110</td>
</tr>
<tr>
<td></td>
<td>Glycine</td>
<td>156</td>
</tr>
<tr>
<td></td>
<td>Leucine</td>
<td>189</td>
</tr>
<tr>
<td>Egyptian Holstein crossbred</td>
<td>Methionine</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Alanine</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Proline</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>Arginine</td>
<td>208</td>
</tr>
</tbody>
</table>

Chinese cattle...
chemical properties. The Tyr residue is present in all tested breeds vs. His residue in Pakistani Latti ovine breed. Because the both histidine and tyrosine are polar amino acids (neutral), therefore Tyr can substitute with His residue [32].

Fig 12. Multiple alignments of the encoded proteins of Egyptian Holstein crossbred and Ossimi breed vs. tested NCBI database breeds.

3.4.6. Phylogeny Tree and Synteny

The phylogenetic tree of EH_GH and Os_GH protein sequences vs. investigated breeds has been shown in Fig. 13. The predicted synteny location between Egyptian Holstein crossbred (Chr 19) and Ossimi breed (Chr 11) is shown in Fig. 14. Comparisons of mammalian gene maps show
large regions of conserved synteny where the same genes map to equivalent regions in different species [4], [5], [42], [43] and [44].

Fig13. Phylogenetic tree of Egyptian Holstein crossbred and Ossimi breed growth hormone protein sequences vs. tested NCBI breeds.

From phylogenetic tree and synteny results have been illustrated that EH_GH gene at chromosome 19 is an orthologous gene to the Os_GH gene at chromosome 11, as expected from their close evolutionary relationship [45]. The results support our assumptions and our previous work [12], and can induce a clue for further studies to produce transgenic sheep carry an expressed copy of bovine growth hormone gene using Sperm Mediated Gene Transfer (SMGT) technique as mentioned previously in different mammalian systems [46], [47] and [48] in mice; [49] in rabbits; [50] in pigs; [51] and [52] in Goat; [53] and [54] in sheep and [55] in cattle.

Fig14. The predicted synteny location of growth hormone protein sequences between Egyptian Holstein crossbred (Chr 19) and Ossimi breed (Chr 11).

4. CONCLUSION

It could be concluded that GH sequence of Egyptian Holstein crossbred cattle and Ossimi sheep breed are orthologous genes (98%). We assume in this study that the novel genetic information about the bovine and ovine GH such Egyptian breeds can support further success possibilities of producing transgenic sheep carrying bovine GH, using sperm mediated gene transfer (SMGT) technique to enhance their performance.

ACKNOWLEDGMENTS

Many Thanks to Dr. Mamdouh Sharaf El-Deen, Professor of Animal Production, Animal Production Department, Faculty of Agriculture, Cairo University and Chairman of Cairo Poultry Company(CPC), for the financial support of this work.
Cloning and Characterisation of Orthologous Genes Encoding Growth Hormone in Egyptian Holstein Crossbred and Ossimi Sheep

REFERENCES

Cloning and Characterisation of Orthologous Genes Encoding Growth Hormone in Egyptian Holstein Crossbred and Ossimi Sheep


Waleid Mohamed El-Sayed Shakweer et al.

AUTHORS’ BIOGRAPHY

Waleid Shakweer is an Assistant Researcher, Animal Production Dept., National Research Centre (NRC), Dokki, Egypt. He obtained his M.Sc. degree in 2008 from Faculty of Agriculture, Cairo University. He is a PhD student at Cairo University, his Ph.D. study aims to produce transgenic spermatzoa carries bovine and/or ovine growth hormone cDNA. His interested fields are animal biotechnology aspects and transgenic animal. He attainted the International Summer Schools of Animal Genomics, Università Cattolica del Sacro Cuore, Piacenza, Italy. He granted 3 months at Vavilove Institute of General Genetics, Russian Academy of Science to isolate GH gene from Holstein cattle and packaged into cloning vector. Also he granted 3 months at Regional Centre of Veterinary Medicine, Faculty of veterinary medicine, Agrarian University, Russia to produce in vitro embryo production of sheep.

Yassein M. Hafez is a professor of Animal Physiology, Animal Production Department, Faculty of Agriculture, Cairo University and supervisor of Hormones and immunology Lab., Cairo University Research Park, Cairo University. He was born in 1961. His Ph. D. was done in IFAS, University of Florida, USA. His fields of interest are animal endocrinology, immunological aspects and assisted reproductive technologies. He is a supervisor of 10 Master of Science and 5 Ph. D. students working in projects related to climatic changes, endocrinology, molecular biology and reproductive physiology.

Marwa N.M.E. Sanad is a Geneticist “Molecular genetics and bioinformatics”. She is a researcher in Genetics and Cytology Department, National Research Centre in Giza, Egypt, and a Research associate in College of Agriculture, Human, and Natural Resources Sciences, Washington State University, WA, USA. She received her BSc, MSc and PhD degrees in Genetics from Faculty of Agriculture, Ain Shams University, Egypt, and a diploma in Bioinformatics in 2006. She earned her Ph.D. in 2011 on comparative genomic studies, her current research areas are on epigenetics, evolutionary genetics and genomic prediction/selection.

Ashraf El-Sayed is an Associate Professor of reproductive physiology, Department of Animal Production, Faculty of Agriculture, Cairo University. Moreover he is the Director of Embryology Lab., and Co-Director of Biotechnology Sector, Cairo University Research Park (CURP), at Faculty of Agriculture, Cairo University since 2008. He is a certified trainer at Cairo University since 2008. He was born in 1971 in Cairo Egypt. He obtained his PhD degree in 2006 from Bonn University, Germany. His field of interest is animal biotechnology aspects which include In vitro embryo production of farm animals; in addition he is an expert of molecular biology tools and its applications in embryonic genome. He is a leader of a group consists of 10 master and 3 PhD students working in three projects related to embryonic genome located at CURP.

Ibrahim Mohamad Awadalla is a professor of animal nutrition, Animal Production Department, National Research Centre. He is born in 1955. Chairman of the upgraded of productivity of farm animals: It is one of the units of the special character of the National Research Center established in accordance with the provisions of presidential Decree No. 50 of 1988 on the executive regulations of the National Research Center and the decision of the President of the National Research Center No. 139 of 10.4.2005 and Unit aims to raise the rate of production performance of animals on the production of milk and meat using modern scientific methods of breeding, care and safety rules for the quality of feed composition.

MamdoughIbrahim Mohamad is a professor of animal nutrition and head of Animal Production Department, National Research Centre. He is born in 1956. His interested fields are animal nutrition and animal breeding. His projects aim to enhance the Baladi goats and sheep by crossing with superior forging breeds such Bore and Damascus goats and Awassi sheep.