

# GOPEN ACCESS

**Citation:** Abdel-Fattah M, Saeed H, El-Shennawy L, Shalaby M, Embaby A, Ataya F, et al. (2019) The Arabian camel, *Camelus dromedarius* interferon epsilon: Functional expression, *in vitro* refolding, purification and cytotoxicity on breast cancer cell lines. PLoS ONE 14(9): e0213880. https://doi.org/ 10.1371/journal.pone.0213880

Editor: Paulo Lee Ho, Instituto Butantan, BRAZIL

Received: February 28, 2019

Accepted: August 9, 2019

Published: September 6, 2019

**Copyright:** © 2019 Abdel-Fattah et al. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: The data underlying this study are within the paper and it Supporting Information files, and have been uploaded to GenBank here: https://www.ncbi.nlm.nih.gov/ nuccore/MH025455.1?report=genbank&to=585.

**Funding:** The authors received no specific funding for this work.

**Competing interests:** The authors have declared that no competing interests exist.

RESEARCH ARTICLE

# The Arabian camel, *Camelus dromedarius* interferon epsilon: Functional expression, *in vitro* refolding, purification and cytotoxicity on breast cancer cell lines

Manal Abdel-Fattah<sup>1</sup>, Hesham Saeed<sup>1</sup>\*, Lamiaa El-Shennawy<sup>2</sup>, Manal Shalaby<sup>3</sup>, Amira Embaby<sup>1</sup>, Farid Ataya<sup>4,5</sup>, Hoda Mahmoud<sup>1</sup>, Ahmed Hussein<sup>1</sup>

1 Department of Biotechnology, Institute of Graduate Studies and Research, Alexandria University, Alexandria, Egypt, 2 Department of Environmental Studies, Institute of Graduate Studies and Research, Alexandria University, Alexandria, Egypt, 3 Genetic Engineering and Biotechnology Research Institute (GEBRI), City for Scientific Research and Technology Applications, New Borg Al-Arab City, Alexandria, Egypt, 4 Biochemistry Department, College of Science, Riyadh, King Saud University, KSA, 5 National Research Centre, Dokki, Giza, Egypt

\* hsaeed1@ksu.edu.sa

## Abstract

The current study highlights, for the first time, cloning, overexpression and purification of the novel interferon epsilon (IFNE), from the Arabian camel Camelus dromedaries. The study then assesses the cytotoxicity of IFN against two human breast cancer cell lines MDA-MB-231 and MCF-7. Full-length cDNA encoding interferon epsilon (IFNE) was isolated and cloned from the liver of the Arabian camel, C. dromedarius using reverse transcription-polymerase chain reaction. The sequence analysis of the camel IFN<sub>E</sub> cDNA showed a 582-bp open reading frame encoding a protein of 193 amino acids with an estimated molecular weight of 21.230 kDa. A BLAST search analysis revealed that the C. dromedarius IFNs shared high sequence identity with the IFN genes of other species, such as Camelus ferus, Vicugna pacos, and Homo sapiens. Expression of the camel IFNE cDNA in Escherichia coli gave a fusion protein band of 24.97 kDa after induction with either isopropyl β-D-1-thiogalactopyranoside or lactose for 5 h. Recombinant IFNs protein was overexpressed in the form of inclusion bodies that were easily solubilized and refolded using SDS and KCI. The solubilized inclusion bodies were purified to apparent homogeneity using nickel affinity chromatography. We examined the effect of IFNs on two breast cancer cell lines MDA-MB-231 and MCF-7. In both cell lines, IFN<sub>ε</sub> inhibited cell survival in a dose dependent manner as observed by MTT assay, morphological changes and apoptosis assay. Caspase-3 expression level was found to be increased in MDA-MB-231 treated cells as compared to untreated cells.

#### Introduction

IFNs are members of a large cytokine family of evolutionarily conserved pleiotropic regulators of cellular functions; they are relatively low-molecular weight signaling proteins (20-25 kDa) usually glycosylated and produced by a variety of cells, such as epithelia, endothelia, stroma, and cells of the immune system [1-3]. The expression of IFNs is induced by a variety of different stimuli associated with viral infections, bacteria, parasites, inflammation, and tumorigenesis [4]. IFNs, therefore, induce a diverse range of biological functions and responses, including cell proliferation and differentiation, inflammation, chemotaxis, immune cell (natural killer cells and macrophages) activation, and apoptosis [5, 6]. The key to understanding these regulatory proteins lies in the recognition of their pleiotropism, overlapping activities, functional redundancies, and side effects [1]. Based on the type of receptors they interact with for signal transduction, IFNs are classified into three major types namely, type I, II, and III, which have different gene and protein structures and biological activities [7]. The mammalian type I IFNs represents a large family of related proteins, mainly virus-inducible, divided into eight subfamilies named  $\alpha$ ,  $\beta$ ,  $\omega$ ,  $\delta$ ,  $\varepsilon$ ,  $\nu$  and  $\kappa$  [8, 9]. Besides the autocrine activation of antiviral responses, type I IFNs function systematically to induce an antiviral state in the surrounding and distal cells [10, 11]. In combination with chemo and radiation therapies, interferon therapy is used as a treatment of some malignant diseases, such as hairy cell leukemia, chronic myeloid leukemia, nodular lymphoma, and cutaneous T-cell lymphoma [12]. The recombinant IFN- $\alpha$ 2b can be used for the treatment of patients with recurrent melanomas [13]. Hepatitis B, hepatitis C, and HIV are treated with IFN- $\alpha$  often in combination with other antiviral drugs [14, 15].

One of the most recently discovered interferon is the interferon epsilon (IFNE). Signal transduction by IFN $\epsilon$  is mediated through binding to the interferon  $\alpha/\beta$  receptor (IFNAR), despite its low sequence homology with  $\alpha$ - and  $\beta$ -type interferons. Although binding to the same heterodimeric receptor pair, they evoke a broad range of cellular activities, affecting the expression of numerous genes and resulting in profound cellular changes [10, 16-18]. The expression of IFN $\varepsilon$  is neither induced by a pattern recognition receptor pathway nor by an exposure to viral infection [19]. Unlike other type-I IFNs, IFN $\varepsilon$  is constitutively expressed in the lung, brain, small intestine, and reproductive tissue; thus, it is thought to play a role in reproductive function, in either viral protection or early placental development in placental mammals [16, 17]. IFNE has high amino acid sequence homology with other type-I interferons, of which IFN- $\beta$  is the closest paralog, and they share 38% identical residues. A common structural feature of IFN $\varepsilon$  is the lack of a disulfide linkage and the presence of two glycosylation sites represented by asparagine 74 and 83. Many IFNs genes have been cloned and characterized from a variety of species such as human, pig, mouse, dog, cat, cattle, chicken, turkey, goose, zebra fish, and Atlantic salmon [20-23]. However, the information about the IFNE from the Arabian one-humped camel, *Camelus dromedarius*, has not been reported yet. This domesticated camel is one of the most important animals in the Arabian Peninsula, having high cultural and economic value. In Saudi Arabia, it comprises 16% of the animal biomass and is considered as the main source of meat [24, 25]. The aim of the present study was the isolation of full-length C. dromedarius IFNE gene, followed by its expression in Escherichia coli, in vitro refolding of the recombinant protein, purification, and characterization of the purified IFNE protein. Cytotoxicity and apoptosis assays were then performed to define the effect of the purified recombinant IFNe protein on human cancer cell lines. The results of this study contribute towards the importance of discovering and characterizing IFNE from this unique Arabian camel, and propose its potential use for the treatment of cancer.

### Materials and methods

#### Chemicals and reagents

All chemicals and reagents were of molecular biology, analytical, or chromatographic grade. Water was de-ionized and milli-Q-grade.

#### Tissue collection and RNA isolation

Liver tissues (2 g) from adult male one-humped Arabian camel, *C. dromedarius*, were collected from a slaughter-house located in the north of Riyadh City, Kingdom of Saudi Arabia. The animals were sacrificed under the observation of a skilled veterinarian, and the liver samples were taken and immediately submerged in 5 mL of RNA later<sup>®</sup> solution (Ambion, Courtabeuf, France) to preserve the integrity of RNA. The samples were kept at 4°C overnight and thereafter stored at -80°C until used for RNA isolation. The liver samples were removed from -80°C and left at room temperature until thawed completely. Fifty milligrams were homogenized in 0.5 mL RLT lysis solution supplemented with 1% 2-mercaptoethanol using a rotor-stator homogenizer (MEDIC TOOLS, Switzerland). Total RNA was isolated and purified using the RNeasy Mini Kit (Qiagen, Germany), with a DNase digestion step following the manufacturer's protocol. The elution step was performed using 50 µL nuclease free water. The concentration, purity, and integrity of the isolated purified total RNA were determined using the Agilent 2100 Bioanalyzer System and Agilent total RNA analysis kit, according to the manufacturer's protocols (Agilent Technologies, Waldbronn, Germany).

#### First strand cDNA synthesis and amplification of camel IFN $\epsilon$ gene

TotalRNA, isolated previously from adult male one-humped Arabian camel, *C. dromedaries*, was used in the current study as a source for camel *IFNe* gene. Two micrograms of total RNA were reverse transcribed into the first strand cDNA using the ImProm-II Reverse Transcription System (A3800, Promega, Madison, USA) according to the manufacturer's protocol and used as a template for the amplification of the full-length camel IFNe cDNA. A polymerase chain reaction (PCR) was conducted in a final volume of 50 µL, containing 25 µL 2X high-fidelity master mix (GE Healthcare, USA), 3 µL (30 pmol) of each *IFNe* gene forward primer that contains an *Eco*RI restriction site (5′ – <u>GAATTC</u> ATGATTAACAAGCCTTTCTT–3′) and a reverse primer that contains a *Hind*III restriction site (5′ – <u>AAGCTTAGGATCCATTCCTT</u> GTTTGC–3′), and 5 µL cDNA. The PCR amplification was performed using the following reaction conditions: 1 cycle at 95°C for 5 min, followed by 30 cycles at 95°C for 30 s, 55°C for 30 s, and 72°C for 1 min. A final extension step was carried out at 72°C for 5 min. The PCR products were resolved in a 1.5% agarose gel stained with 0.5 µg/mL ethidium bromide.

## Cloning and sequencing of the full-length camel IFNE cDNA

The PCR product was first cloned into the pGEM<sup>®</sup>-T Easy vector (Promega Co. Cat #A1360) to facilitate the sequencing process and subcloning into the pET28a (+) expression vector. The ligation reaction was carried out in a clean sterile 1.5-mL Eppendorf tube containing 4  $\mu$ L of the PCR product, 1  $\mu$ L (50 ng) of pGEM-T-Easy vector (Promega, USA), 1  $\mu$ L of 10X ligase buffer, and 1 U of ligase enzyme. The final volume of the reaction was adjusted to 10  $\mu$ L by the addition of nuclease free water. The reaction tubes were kept at 16°C overnight, after which 5  $\mu$ L was used to transform the *E. coli* JM109 competent cells, according to Sambrook et al. (1989) [26]. Screening was carried out on the selective LB/IPTG/X-gal/Ampicillin/agar plates. The recombinant plasmids were prepared from some positive clones using the PureYield Plasmid Miniprep System (Cat #A1222, Promega, Madison, USA). The sequencing of the cloned

insert was carried out according to Sanger et al. 1977 [27] using the T7 (5'-TAATACGACTCA CTATAGGG-3') and SP6 (5'-TATTTAGGTGACACTATAG-3') sequencing primers. The sequence analysis was carried out using the DNAStar, BioEdit, and Clustal W programs.

#### Phylogenetic tree and structure modeling analysis

A phylogenetic tree analysis was constructed according to Dereeper et al. [28], using the Phylogeny.fr software (http://www.Phylogeny.fr). The nucleotide sequences for the Arabian camel IFNɛ cDNA was analyzed using the basic local alignment search tool (BLAST) programs BLASTn, BLASTp (http://www.ncbi.nlm.nih.gov), and a multiple sequence alignment was carried out using the ClustalW, BioEdit, DNAStar, and Jalview programs. The protein sequence was obtained by translating the cDNA nucleotides sequence by using a translation tool at the ExPasy server (http://web.expasy.org/translate/). The protein sequence was submitted to the Swiss-Model server for structure prediction, and the structural data were analyzed by the PDB viewer program. Finally, the predicted 3D structure models were built based on the multiple threading alignments by using the local threading meta-server (LOMET) and iterative TAS-SER assembly simulation [29, 30].

### Sub-cloning into pET-28a (+) vector

The IFNe cDNA insert cloned into the pGEM-T-Easy plasmid was released using the EcoRI and *Hind*III restriction enzymes (2 units each) according to Sambrook et al. (1989) [26]. The released insert was purified from the agarose gel using the QIAquick Gel Extraction Kit (Cat. # 28704, QIAGEN) and sub-cloned into the pET-28a (+) expression vector. The plasmid pET-28a (+) (Novagen) carries an N-terminal His-Tag/thrombin/T7 configuration, and the expression of the cloned gene is under the control of a T7 promoter. A 2-µg aliquot of plasmid pET-28a (+) was digested with 2 units of EcoRI and HindIII at 37°C overnight, after which the digestion reaction was terminated by heating the tubes at 65°C for 15 min. The linearized plasmid was treated with 2 units of shrimp alkaline phosphatase (Promega, Madison, USA) at 37°C for 30 min. Finally, the reaction was terminated by incubation at 70°C for 10 min. The ligation reaction was carried out in a tube containing  $2 \mu L$  (50 ng) of pET28a (+),  $2 \mu L$  (100 ng) of IFNE cDNA insert, 1 µL 10X ligase buffer, and 1 µL (2 units) of ligase enzyme. The final volume was adjusted to 10  $\mu$ L by the addition of nuclease free water, and the tube was incubated at 16°C overnight. Subsequently, 5 μL of the ligation reaction was used to transform E. coli BL21(DE3) pLysS (Cat. # P9801, Promega, USA) competent cells, according to Sambrook et al. (1989) [26]. The recombinant E. coli BL21(DE3) pLysS harboring the pET-28a (+) vector was screened on the selective LB/IPTG/X-gal/Kanamycin/agar plates and by using the colony PCR strategy utilizing the IFNe gene-specific primers. The recombinant plasmids were isolated from the positive clones using the Pure Yield Plasmid Miniprep System (A1222, Promega, USA), and some potential positive plasmids containing the cDNA insert were digested with EcoRI and HindIII to confirm the presence of the IFNE cDNA insert.

## Expression of camel IFNE cDNA in E. coli BL21(DE3) pLysS

The transformed *E. coli* BL21(DE3) pLysS harbouring the recombinant plasmid were cultured in 1 L of Luria broth medium supplemented with 34  $\mu$ g/mL kanamycin and incubated at 37°C for 4 h at 250 rpm. When the optical density at 600 nm reached 0.6, isopropyl- $\beta$ -D-1-thiogalactopyranoside (IPTG) was added to the culture at a concentration of 1 mM. The culture flask was incubated at 37°C with shaking at 250 rpm for 5 h, after which the bacterial cells were harvested by centrifugation at 8000 rpm for 20 min at 4°C. The bacterial pellets were re-suspended in 10 mL of 0.1 M potassium phosphate buffer, pH 7.5, containing 50% glycerol. The bacterial cell suspension was then sonicated on an ice-bath using 4x 30-s pulses, and the cell debris were removed by centrifugation at 10,000 rpm for 10 min at 4°C, after which the supernatant and pellets were collected in separate Eppendorf tubes. The pellets were re-suspended in 5 mL of 0.1 M potassium phosphate buffer, pH 7.5, containing 50% glycerol and both supernatant and pellets were used for further analysis. The gene expression was also analysed using lactose as an inducer at a concentration of 2 g/L in the fermentation medium.

#### **Protein determination**

Protein concentration was determined according to Bradford (1976) [31], using 0.5 mg/mL of bovine serum albumin (BSA) as a standard.

# Sodium dodecyl sulfate gel electrophoresis (SDS-PAGE) and western blotting analysis

The expression of the camel recombinant *IFNe* gene in *E. coli* was checked by performed a 12% SDS-PAGE according to Laemmli, 1970 [32]. After electrophoresis, the gel was stained with Coomassie Brilliant Blue R-250 followed by de-staining in a solution of 10% (v/v) methanol and 10% (v/v) acetic acid. Recombinant *C. dromedarius* IFNe protein was detected by wesetrn blotting using 6x-His-Tag monoclonal antibody (His.H8, Cat# MA1-21315, Thermo-Fischer Scientific) at a 1:1000 dilution according to Towbin et al. [33]. Secondary antibody used was goat anti-mouse IgG labeled with horse radish peroxidase (Invitrogen Cat# G-21040) at a 1:2000 dilution. The membrane was developed using chromogenic substrate 3,3',5,5'-Tetramethyl benzidine (TMB) liquid substrate system (Sigma-Aldrich, Cat# T0565).

#### Refolding of C. dromedarius recombinant IFNE protein inclusion bodies

The transformed *E. coli* BL21(DE3) pLysS cells containing over-expressed camel IFNε protein was disrupted by sonication, and the inclusion bodies were recovered by centrifugation at 10,000 rpm for 30 min at 4°C. The pellets were washed three times with 20 mM Tris-HCl, pH 8.0, and after the final wash, the pellets were resuspended in denaturation buffer containing 50 mM M Tris-HCl (pH 8.0), 0.3 M NaCl, and 2% SDS with continuous stirring on an ice-bath until the solution becomes clear. The protein solution was then kept at 4°C overnight, followed by centrifugation for 10 min at 10,000 rpm and 4°C to precipitate the excess SDS. Subsequently, KCl was added to the supernatant at a final concentration of 400 mM, and the solution was kept at 4°C overnight. Thereafter, centrifugation was carried out for 10 min at 10,000 rpm and 4°C, and the clear solution was dialyzed overnight against 50 mM potassium phosphate buffer (pH 7.5) and applied to a nickel affinity column [33, 34].

#### Affinity purification of C. dromedarius recombinant IFNE

The recombinant IFN $\epsilon$  protein was purified using a single-step High-Select High Flow (HF) nickel affinity chromatographic gel (Sigma-Aldrich, Cat. # H0537). The nickel affinity column (1.0 cm × 1.0 cm) was packed with the affinity matrix and washed thoroughly with 30 mL of de-ionized water, followed by equilibration with the 5-bed volumes of 50 mM potassium phosphate buffer (pH 7.5) containing 20 mM imidazole. A solution of solubilized inclusion bodies (5 mL) was loaded onto the column, and the column was washed with 5-bed volumes of 50 mM potassium phosphate buffer (pH 7.5) containing 20 mM imidazole. The recombinant IFN $\epsilon$  protein was eluted with 50 mM potassium phosphate buffer (pH 7.5) containing 500 mM imidazole. The collected fractions were measured at 280 nm, and the fractions presented in the second peak were pooled together and dialyzed overnight against 50 mM potassium

phosphate buffer (pH 7.5). The purity of the dialyzed recombinant IFNE protein was checked by performing 12% SDS-PAGE.

#### **Electron microscopy analysis**

The recombinant *C. dromedarius* IFNe inclusion bodies were fixed in a solution of formaldehyde and glutaraldehyde (4:1) and observed and analyzed by transmission electron microscopy (TEM; JEOL-JSM 1400 plus) and scanning electron microscopy (SEM; JEOL-JSM 5300).

# Cytotoxicity of *C. dromedarius* recombinant IFN epsilon on breast cancer cell lines

Human breast cancer cell lines, MDA-MB-231 and MCF-7, were obtained from the lab of Professor Stig Linder, Karolinska Institute, Sweden. Cells were cultured in Dulbecco's Modified Eagle's Medium supplemented with 10% Fetal Bovine Serum (Sigma), 100 U/mL penicillin, and 100 mg/mL streptomycin. Cells were maintained in 5% CO<sub>2</sub> at 37°C.

#### MTT assay

MDA-MB-231 and MCF-7 cells were seeded in 96 well plates (15,000 and 10,000 cells/ well, respectively). After 24 h, cells were treated with different concentrations of recombinant interferon epsilon and the control cells received untreated medium in the same buffer. Cells were washed twice with PBS after 48 h of incubation, and 3-(4,5-Dimethyl-2-thiazolyl)-2,5-diphenyl-2H-tetrazolium bromide (MTT) reagent (10  $\mu$ L of 5 mg/mL) (Serva) in 100  $\mu$ L serum free medium was added to each well. After 3–4 h of incubation at 37°C, the medium was discarded, and cells were incubated with 100  $\mu$ L of DMSO. Plates were shaken, then absorbance was measured at 490 nm [35].

#### Apoptosis assay

Apoptosis was analyzed using Annexin V-FITC apoptosis detection kit (Miltenyi Biotec). MDA-MB-231 and MCF-7 were incubated with recombinant IFNɛ for 48 h. The floating cells were detached from the plate surface and attaching cells were harvested by trypsinization and pelleted by centrifugation. The cell pellets were resuspended in binding buffer and incubated with fluorescein isothiocyanate (FITC)-labeled Annexin V for 15 min in the dark at room temperature. Cells were washed and resuspended in binding buffer, then and Propidium Iodide was added. The stained cells were analyzed in Flow Cytometry Service core at Center of Excellence for Research in Regenerative Medicine and its Applications using BD FACSCalibur flow cytometer (BD Biosciences).

#### Caspase-3 assay

Caspase-3 expression level was detected in MDA-MB-231 untreated and camel recombinant IFNɛ treated cell line using Human Caspase-3 (Casp-3) sandwich ELISA Kit (SinoGeneClon Biotech Co., Ltd) according to manufacturer's instructions.

#### Detection of gram-negative bacterial endotoxins

Gram negative bacterial endotoxins were determined using Thermo Scientific Pierce LAL Chromogenic Quantitation kit (ThermoFischer Cat. # Cat. # 88282) according to the manufacturer manual. The developed yellow color was measured at 405 nm using a microplate plate reader and the amount of endotoxin in the samples was calculated from the standard curve prepared from *E. coli* endotoxin included with the kit.

#### Statistical analysis

GraphPad Prism 6.0 Software was used to perform statistical analyses. One way or two ways ANOVA (followed by Tukey or Sidak's posttest) were used where appropriate. Data are presented as mean  $\pm$  SEM or  $\pm$  SD from at least two independent experiments.

### **Results and discussion**

#### C. dromedarius IFNE full-length cDNA isolation and sequence analysis

By far, most information about type I IFNs has stemmed from the studies of IFNs from other species such as human, turkey, zebra fish, and bovine, but no published data is available on the Arabian camel IFNs [11, 20, 21]. In the present study, the full-length IFNE cDNA of the Arabian camel, C. dromedarius, was isolated by RT-PCR using gene-specific primers designed from the available expressed sequence tag (EST) camel genome project database (http://camel. kacst.edu.sa/). The PCR product corresponding to the 582 nucleotides represents the fulllength IFNE cDNA (Fig 1). The PCR product was cloned into the pGEM-T-Easy vector, and the cDNA insert was sequenced using the T7 and SP6 primers. The nucleotide sequence was deposited in the GenBank database under the accession number MHO25455. Comparing the nucleotide sequence of the Arabian camel IFN $\varepsilon$  cDNA with the nucleotide sequences of other species deposited in the GenBank database using the Blastn and Blastp programs available on the National Center for Biotechnology Information (NCBI) server revealed that the putative camel IFNe gene has high statistically significant similarity scores to numerous IFNe genes from other species (Table 1). To determine the relatedness of C. dromedarius IFNE with known amino acid sequences available in the GenBank database, a multiple sequence alignment was conducted (Fig 2). It was observed that the percentage identity of C. dromedaries IFNE with IFNE from other species was 100% for Camelus ferus (GenBank accession no. XP\_006179655), 95% for Vicugna pacos (GenBank accession no. XP\_006215195), 82% for Sus scrofa (GenBank accession no. NP 001098780), 78% for Bos taurus (GenBank accession no. XP 005209958), and 75% for Homo sapiens (GenBank accession no. NP 795372). Moreover, the camel IFNE has high amino acid sequence homology with other type I IFNs, of which the closest paralog is IFN $\beta$ , and they share 38% identical residues [10]. A phylogenetic tree constructed (Fig 3) from the amino acid sequences of the predicted IFNE proteins deposited in the GenBank indicated that the Arabian camel IFNE took a separate evolutionary line distinct from other ungulates and mammalian species, including H. sapiens.

#### C. dromedarius IFNE structure annotations and predicted 3D structure

The Arabian camel IFNE primary structure and the protein motif secondary structure annotation prediction are shown in Figs 4 and 5A. The nucleotides and the deduced amino acid sequence showed an open reading frame consisting of 582 nucleotides and 193 amino acid residues with a calculated molecular weight of 21,230 kDa. The isoelectric point, predicted using a computer algorithm, was found to be 9.03. From the primary structure and the multiple sequence alignment of camel IFNE with other ungulates and human, several observations merit discussion. First, the primary structure homology was greater than 75% among type I IFNs of different species. The high degree of amino acid sequence identity and conservation is presumably due to the functional constraints during evolution, although it was clear from the phylogenetic tree analysis (Fig 3) that the camel IFNE took a separate evolutionary line away from other species having type I IFNs. Second, the putative Arabian camel IFNE protein is characterized by the presence of amino acid residues Ser<sup>38</sup>, Glu<sup>112</sup>, and Ile<sup>167</sup> that are highly conserved among type I INFE. Third, the Arabian camel IFNE putative protein contained



Fig 1. Agarose gel (1.5%) electrophoresis of PCR product for *C. dromedarius* interferon epsilon gene (Lane 2). Lane 1 represents 100 base pair DNA ladder.

https://doi.org/10.1371/journal.pone.0213880.g001

three cysteine residues (Cys<sup>53</sup>, Cys<sup>163</sup>, and Cys<sup>175</sup>), like those found in the bovine IFN<sub>E</sub>, and two of them, probably Cys<sup>53</sup> and Cys<sup>175</sup>, might be involved in the formation of intramolecular disulfide bonds that link the N-terminus to the end of helix F. These cysteine residues are highly conserved amongst other members of the examined type-I IFN homologs, such as human IFN $\lambda$  and  $\beta$  and rabbit interferon- $\gamma$  [36–38]. Fourth, the analysis of putative glycation sites in the camel IFN $\varepsilon$  protein (Figs 3 and 4) led to the prediction of seven potential glycation sites, although not occurring within the conserved signal for glycosylation, Asn-Xaa-Ser/Thr; these sites are 3NKPF, 35NRES, 43NKLR, 59NFLL, 90NLFR, 139NLRL, and 173NRCL. These glycation sites might act as the sites of protection against proteases-mediated hydrolysis and contributing to the process of folding, oligomerization, and stability of the protein. The identification of such sites raised the possibility that the putative camel IFNE might form a glycoprotein [39]. Fifth, the Arabian camel IFNe amino acid sequence was characterized by the presence of IFNAR-1- and IFNAR-2-binding domains. The putative IFNAR-1-binding domain is critical for receptor recognition and biological activity, and this domain was represented by the amino acid residues,  $F^{29}$ ,  $Q^{30}$ ,  $R^{33}$ ,  $R^{36}$ ,  $E^{37}$ ,  $K^{40}$ ,  $N^{43}$ , and  $K^{44}$ , located in the first  $\alpha$ -helix of the camel IFNE protein (Fig 5A). The IFNAR-2-binding site contained the amino acid residues, L<sup>54</sup>, P<sup>55</sup>, H<sup>56</sup>, R<sup>57</sup>, K<sup>58</sup>, N<sup>59</sup>, F<sup>60</sup>, L<sup>61</sup>, P<sup>63</sup>, Q<sup>64</sup>, K<sup>65</sup>, Q<sup>71</sup>, and Y<sup>72</sup>. Other conserved amino acids residues involved in the binding of different ligands and DNA are shown in Table 2.

The predicted 3D structure of the Arabian camel IFN $\epsilon$  indicated that the protein secondary structure consisted of six  $\alpha$ -helices labelled from A to F. The composition of the predicted secondary structure revealed 61.5%  $\alpha$ -helices, 32.6% coils, and 2.1% turns. Compared with other type I $\alpha$  IFNs, the camel protein showed an extended C-terminus (Fig 5B, 5C and 5D). It was observed that the overall folding in the 3D structure of camel IFN $\epsilon$  was quite similar to that of the bovine-type IFN $\epsilon$  [8]. Moreover, the alignment template model (Fig 5E and 5F) showed 36.36% similarity between the camel IFN $\epsilon$  and *H. sapiens* type-I  $\alpha$ 2 IFN, with the preservation of the components of the secondary structures,  $\alpha$ -helices, coils, and turns.

Animal species	Accession no.	% Identity
Camelus ferus	XP_006179655	100
Vicugna pacos	XP_006215195	95
Balaenoptera acutorostrata	XP_007176883	82
Sus scrofa	NP_001098780	82
Hipposideros armiger	XP_019484975	81
Orcinus orca	XP_004275093	80
Delphinapterus leucas	XP_022407268	80
Lipotes vexillifer	XP_007455001	80
Tursiops truncates	XP_019790467	79
Bos mutus	XP_005887920	79
Bos taurus	XP_005209958	78
Bison bison bison	XP_010851614	78
Ovis aries	XP_011982517	78
Macaca nemestrina	XP_011768789	77
Papio anubis	XP_021783163	77
Homo sapiens	NP_795372	75

Table 1. Homology of the deduced amino acids of C. dromedarius interferon epsilon with other species.

https://doi.org/10.1371/journal.pone.0213880.t001

# PLOS ONE

Canclas downdarias 1 MIN NR OF FE LV. VLLA FST. IF SR. EX. NPLL FOORNUM R. S. L. KLLNKLR, S. S. I. OC. C. PHR. KH. FLL P. OK. SM. D. PHOY OK, G. H. LA I. 90 Canclas down and the second seco					10			20			30			40			50			60			70			
Candual prova 1 MI NKOP FF E I VLVL LA FST IF SR EL KP LLF00 RRV WR ESLK LLNKL RY SS I 00 CL PHRKN FLLP0 RSVD PH0Y 0K (0H LA 18 00 Plagma pace 1 MI NKOP FF E I VLVL LA SS SVC SR EL KLVLF00 RRV WR ESLK LLNKL OT SS I 00 CL PHRKN FLLP0 RSVD PH0Y 0K (0A LA 18 00 Balamoptera actuarostam 1 MI NKS FF E I VLVL LA SS SVC SR EL KLVLF00 RRV WR ESLK LLNKL OT SS I 00 CL PHRKN FLLP0 RSVM PH0Y 0K (0A LA 18 00 Bos mutus 1 MI NKS FF E I VLVL A SS SVC SR EL KLVLF00 RRV WR ESLK LLNKL OT SS I 00 CL PHRKN FLLP0 RSVM PH0Y 0K (0A LA 18 00 Bos mutus 1 MI NKS FF E I VLVL A SS SVC SR EL KLVLC00 RRV MO ESLK LLNKL OT SS I 00 CL PHRKN FLLP0 RSVM PH0Y 0K (0A VLA 18 00 Ovis arise 1 MI NKA FF E I VLVL A SS TVC SO EL KLVLC00 RRV MO ESLK LLNKL OT SS V00 CL PHRKN FLLP0 RSVM PH0Y 0K (0A VLA 18 00 Ovis arise 1 MI NKA FF E I VLVL A SS TVC SO EL KLVLC00 RRV MO ESLK LLNKL OT SS I 00 CL PHRKN FLLP0 RSVM PH0Y 0K (0A VLA 18 00 Ovis arise 1 MI NKA FF E I VLVL LA SS TVC SO EL KLVLC00 RRV MO ESLK LLNKL OT SS I 00 CL PHRKN FLLP0 RSVM PH0Y 0K (0A VLA 18 00 Macaca mututa 1 MI NKA FF FE I VLVL LA SS TVC SO EL KLVLC00 RRV MO ESLK LLNKL OT SS I 00 CL PHRKN FLLP0 RSVM PH0Y 0K (0A VLA 18 00 Macaca mututa 1 MI NKA FF FE I VLVL LA ST TVF FLD KLLF A DOR RSVM RESLK LLNKL OT SS I 00 CL PHRKN FLLP0 RSVM PH0Y 0K (0A VLA 18 00 Macaca mututa 1 MI NKA FF FE I VLVL LA ST TVF FLD KLLF A DOR RSVM RESLK LLNKL OT SS I 10 CL PHRKN FLLP0 RSVM PR0Y 0K (0A AL 79 Homo sapiens 1 MI I KHFF FG I VLVL LA ST TVF FLD KLLF A DOR RSVM RESLK LLNKL OT SS I 10 CL PHRKN FLLP0 RSVM PR0Y 0K (0A AL 79 Homo sapiens 1 MI I KHFF FG I VLVL LA ST TVF FLD KLLF A DOR VR RS KLLSTL DA DOL SS I 10 CL PHRKN FLLP0 RSVM PR0Y 0K (0A AL 79 HOMO Sapiens 11 HEML 00 I FLF FA VI SLD 0WE E I MODI FI SE LH00 LEYLE TL RLD A DA DOK SVM RC NV OK (1H DVL E SOEY 160 D Candual forma 11 H HFF FG I VLVL LA ST TVF FLD I MODI FI SE LH00 LEYLE TL I RLD A DA DOK SVM RC NV OK (1H DVL E SOEY 160 D Candual forma 11 H HFF FG I VLVL LA ST TVF FLD I MODI FI SE LH00 LEYLE TL I RLD A DA DOK SVM RC NV NV OK (1H DVL E SOEY 1	Camelus dromedarius	1	MINK	FFE	IVI	VLLA	FST	IFSR	ELK	PILI	FQQR	RVN	RES	LKLI		RTS	SIQ	QCL	PHRK	NFL		SMD	PHO	YQK	HILAI	80
Image appears   1   1   MIL MK PF FE I VLVL LASSTVF SRELK PVLF00 KRVVR RESLKLLNKL RSSI 00 CLPHKKNFLLP0KSVDPH0Y0 GOALA 1   300     Sussorgia   1   MIL MK PF FE I MLVL LASST 0F SRELK LVLF00 KRVVR RESLKLLNKL 0TSSI 00 CLPHKKNFLLP0KSVDPH0Y0 GOALA 1   300     Sussorgia   1   MIL MK AF FE I VLVL LASST 0F SRELK LVLC00 KRVVR SLKLLNKL 0TSSI 00 CLPHKKNFLLP0KSVDPH0Y0 GOVLA 1   300     Sonstatus   1   MIL MK AF FE I VLVL LASST 0F SRELK LVLC00 KRVVR 0F SLKLLNKL 0TSSI 00 CLPHKKNFLLP0KSVDPH0Y0 GOVLA 1   300     Sonstatus   1   MIL MK AF FE I VLVL LASST 0F SRELK LVLC00 KRVVR 0F SLKLLNKL 0TSSI 00 CLPHKKNFLLP0KSVDPH0Y0 GOVLA 1   300     Sonstatus   1   MIL MK AF FE I VLVL LASST 0F SRELK LVLC00 KRVVR 0F SLKLLNKL 0TSSI 00 CLPHKKNFLLP0KSVDPH0Y0 GOVLA 1   300     Gaacaa and and and the MK AF FE I VLVL LASST 0F SRELK LVLC00 KRVVR 0F SLKLLNKL 0TSSI 00 CLPHKKNFLLP0KSVDPH0Y0 GOVLA 1   300     Gaacaa and and the MK AF FE I VLVL LASST 0F SRELK LVLF00 KRVVR SLKLLNKL 0TSSI 00 CLPHKKNFLLP0KSVDPH0Y0 GOVLA 1   300     Gaacaa and the MK AF FE I VLVL LASST 0F SRELK 1   FE I KOLVL VLLASST 0F SRELK 1   300 CL SHKKN FLLP0KSVDPH0Y0 GOVLA 1   300     Gaacaa and the MK AF FE I VLVL LASST 0F SRELK 1   SRE 0 CL SRELK 1   SRE 0 CL SRELK 1   300 CL SHKKN FLLP0KSVDPH0Y0 GOVLA 1   300     Gaacaa and the MK AF FE I VLVL LASST 0F SRELK 1   SRE 0 CL SRELK 1	Camelus ferus	1	MINK	PFFE	IVI	VLLA	FST	IFSR	ELKI	PILI	FQQR	RVN	RES	LKLI	LNKL	RTS	SIQ	QCL	PHRK	NFL	LPQ	SMD	РНО	YQK	HILAI	80
Balamopiera acutorostanii   111 NNR FF FE IVLVLLASS SVC R ELKLVLF GOKKYNRES LKLLNKLOTS S IGOCLPHRKK NELL POKSMNPHOYG GOALA I 80     Bos sus srofa   111 NNR FF FE IVLVLLASS SVC R ELKLVLS GOKKYNRES LKLLNKLOTS S IGOCLPHRKK NELL POKSMNPHOYG GOALA I 80     Bos mutas   111 NNR FF FE IVLVLLASS SVC R ELKLVLS GOKKYNRES LKLLNKLOTS S IGOCLPHRKK NELP DKSMNPHOYG GOALA I 80     Bos mutas   111 NNR FF FE IVLVLLASS SVC R ELKLVLC GORKYNOE SLKLLNKLOTS SVG CLPHRKK NELP DKSMNPHOYG GOALA I 80     Ovis aries   111 NNR FF FE IVLVLLASS SVC R ELKLVLC GORKYNOE SLKLLNKLOTS SVG CLPHRKK NELP DKSMNPHOYG GOALA I 80     Ovis aries   111 NNR FF FE IVLVLLASS TVC R ELKLVLC GORKYNOE SLKLLNKLOTS SVG CLPHRKK NELP DKSMNPHOYG GOALA I 80     Ovis aries   111 NNR FF FE IVLVLLASS TVC R ELKLVLC GORKYNOE SLKLLNKLOTS SVG CLEHRKK NELP DKSMNPHOYG GOALA I 73     Macacamulata   111 NNR FF FE IVLVLLAST TFLELLKLL ILF GORKYNOE SLKLLNKLOTS SS IG CLEHRKK NELP DKSMNPHOYG GOALA I 73     Folis catus   111 NNR FF FE IVLVLLAST TFLELLKLIL IF GORKYNOE SLKLLNKLOTS SS IG CLEHRKK NELP DKSMNPHOYG GOALA I 73     Folis catus   111 NNR FF FE IVLVLLAST TFLELLKLIL IF GORKYNOE SLKLLNKLOTS SS IG CLEHRKK NELP DKSMNPHOYG GOALA I 73     Folis catus   111 NNR FF FE IVL KLAST TFLELLKLIL ILF GORKYNOE SLKLLNKLOTS SS IG CLEHRKK NELP DKSMNPHOYG GOALA I 73     Folis catus   111 NNR FF FE IVL KLAST TFLELLKLIL ILF GOKKYNOE SLKLLNKLOTS SS IG CLEHRKK NELP DKSMNPHOYG GOALA I 73     Folis catus   111 NNR FFFE IVL KLAST TFLELKK   111 FOK	Vicugna pacos	1	MINK	PFFE	IVI	VLLA	SST	VFSR	ELKI	PVL	FQQR	RVN	RES	LKLI	LNKL	RTS	SIQ	QCL	PHRK	NFL	LPQ	SVD	РНО	YQK	HILAI	80
Sus scroja   1   1   N N N A FF E I MU, VL LA S S O F R E L K, VL SQQ REV N R E S L KL L N LO D S I QQ C L PHR M FEL P D KS N PHOY QQ A LA BO Bos mutus   1   N N A FF E I VL VL LA S S I VC SQ E L K, VL CQQ REV N QE S L KL L N LQ D S S VQ G L PHR M FEL P D KS N PHOY QQ A LA BO O sis aris   1   N N A FF E I VL VL LA S S I VC SQ E L K, VL CQQ REV N QE S L KL L N LQ D S S VQ G L PHR M FEL P D KS N PHOY QQ A LA BO O sis aris   1   N N A FF E I VL VL LA S S I VC SQ E L K, VL CQQ REV N QE S L KL L N LQ D S S VQ G L PHR M FEL P D KS N PHOY QQ A LA BO O sis aris   1   N N A FF E I VL VL LA S S I VC SQ E L K, VL CQQ REV N QE S L KL L N LQ D S S VQ G C L H R KM FEL P D KS N PHOY QQ A LA BO N acaca mulata   1   N N A FF E I VL VL LA S S I FEL E K (VL CQQ REV N QE S L KL L N LQ D S S VQ G C L H R KM FEL P D KS N PHOY QG A LA BO N acaca mulata   1   N N A FF F E I VL VL LA S S I FEL E K (VL CQQ REV N QE S L KL N R LO T S I H E L D P KS N E POY VA G H H A BO Folic catus   1   1   N H A FF F E I VL VL LA S S I FEL E K (VL CQQ REV N R E S L KL N R LO T S I H E L D P KS N E POY VA G H H A BO P acaca mulata   1   1   1   1   0   1   0   1   0   1   0   1   0   1   0   1   0   1   0   1   0   1   0   1   0   1   0   1   0   1   0   1   0   1   0   1   0   1   0   <	Balaenoptera acutorostrata	1	MINK	PFFE	IVI	VLLA	SSS	VCSR	ELKI	LVL	FQQ <mark>K</mark>	RVN	RES	LKLI	LNKL	QTS	SIQ	QCL	PHWK	NFL	LPQ	<b>SMN</b>	РНО	YQK	QALAI	80
Bos mutas   1 MIN NA FFE IVLVLLAYSTVC NO EL KLVVL COORNYNOE SLKLL NKLDTSSVOG CLPHRK HFLLPORSVNPHOYD KOQVLA 180     Ovis aries   1 MIN NA FFE IVLVLLASSTVC NO EL KLVVL COORNYNOE SLKLL NKLDTSSVOG CLPHRK HFLLPORSVNPHOYD KOQVLA 180     Gauss caballus   1 MIN NA FFE IVLVLLASSTVC NO EL KLVVL COORNYNOE SLKLL NKLDTSSVOG CLPHRK HFLLPORSVNPHOYD KOQVLA 180     Maccacamulata   1 MIN NA FFE IVLVLLASSTVC NO EL KLVVL COORNYNNE SLKLL NKLDTSSI OC CLPHRK NFLLPORSVNPHOYD KOQVLA 180     Maccacamulata   1 MIN KHFFE INLVLLASSTVF FD LLKLI IF FOORNYN RE SLKLL NKLDTSSI HC CPHRK NFLLPORSVNPHOYD KOQVLA 180     Promo sapiens   1 MIN KHFFE INLVLLASSTVF FD LLKLI IF FOORNYN RE SLKLL NKLDTLSI OG CLPHRK NFLLPORSVNPHOYD KOQVLA 180     Polo   100   110   120   130   140   150     Camelus dromedarius   81   LHEML OO TFRUE FRAVISL DGWE EI OND FFL SEL HOU EYLET LI RLOAFORS GILGSEN LHOVK SY FOR HOVLESOEY 160     Camelus dromedarius   81   HEML OO TFRUE FRAVISL DGWE EI OND FFL SEL HOU EYLET LI RLOAFORS GILGSEN LHOVK SY FOR HOVLESOEY 160     Camelus dromedarius   81   HEML OO TFRUE FRAVISL DGWE EI OND FFL SEL HOU EYLET LI RLOAFORS GILGSEN LHOVK SY FOR HOVLESOEY 160     Sas crofa   81   HEML OO TFRUE FRAVISL DGWE EI OND FFL SEL HOU EYLET LI RLOAFORS GILGSEN LHOVK SY FOR HOVLESOEY 160     Sas crofa   81   HEML OO TFRUE FRAVISL DGWE END HOW ELSEL HOU EYLET LI ROAFORS GILGSEN LHOVK SY FOR HOVLESOEY 160	Sus scrofa	1	MINKS	SFFE	IM	VLLA	SST	GFSR	ELKI		SQQR	RVN	RES		LNKL	QTS	SIQ	QCL	PHRK	NFL	LPQ	SMN	рна	YQK	QALAI	80
Bostaurus   1 MIN NA FFEIVLVLLASSTVCKO ELKUVLCQORRYNOESUKLLNKLDTSSV0QCLPHRKHFLLPOKSVNPHOYGKGOVLA 180     Ovisuris   1 MIN NA FFEIVLVLLASSTVCKO ELKUVLCQORRYNOESUKLLNKLDTSSV0QCLPHRKHFLLPOKSVNPHOYGKGOVLA 180     Equis caballus   1 MIN KA FFEIVLVLLASSTVCKO ELKUVLCQORRYNOESUKLLNKLDTSSINCOCLPHRKHFLLPOKSVNPHOYGKGOVLA 180     Paus caballus   1 MIN KA FFEIVLVLLASSTVCKO ELKUVLFOQRKYNOESUKLLNKLDTSSINCOCLPHRKHFLLPOKSVNPHOYGKGOVLA 180     Preizeatus   1 MIN KHFFAIVLSVLLASSTIFFE ELKUVLFOQRKYNOESUKLLNKLDTSSINCOCLPHRKNFLLPOKSVNPHOYGKGOVLA 179     Jono sagiens   1 MIN KHFFAIVLSVLLASSTIFFE ELKULIFFOQROVNOESUKLLNKLDTLSIOQCLPHRKNFLLPOKSVNPHOYGKGOVLA 179     Jono sagiens   1 MIN KHFFAIVLSULASSTIFFE ELKULIFFOQROVNOESUKLLNKLDTLSIOQCLPHRKNFLLPOKSVNPHOYGKGOVLA 179     Jono sagiens   1 MIN KHFFAIVLSULASSTIFFE ELKULIFFOQROVNOESUKLLNKLDTLSIOQCLPHRKNFLLPOKSVNPHOYGKGOVLA 179     Jono sagiens   1 MIN KHFFAIVLSULASSTIFFE ELKULIFFE ELKOLEYLETLIKLOTSION   140   150     Camelus forms   81   LHEMLOG IFNLFRAVISLDGWEE IOMD RFLSELHOOLEYLETLIKLOTSIG EGEN RLOVKYFOR HOYLESGEY 160   160     Galaenopera acutorostrus   81   LHEMLOG IFNLFRAVISLDGWEE IOMD RFLSELHOOLEYLETLIKLO QLASGRSGILGSEHLRUVKYFOR HOYLESGEY 160   160     Sas scrigt   81   LHEMLOG IFNLFRAVISLDGWEE IOMD RFLSELHOOLEYLETLIKLO QLASGRSGILGSEHLRUVKYFOR HOYLESGEY 160   160   170   170   180   170   170   180 <td>Bos mutus</td> <td>1</td> <td></td> <td>AFFE</td> <td>IVI</td> <td>VLLA</td> <td>YST</td> <td>vcsa</td> <td>ELKI</td> <td>LVL</td> <td>CQQR</td> <td>RVN</td> <td>DES</td> <td></td> <td></td> <td>QTS</td> <td>sva</td> <td>QCL</td> <td>PHRK</td> <td>HFL</td> <td>LPQ</td> <td>SVN</td> <td>рна</td> <td>YQK</td> <td>GOVLAI</td> <td>80</td>	Bos mutus	1		AFFE	IVI	VLLA	YST	vcsa	ELKI	LVL	CQQR	RVN	DES			QTS	sva	QCL	PHRK	HFL	LPQ	SVN	рна	YQK	GOVLAI	80
Ovis arise   1 MIL NA FFFE IVILUL ASST VOSO EL NEUL COORREY UNES EN LLINKLOTS IN OCCLEHRKY HELPOKS VNPHOVOK GOLLA 1800     Macca nulata   1 MIL NA FFFE IMLUL ASST IVIS EN LEUL VLFOORREY UNES EN LLINKLOTS IN OCCLEHRKY HELPOKS UNPHOVOK GOLLA 1800     Macca nulata   1 MIL NA FFFE IMLUL ASST IVIS EN LEUL VLFOORREY UNES EN LLINKLOTS SINL CLPHRKY FELPOKS UNPHOVOK GOLLA 1800     Macca nulata   1 MIL NA FFFE IMLUL ASST IVIS EN LA LIFOORREY UNES EN LLINKLOTS SINL CLPHRKY FELPOKS UNPHOVOK GOLLA 179     Folis catus   1 MIL NA FFFE IMLUL ASST IVIS EN LA LIFOORREY UNES EN LLINKLOT EN CORPT NESS IS DOVOK GH LA 1800     90   100   110   120   130   140   150     Camelus dromedarius   81   LHENL GOT FNUE FRAVI SLOGWE FLOMDEFF SEL MOLE EVICET I RUA AFORS GLOS SEN RELOVINS VERON HOVE ES CEVE 160     Scanapazon   81   LHENL GOT FNUE FRAVI SLOGWE FLOMDEFF SEL MOLE EVICET I RUA AFORS GLOS SEN RELOVINS VERON HOVE ES CEVE 160     Scanapazon   81   LHENL GOT FNUE FRAVI SLOGWE FLOMDEFF SEL MOLE EVICET I RUA AFORS GLOS SEN RELOVINS VERON HOVE ES CEVE 160     Scanapazon   81   LHENL GOT FNUE FRAVI SLOGWE ES HERE FLOVE LA CUE VEL LA LA RUA OA SCANA RES SLOGUE RES CLOS SEN RELOVINS VERON HOVE ES CEVE 160     Sas scróa   81   LHENL GOT FNUE FRAVI SLOGWE ES HERE FLOVE LA CUE VEL LA LA RUA OA SCANA RES DI L CS EN RELOVINS VERON HOVE ES CEVE 160     Sas scróa   81   LHENL GOT FSLF RA VIS SLOGWE E	Bos taurus	1	MINK	AFFE	IVI	VLLA	SST	vcsa	ELKI	LVL	CQQR	RVN	QES			QTS	sva	QCL	PHRK	HFL	LPQ	SVN	рна	YQK	GOVLAI	80
Equas caballas   1 MI NG FFE IMULLASSTIFS LELALVI.FOORRYNNESS LILLNRLOTSSTIG COLSHRRMFLEPORSVMPGVOKRHLAB     Macacanulatia   1 MI NG FFE IMULLASSTIFS LELALVI.FOORRYNNESS LILLNRLOTSSTIG COLPHRKMFLEPORSVMPGVOKRHLAB     Pdis catas   1 MI NG HFFE IMULLASSTIFF LELALAFOOR.VNRESS LILLSTLOSSS IO COLPHRKMFLEPORSVMPGVOKRGALAB     90   100   110   120   130   140   150     Camelas dromedarias   11 LHENG OF FULFULASSTIFS LELALVIFOR NORES LILLNKLOTSS LOC CLPHRKMFLEPORS VSPOR VKGGALAB   79     Idenospras   11 LHENLOOI FULFFRAVIS LOMME FI SELHOOLEYLETI RLOAGONS GI CSENN RLOVK SYFOR IHDYLESOEY 160     Camelas dromedarias   11 LHENLOOI FNLFFRAVIS LOMME FI SELHOOLEYLETI RLOAGONS GI CSENN RLOVK SYFOR IHDYLESOEY 160     Balanopera acuterostraus   11 LHENLOOI FNLFFRAVIS LOMME FI SELHOOLEYLETI RLOAGONS GI CSENN RLOVK SYFOR IHDYLESOEY 160     Balanopera acuterostraus   11 LHENLOOI FNLFFRAVIS LOMME FI SELHOOLEYLEALMEN CALON SYFOR IHDYLESOEY 160     Balanopera acuterostraus   11 LHENLOOI FNLFFRAVIS LOMME FI SELHOOLEYLEALMEN CALON SOTI CSENN RLOVK SYFOR IHDYLESOEY 160     Bornauta   11 LHENLOOI FNLFFRAVIS LOMME FI SELHOOLEYLEALMEN CALON SOTI CSENN RLOVK SYFOR IHDYLESOEY 160     Bornauta   11 LHENLOOI FNLFFRAVIS LOMME FI SELHOOLEYLEALMEN CALON SOTI CSENN RLOVK SYFOR IHDYLESOEY 160     Bornauta   11 LHENLOOI FNLFFRAVIS LOMME SINCE FI SELHOULEYLEALMEN CALON SOTI CSENN RLOVK SYFOR IHDYLENDIY	Ovis aries	1	MINK	AFFE	IVI	VLLA	SST	vcsa	ELKI		CQQR	RVN	QES	LKL		QTS	SIQ	QCL	LHRK	NFL	LPQ	SVN	рна	YQK	QVLAI	80
Macaca mulata   1 MIL NIT FEI INUULI ASTUFFLOLIKUL FOORKUNNESS KULLINKUGUSS IN CLUPHENNELL PORSUNPROVING ALA 199     Feliscatas   1 MIL KHFFGTVUVULASTUFFLOLIKUL FOORVINGESKULLINKUGUSS IN CCUPHENNELL PORSUNPROVING ALA 179     Homo sapiens   1 MIL KHFFGTVUVULASTUFFSLOLIKULI FOORVINGESKULLINKUGUSS IN CCUPHENNELL PORSUNPROVING ALA 179     Canelus dromedarius   51 LHEML OO IFNLFRAVISLOGWEE I OMDRELSELHOOLEYLETLIR LOAFEORSGIL GSENL RUOVKSYFOR INDULESGEY 160     Canelus dromedarius   51 LHEML OO IFNLFRAVISLOGWEE I OMDRELSELHOOLEYLETLIR LOAFEORSGIL GSENL RUOVKSYFOR INDULESGEY 160     Canelus dromedarius   51 LHEML OO IFNLFRAVISLOGWEE I OMDRELSELHOOLEYLETLIR LOAFEORSGIL GSENL RUOVKSYFOR INDULESGEY 160     Sisserofa   81 LHEML OO IFNLFRAVISLOGWEE I OMDRELSELHOOLEYLETLIR LOAFEORSGIL GSENL RUOVKSYFOR INDULESGEY 160     Balaenopterna acutorostata   81 LHEML OO IFNLFRAVISLOGWEE I OMDRELSELHOOLEYLETLIR LOAFEORROTIG GSENL RUOVKSYFOR INDULESOEY 160     Bostaurus   81 LHEML OO IFNLFRAVISLOGWEE SHTEKEK LVELHOOLEYLEAL MRLOAFEORROTIG GSENL RUOVKSYFOR INDULESOEY 160     Bostaurus   81 LHEML OO IFSLERANIVSLO DOWEE SHTEKEK LVELHOOLEYLEAL MRLOAKS STILG GSENL RUOVKMYFOR INDULESOEY 160     Bostaurus   81 LHEML OO IFSLERANIVSLO DOWEE SHTEKEK LVELHOOLEYLEAL MRLOAKS STILG GSENL RUOVKMYFOR INDULESOEY 160     Oris aries   81 LHEML OO IFSLERANISLO DOWEE SHTEKEK LVELHOULEYLEAL MRLOAKS STILG GSENL RUOVKMYFOR INDULESOEY 160     Marcaca mulata   81 LHEML OO IFNLFRAVISLO DOWEE SHTEKEK LVELHOULEYLEAL	Equus caballus	1	MINKO	2 F F E	IM	VLLA	SSI	IFSL	ELKI	LVL	FQQR	RVN	RES	LKL		QPS	AIQ	QCL	SHRR	NFL	LPQ	SVN	РНО	YQK	RHALAI	80
Feliscatus   1	Macaca mulatta	1	MIK	IFFE	IM	VLLA	SII	VFFL	DLK		FQQK	RVN	DES			QIS	SIH	LCL	PHRK	NFL	LPQ	SLS	PQC	YQK	HILAI	80
Homo sagiens   Image: Marker Given Lasser Figure Link Link Girls and Link Link Girls and Link Link Girls and Link Construction Link Girls and Link Construction Link Girls and Link Construction Link Constructin Link Construction Link Construction Link Construction Link Const	Felis catus	1			IV	VLLA	SSI	TEPL	ELK	LAL	FQQR	- VNI	KES		STL	QSS	SIQ	QCL	PHRK	NFL	LPQ	SVN	PRO	YQKO	QALAI	79
90   100   110   120   130   140   150     Camelus dromedarius   81   L HEML 00 I FNL FRAVI SL DGWE E I OMDR FL SEL HOOL EYL ETL IRL 0 AE ORS GIL GSENL RL 0VK SY FOR I HDYL E SOEY 160     Camelus forus   81   L HEML 00 I FNL FRAVI SL DGWE E I OMDR FL SEL HOOL EYL ETL IRL 0 AE ORS GIL GSENL RL 0VK SY FOR I HDYL E SOEY 160     Vicugna paco   81   L HEML 00 I FNL FRAVI SL DGWE E I OMDR FL SEL HOOL EYL ETL IRL 0 AE ORS GIL GSENL RL 0VK SY FOR I HDYL E SOEY 160     Salaenoptera acutorotrus   81   L HEML 00 I FNL FRAVI SL DGWE E TIME K FL IE LHOOL EYL E AL MRL 0 AE OR SO TL GSENL RL 0VK SY FOR I HDYL E SOEY 160     Sus scrofa   81   L HEML 00 I FNL FRAVI SL DGWE E SHTEK FL VEL HOOL EYL E AL MRL 0 AE OR SO TL GSENL RL 0VK SY FOR I RDYL E NOOY 160     Sus scrofa   81   L HEML 00 I FS L FRAVI SL DGWE E SHTEK FL VEL HOOL EYL E AL MRL 0 AK 0R SO TL GSENL RL 0VK MY FOR I RDYL E NOOY 160     Bos mutus   81   L HEML 00 I FNL FRAVI SL DGWE E SHTEK FL VEL HOOL EYL E AL MRL 0 AK 0R SO TL GSENL RL 0VK MY FOR I HDYL E SOOY 160     Bos mutus   81   L HEML 00 I FNL FRAN SU DGWE E SHTEK FL VEL HOOL EYL E AL MRL 0 AK 0R SO TL GSENL RL 0VK MY FOR I HDYL E SOOY 160     Carsea subaltas   81   L HEML 00 I FNL FRAN SS DGWE E SHTEK FL VEL LOOL EYL E AL MRL 0 AK 0R SO TL GSENL RL 0VK MY FOR I HDYL E SOOY 160     Fausacas   80   L HEML 00 I FNL FRAN SS DGWE E SHTEK FL VEL LOOL EYL E AL	Homo sapiens	1	MIII	1		VLLA	SII	TESL			FQQR	QVN	JES			QIL	SIQ	QCL	PHRK	NFL	LPQ	SL S	PQC	YQK	HLAI	80
90100110120130140150Camelus formedarius31L HEML DO I FNL F RAV I SLD GWE E I OMD RFL SEL HOOL EYL ET LI RL OA FOR SG IL G SENL RL OVKSY FOR I HDYL E SOEY 160Vicugna pacos81L HEML OO I FNL F RAV I SLD GWE E I OMD RFL SEL HOOL EYL ET LI RL OA FOR SG IL G SENL RL OVKSY FOR I HDYL E SOEY 160Balaenoptera acutorostrati81L HEML OO I FNL F RAV I SLD GWE E I OMD RFL SEL HOOL EYL ET LI OL OA FOR SG IL G SENL RL OVKSY FOR I HDYL E SOEY 160Balaenoptera acutorostrati81L HEML OO I FNL F RAV I SLD GWE E SIME FL OVL EYL E ALMRL OA AE G K RD IL G SENL RL OVKSY FOR I HDYL E SOEY 160Bos staurus81L HEML OO I F SL F RA I V SLD GWE E SIM E FL VEL HOOL EYL E ALMRL OA AE G K RD IL G SENL RL OVKMY FOR I RDYL E NODY 160Bos staurus81L HEML OO I F SL F RA I V SLD GWE E SIM E FL VEL HOUL EYL E ALMRL OA K K SD IL G SENL RL OVKMY FOR I HDYL E SOEY 160Ovis aries81L HEML OO I F NL F RA I S SLD GWE E SIM E FL VEL HOUL EYL E ALMRL OA K K SD IL G SENL RL OVKMY FOR I HDYL E SOEY 160 <i>Gausa</i> mulatin81L HEML OO I F NL F RA N I SLD GWE E SIM E FL VEL LOOL EYL E AL MRL OA K K SD IL G SENL RL OVKMY FOR I HDYL E SOEY 160 <i>Gausa</i> mulatin81L HEML OO I F NL F RA N I SLD GWE E SIM E FL U EL OU LEYL E AL MG LE AA EKL SG TL G SD NL RL OVKMY FOR I HDYL E SOEY 160 <i>Gausa</i> solaris81L HEML OO I F NL F RA N I SLD GWE E SIM E FL I EL HOOL EYL E AL MG LE AA EKL SG TL G SD NL RL OVKMY FOR I HDYL E SOEY 160 <i>Felis</i> catus80L HEML OO I F NL F RA N I SLD GWE E SIM E FL I OL HOUL EYL E AL MG LE AA EKL SG TL G SD NL RL OVKMY FOR I HDYL E SOEY 160 <i>Felis</i> catus81L HEML OO I F NL F RA N																										
Camelus dromedarius   81   L HEML QQ I F N L F R AVI S L D GWE E I OMDR F L SE L HQQL EYL ET L I R L QA E OR SG I L GSE N L R L QVK SY FOR I HDYL ESOEY 160     Camelus forus   81   L HEML QQ I F N L F R AVI S L D GWE E I OMDR F L SE L HQQL EYL ET L I R L QA E OR SG I L GSE N L R L QVK SY FOR I HDYL ESOEY 160     Vicengna pacos   81   L HEML QQ I F N L F R AVI S L D GWE E I OMDR F L SE L HQQL EYL ET L I R L QA E OR SG I L GSE N L R L QVK SY FOR I HDYL ESOEY 160     Balaenoptera acutrostrata   81   L HEML QQ I F N L F R AVI S L D GWE E TIMBE Y FL I EL HQQL EYL E AL MR L QA E OR SG I L GSE N L R L QVK SY FOR I HDYL ESOEY 160     Sus scrofa   81   L HEML QQ I F SL F R AVI SL D GWE E SHT BE Y FL Y EL HQQL EYL E AL MR L QA E OR SD I L GSE N L R L QVK MY FOR I HDYL ESOEY 160     Bos mutus   81   L HEML QQ I F SL F R A I VSL D GWE E SHT E Y FL Y EL HQQL EYL E AL MR L QA K OK SD T L GSE N L R L QVK MY FOR I HDYL ESODY 160     Ovis aries   81   L HEML QQ I F N L F R AN SL D GWE E SHT E Y FL Y EL HQQL EYL E AL MR L QA K K SD T L GSE N L R L QVK MY FOR I HDYL E SODY 160     Ovis aries   81   L HEML QQ I F N L F R AN I SL D GWE E SHT E Y FL Y EL L QQ L EYL E AL MR L QA K K SD T L GSE N L R L QVK MY FOR I HDYL E SODY 160     Ovis aries   81   L HEML QQ I F N L F R AN I SS D GWE E SHT E Y FL Y E L L QQ L EYL E AL MR L QA K K SD T L GSE N L R L QVK MY FOR I HDYL E SODY 160     Proto sagiens   81   L HEML QQ I F N L F R AN I SS D GWE E SHT E Y FL Y E L L QQ L EYL E AL MG L E AF					90			100			110			120			130			140			150	)		
Camelus ferus   81   LHEMLQQIFNLFRAVISL DGWE EIOMDRFLSELHOQLEYLETLIRLOAEORSGILGSENLRLOVKSYFORIHDYLESOEY 160     Balaenoptera acutrostatu   81   LHEMLQQIFNLFRAVISL HGWE EIOMDRFLSELHOQLEYLETLIRLOAFORSGILGSENLRLOVKIYFORIHDYLESOEY 160     Salaenoptera acutrostatu   81   LHEMLQQIFSLFRAVISL HGWE ESIME EFLUELHOQLEYLETLIRLOAFORSGILGSENLRLOVKIYFORIHDYLESOEY 160     Sans acutrostatu   81   LHEMLQQIFSLFRAVISL DGWE ESIME EFLUELHOQLEYLE ALMRLOAFOK SDILGSENLRLOVKIYFORIHDYLESOEY 160     Bos mutus   81   LHEMLQQIFSLFRAVISL DGWE ESIME EFLUELHOQLEYLE ALMRLOAFOK SDILGSENLRLOVKIYFORIHDYLESOEY 160     Bos mutus   81   LHEMLQQIFSLFRAVISL DGWE ESIME FKIVELHOQLEYLE ALMRLOAFOK SDILGSENLRLOVKIYFORIHDYLESOEY 160     Bos mutus   81   LHEMLQQIFSLFRANISL DGWE ESIME FKIVELHOQLEYLE ALMRLOAFOK SDILGSENLRLOVKIYFORIHDYLESOEY 160     Bos mutus   81   LHEMLQQIFSLFRANISL DGWE ENHE KFLVELHOQLEYLE ALMRLOAFOK SDILGSENLRLOVKIYFORIHDYLESOEY 160     Equus cabilus   81   LHEMLQQIFSLFRANISL DGWE ENHE KFLVELHOQLEYLE ALMGLE AEKL SGILGSENLRLOVKIYFORIHDYLESOEY 160     Macaca mulatu   81   LHEMLQQIFSLFRANISL DGWE ENHE KFLVELHOQLEYLE ALMGLE AEKL SGILGSENLRLOVKIYFORIHDYLESOEY 160     Falis catus   80   LHEMLQQIFSLFRANISL DGWE ENHE KFLVELHOULEYLE ALMGLE AEKL SGILGSENLRLOVKIYFORIHDYLESOEY 159     Homo sapiens   81   LHEMLQQIFSLFRANISL DGWE ENHME KFLIQLHOQLEYLE ALMGLE AEKL SGILGSENLRLOVKIYFORIHDYLESOEY 159	Camelus dromedarius	81		001	ENI	FRAV	ISL	DGWE	EIO	MDR	FLSE	LHO		YLE	LIB		EOR	SGI		NLR		SYF	ORI	HDYI	ESOEY	160
Vicugnapacos   81   LHEMLOQIFNLFRAVISLHGWE IOMDRFLSELHOQLEYLETLIOLOAGORSGTLGSEHLRLOVKSYFORIHDYLENDY   160     Balaenoptera acutorostrata   81   LHEMLOQIFNLFRAVISLHGWE THMEKFLIELHOQLEYLEALMRLOAGOKRDTLGSENLRLOVKIYFORIRDYLENDY   160     Susscrig   81   LHEMLOQIFSLFRAVISL   GGWE ESHTEKFLIVELHOQLEYLEALMRLOAGOKRDTLGSENLRLOVKMYFORIRDYLENDY   160     Bostaurus   81   LHEMLOQIFSLFRAVISL   GGWE ESHTEKFLIVELHOQLEYLEALMRLOAKOKDTLGSENLRLOVKMYFORIHDYLESODY   160     Ovis aries   81   LHEMLOQIFSLFRAVISL   GGWE ESHTEKFLIVELHOQLEYLEALMRLOAKOKDTLGSENLRLOVKMYFORIHDYLESODY   160     Ovis aries   81   LHEMLOQIFSLFRAVISL   GGWE ESHTEKFLIVELHOQLEYLEALMRLOAKOKSDTLGSENLRLOVKMYFORIHDYLESODY   160     Acacaa mulata   81   LHEMLOQIFSLFRANISLOGWE ESHTEKFLIVELHOQLEYLEALMRLOAKOKSDTLGSENLRLOVKMYFORIHDYLESODY   160     Felis catus   80   LHEMLOQIFSLFRANISLOGWE ESHTEKFLIVELHOQLEYLEALMGLEAECHCGPLGSENLRLOVKMYFORIHDYLESODY   160     Felis catus   80   LHEMLOQIFSLFRANISLOGWE ESHTEKFLIVELHOQLEYLEALMRLOAKKOSDTLGSENLRLOVKMYFORIHDYLESODY   160     Felis catus   80   LHEMLOQIFSLFRANISLOGWE ESHTEKFLIVELHOQLEYLEALMGLEAECHCGPLGSENLRLOVKMYFORIHDYLESODY   160     Felis catus   80   LHEMLOQIFSLFRANISLOGWE ESHTEKFLIVELHOQLEYLEALMGLEAECHCGPLGSENLRLOVKMYFORIHDYLESODY	Camelus ferus	81	LHEMI	001	FNI	FRAV	ISL	DGWE	EIQ	MDR	FLSE	LHQ	DLE	YLE	TLIR	LQA	EOR	SGI	LGSE	NLR	LOV	SYF	ORI	HDYI	ESQEY	160
Balemontera acutorostrata   81   L HEMLQQ I FNL FRAIIISL NGWEETHMEEKFLVEL HOQLEYLEALMRLQAEQKSDTLGSENL RLQVKMYFOR IRDYLENDY     Susserofa   81   L HEMLQQ I FSL FRAIVISLD GWEESHMEEFLVEL HOQLEYLEALMRLQAEQKSDTLGSENL RLQVKMYFOR IRDYLENDY     Bosnutus   81   L HEMLQQ I FSL FRAIVSLD GWEESHMEEFLVEL HOQLEYLEALMRLQAKGKSDTLGSENL RLQVKMYFOR IRDYLENDY     Bostaurus   81   L HEMLQQ I FSL FRAIVSLD GWEESHTEKFLVEL HOQLEYLEALMRLQAKGKSDTLGSENL RLQVKMYFOR I HDYLESODY     Bostaurus   81   L HEMLQQ I FNL FRATSSLD GWEESHTEKFLVEL HOQLEYLEALMRLQAKKSDTLGSENL RLQVKMYFOR I HDYLESODY     Govis arise   81   L HEMLQQ I FNL FRATSSLD GWEESHTEKFLVEL HOQLEYLEALMRLQAKKSDTLGSENL RLQVKMYFOR I HDYLESODY <i>Equus caballus</i> 81   L HEMLQQ I FNL FRAN I SLD GWEESHMET FLI EL HOQLEYLE AL MGLE AEGNCGPL GSENL RLQVKMYFOR I HDYLE SODY <i>Macaca mulatta</i> 81   L HEMLQQ I FNL FRAN I SLD GWEESHMET FLI DL HOQLEYLE AL MGLE AEGNCGPL GSENL RLQVKMYFOR I HDYLE NOEY <i>Macaca mulatta</i> 81   L HEMLQQ I FSL FRAN I SLD GWEESHMET FLI DL HOQLEYLE AL MGLE AEKLSGTL GSDNL RLQVKMYFOR I HDYLE NOEY <i>Macaca mulatta</i> 81   L HEMLQQ I FNL FRAN I SLD GWEE NHEK FLI DL HOQLEYLE AL MGLE AEKLSGTL GSDNL RLQVKMYFOR I HDYLE NOEY <i>Macaca mulatta</i> 81   LHEMLQQ I FSL FRAN I SLD GWEE NHEK FLI DL HOQLEYLE AL MGLE AEKLSGTL GSDNL RLQVKMYFOR I HDYLE NOEY <i>Too</i> 180   190   200 <tr< td=""><td>Vicugna pacos</td><td>81</td><td></td><td>001</td><td>FNI</td><td>FRAV</td><td>ISL</td><td>HGWE</td><td>EIQ</td><td>MDR</td><td>FLSE</td><td>LHQ</td><td>DLE</td><td>YLE</td><td>TLIQ</td><td>LQA</td><td>EOR</td><td>SGT</td><td>LGSE</td><td>HLR</td><td>LOV</td><td>SYF</td><td>ORI</td><td>HDYL</td><td>ESQEY</td><td>160</td></tr<>	Vicugna pacos	81		001	FNI	FRAV	ISL	HGWE	EIQ	MDR	FLSE	LHQ	DLE	YLE	TLIQ	LQA	EOR	SGT	LGSE	HLR	LOV	SYF	ORI	HDYL	ESQEY	160
Sus scrofa   81   LHEMLOQIFSLFRAVISLDGWEESHMEEFLVELHOQLEYLEALMRLOAKOKSDTLCSENLTLOVKMYFORIRDYLENDDY   160     Bos mutus   81   LHEMLOQIFSLFRAVISLDGWEESHTEKFLVELHOQLEYLEALMRLOAKOKSDTLCSENLRLOVKMYFORIHDYLESODY   160     Ovis aries   81   LHEMLOQIFSLFRAVISLDGWEESHTEKFLVELHOQLEYLEALMRLOAKOKSDTLGSENLRLOVKMYFORIHDYLESODY   160     Ovis aries   81   LHEMLOQIFSLFRAVISLDGWEESHTEKFLVELHOQLEYLEALMRLOAKKSDTLGSENLRLOVKMYFORIHDYLESODY   160     Equus caballus   81   LHEMLOQIFSLFRANISLDGWEESHMEFFLIELHOQLEYLEALMRLOAKKSDTLGSENLRLOVKMYFORIHDYLESODY   160     Macaca mulata   81   LHEMLOQIFSLFRANISLDGWEESHMEFFLIELHOQLEYLEALMRLOAKKSDTLGSENLRLOVKMYFORIHDYLESODY   160     Felis catus   80   LHEMLOQIFSLFRANISLDGWEESHMEFFLIELHOQLEYLEALMGLEAEKLSGTLGSDNLRLOVKMYFORIHDYLESOPY   160     Felis catus   81   LHEMLOQIFSLFRANISLDGWEESHVEKFLITELHOQLEYLEALMGLEAEKLSGTLGSDNLRLOVKMYFORIHDYLESOPY   190     Felis catus   81   LHEMLOQIFSLFRANISLDGWEENHTEKFLIGLHOQLEYLEALMGLEAEKLSGTLGSDNLRLOVKMYFORIHDYLESOPY   193     Camelus dromedarius   161   SSCAWTIVQIENCLFFMIQLTGKLSKQGMDP   190   200   193     Camelus dromedarius   161   SSCAWTIVQIEINRCLFFMIQLTGKLSKQGMDP   193   193   193   193   193   193   193	Balaenoptera acutorostrata	81	LHEML	QQI	FNI	FRAI	I SL	N GWE	ETH	MEK	FLIE	LHQ	LK	YLE	ALMR	LQA	EQK	RDT	LGSE	NLR	LQV	IYF	QRI	RDYI	ENQDY	160
Bos mutus   81   LHEMLQQIFSLFRAIVSLDGWEESHTEKFLVELHQQLEYLEALMRLQAKQKSDTLGSENLRLQVKMYFQRIHDYLESQDY   160     Bos taurus   81   LHEMLQQIFSLFRAIVSLDGWEESHTEKFLVELHQQLEYLEALMRLQAKGKSDTLGSENLRLQVKMYFQRIHDYLESQDY   160     Ovis aries   81   LHEMLQQIFNLFRANSLDGWEESHTEKFLVELLQQLEYLEALMRLQAKKKSDTLGSENLRLQVKMYFQRIHDYLESQDY   160     Equus caballus   81   LHEMLQQIFNLFRANSLDGWEESHTEKFLVELLQQLEYLEALMGLEAEGKCGPLGSENLRLQVKMYFQRIHDYLESQDY   160     Macaca mulatta   81   LHEMLQQIFSLFRANSLDGWEESHTEKFLVELLQQLEYLEALMGLEAEKLSGTLGSDNLRLQVKMYFQRIHDYLESQDY   160     Macaca mulatta   81   LHEMLQQIFSLFRANSLDGWEESHTEKFLTELHQQLEYLEALMGLEAEKLSGTLGSDNLRLQVKMYFQRIHDYLENODY   160     Macaca mulatta   81   LHEMLQQIFSLFRANSLDGWEESHTEKFLTELHQQLEYLEALMGLEAEKLSGTLGSDNLRLQVKMYFQRIHDYLENODY   160     Macaca mulatta   81   LHEMLQQIFSLFRANSLDGWEESHTEKFLTELHQQLEYLEALMGLEAEKLSGTLGSDNLRLQVKMYFQRIHDYLENODY   160     Camelus formo sapiens   81   LHEMLQQIFSLFRANSLDGWEESHTEKFLTELHQULEYLEALMGLEAEKLSGTLGSDNLRLQVKMYFQRIHDYLENODY   160     Camelus forus   161   SSCAWTIVQIFENRCLFFWIQLTGKLSKQGMDP   193   193     Camelus forus   161   SSCAWTIVQIFENRCLFFWIQLTGKLSKQGMDP   193   193   193   193   193     Salaenoptera acutorostr	Sus scrofa	81	LHEML	001	FSI	FRAV	ISL	DGWE	ESH	MEE	FLVE	LHQ	LE	YLE	ALMR		EQK	SDT	LCSE	NLT	LQV	MY F	QRI	RDYI	ENQDY	160
Bostaurus   81   L HEML QQ I F SL F RAIVSL DGWE E SHTE K FLVEL HQQL E YLE ALMRL QAKQKS DT LGSENL RLQVKMY FOR I HDYLE SQDY 160     Ovis aries   81   L HEML QQ I F NL F RATSLD GWE E SHTE K FLVEL LQQ LE YLE ALMRL QAKKKS DT LGSENL RLQVKMY FOR I HDYLE SQDY 160     Macaca mulatta   81   L HEML QQ I F NL F RATSLD GWE E SHTE K FLVEL LQQ LE YLE ALMGL E AEKL SGTL GSENL RLQVKMY FOR I HDYLE SQDY 160     Macaca mulatta   81   L HEML QQ I F SL F RATSLD GWE E SHME K FL I QL HQQ LE YLE ALMGL E AEKL SGTL GSENL RLQVKMY F R I HDYLE SQDY 160     Felis catus   80   L HEML QQ I F SL F RATSLD GWE E SHVE K FL T ELH QQ LE YLE ALMGL E AEKL SGTL GSENL RLQVKMY F R I HDYLE SOEY 159     Homo sapiens   81   L HEML QQ I F SL F RATSLD GWE E SHVE K FL T ELH QU LE YLE ALMGL E AEKL SGTL GSENL RLQVKMY F R I HDYLE SOEY 159     Homo sapiens   81   L HEML QQ I F SL F RATSLD GWE E SHVE K FL T ELH QU LE YLE ALMGL E AEKL SGTL GSENL RLQVKMY F R I HDYLE NOEY 160     Felis catus   80   L HEML QQ I F SL F RATSLD GWE E SHVE K FL T ELH QU LE YLE ALMGL E AEKL SGTL GSENL RLQVKMY F R I HDYLE NOEY 160     Felis catus   161   SS C AWT I VQ I E I NR CL F F MI QL T GKL SKQ GMD P   193     Camelus dromedarius   161   SS C AWT I VQ I E I NR CL F F MI QL T GKL SKQ GME P   193     Salaenoptera acutorostrata   161   SS C AWT I VQ VE I NR CL F F V F RL T GKL SKQ GME P   193     Sos nutus	Bos mutus	81	LHEML	QQI	FSI	FRAI	VSL	D GWE	ESH	TEK	FLVE	LHQ	LE	YLE	ALMR	LQA	KOK	SDT	LGSE	NLR	LQV	MY F	QRI	HDYL	ESQDY	160
Ovis aries   81   LHEMLQQ IFNLFRATSSLDGWESHMETFLFKT.VELLQQLEYLEALMRLQAKKKSDTLGSENLRLQVKMYFREIHDYLESQDY   160     Equus caballus   81   LHEMLQQ IFNLFRAN IPLDAWEESHMETFLIELHQQLEYLEALMGLEAEGM CGPLGSENLRLQVKMYFREIHDYLESQDY   160     Macaca mulatu   81   LHEMLQQ IFNLFRAN ISLDGWEESHMETFLIELHQQLEYLEALMGLEAEKLSGTLGSENLRLQVKMYFREIHDYLESQDY   160     Felis catus   80   LHEMLQQ IFSLFRAN ISLDGWEESHVEKFLIELHQQLEYLEELTGPEAEQDSCILGSENVRLQIKMYFREIHDYLESQEY   159     Homo sapiens   81   LHEMLQQ IFSLFRAN ISLDGWEENHTEKFLIQLHQQLEYLEELTGPEAEQDSCILGSENVRLQIKMYFREIHDYLESQEY   159     Homo sapiens   81   LHEMLQQ IFSLFRAN ISLDGWEENHTEKFLIQLHQQLEYLEALMGLEAEKLSGTLGSDNLRLQVKMYFREIHDYLESQEY   159     Homo sapiens   81   LHEMLQQ IFSLFRAN ISLDGWEENHTEKFLIQLHQQLEYLEALMGLEAEKLSGTLGSDNLRLQVKMYFREIHDYLESQEY   159     Homo sapiens   81   LHEMLQQ IFSLFRAN ISLDGWEENHTEKFLIQLHQQLEYLEALMGLEAEKLSGTLGSDNLRLQVKMYFREIHDYLESQEY   159     Komo sapiens   161   SCAWTIVQ IELNRCLFFWIQLTGKLSKQGMDP   160     Scamelus forus   161   SCAWTIVQ IELNRCLFFWIQLTGKLSKQGMDP   193     Scamelus forus   161   SCAWTIVQ IELNRCLFFWIQLTGKLSKQGMDP   193     Scamelus forus   161   SCAWTIVQ IELNRCLFFWIGLTGKLSKQGMDP   193     Scasorofa	Bostaurus	81	LHEMI		FSI	FRAI	VSL	D GWE	ESH	TEK	FLVE	LHQ	<b>LE</b>	YLE	ALMR	LQA	KOK	SDT	LGSE	NLR	LQV	<b>MY</b> F	QRI	HDYL	ESQDY	160
Equus caballus   81   L HEML QQ I FNL FRANIPLD AWEESHME TFLIELHQQLEYLEALMGLEAEQKC GPL GSENL RLQVKMY FR INDYLENQEY 160     Macaca mulatta   81   L HEML QQ I FNL FRANISLD GWEENHMEKFLIQLHQQLEYLEALMGLEAEKLSGTLGSDNLRLQVKMY FR INDYLENQEY 160     Felis catus   80   L HEML QQ I FNL FRANISLD GWEENHMEKFLIQLHQQLEYLEALMGLEAEKLSGTLGSDNLRLQVKMY FR INDYLENQEY 159     Homo sapiens   81   L HEML QQ I FSL FRANISLD GWEENHMEKFLIQLHQQLEYLEALMGLEAEKLSGTLGSDNLRLQVKMY FR INDYLENQEY 159     Homo sapiens   81   L HEML QQ I FSL FRANISLD GWEENHTEKFLIQLHQQLEYLEALMGLEAEKLSGTLGSDNLRLQVKMY FR INDYLENQEY 159     Homo sapiens   81   L HEML QQ I FSL FRANISLD GWEENHTEKFLIQLHQQLEYLEALMGLEAEKLSGTLGSDNLRLQVKMY FR INDYLENQEY 159     Homo sapiens   81   L HEML QQ I FSL FRANISLD GWEENHTEKFLIQLHQQLEYLEALMGLEAEKLSGTLGSDNLRLQVKMY FR INDYLENQEY 159     Gamelus dromedarius   161   SS C AWT I VQ IE INRCLFFWI QLTGKLSKQ GMD P   160     Camelus ferus   161   SS C AWT I VQ IE INRCLFFVI QLTGKLSKQ GMD P   193     Salaenoptera acutorostratu   161   SS C AWT I VQ VE INRCLFFV FQLTGKLSKQ GME P   193     Sus scrofa   161   SS C AWT I VQ VE INRCLFFV FQLTGKLSKQ GME F   193     Sos taurus   161   SS C AWT I VQ VE INRCLFFV FQLTGKLSKQ GME F   193     Sos taurus   161   SS C AWT I VQ VE INRCLFFV	Ovis aries	81	LHEMI		FNI	FRAT	SSL	D GWE	ESH	TEK	FLVE	LLQ	Q L E	YLE	ALMR	LQA	KKK	S D T	LGSE	NLR	LQV	<b>MY</b> F	QRI	HDYI	ESQDY	160
Macaca mulatta   81   L HEML QQ I F SL F RAN I SLD GWE E NHME K FLI QL HQQL EYL E ALMGL E AE KL SGTL G SDNL RL QV KMY F R I HDYL E NQDY 160     Felis catus   80   L HEML QQ I F NL F RAN T SSD GWE E SHVE K FLT E L HQQL EYL E ALMGL E AE KL SGTL G SDNL RL QV KMY F R I HDYL E NQDY 160     Homo sapiens   81   L HEML QQ I F SL F RAN I SLD GWE E NHVE K FL T E L HQQL EYL E ALMGL E AE KL SGTL G SDNL RL QV KMY F R I HDYL E NQDY 160     100   ramelus dromedarius   161   SS C AWT I VQ I E I NRCL F FM I QL T G KL SKQ GMD P   193     Camelus ferus   161   SS C AWT I VQ I E I NRCL F FM I QL T G KL SKQ GMD P   193     Salaenoptera acutorostrata   161   SS C AWT I VQ VE I NRCL F FV F QL T G KL SKQ GMD P   193     Sus scrofa   161   SS C AWT I VQ VE I NRCL F FV F QL T G KL SKQ GME P   193     Sos taurus   161   SS C AWT I VQ VE I NRCL F FV F QL T G KL SKQ GME P   193     Sos taurus   161   SS C AWT I VQ VE I NRCL F LV F R L T R K L SE Q G ME T   193     Sos taurus   161   SS C AWT I VQ VE I NRCL F LV F R L T R K L SE Q G ME T   193     Sos taurus   161   SS C AWT I VQ VE I NRCL F LV F R L T R K L SE Q G ME T   193     Sos taurus   161   SS C AWT I VQ VE I NRCL F LV F R L T R K L SE Q G ME T   193     Bots aries<	Equus caballus	81	L HEMI	QQI	FNI	FRAN	I PL	DAWE	ESH	MET	FLIE	LHQ	a L E	YLE	ALMG	LEA	EQK	CGP	LGSE	NLR	LQV	<b>MY</b> F	RRI	RDYI	ENQEY	160
Felis catus   80 L HEML 00 I FNL FRANTSSDGWEESHVEKFLTEL HOOLEYL EELTGPEAEODSCILGSENVRLOIKMYFOR I HDYLESQEY 159     Homo sapiens   81 LHEML 00 I FNL FRANTSSDGWEESHVEKFLTEL HOOLEYL EELTGPEAEODSCILGSENVRLOIKMYFOR I HDYLESQEY 159     Homo sapiens   81 LHEML 00 I FNL FRANTSSDGWEESHVEKFLTEL HOOLEYL EELTGPEAEODSCILGSENVRLOIKMYFOR I HDYLESQEY 159     Camelus dromedarius   161 SSCAWT I VOIE I NRCL FFMI 0LTGKL SKOGMDP   193     Camelus ferus   161 SSCAWT I VOIE I NRCL FFMI 0LTGKL SKOGMDP   193     Salaenoptera acutorostrati   161 SSCAWT I VOVE I NRCL FFVI OLTGKL SKOGMDP   193     Sus scrofa   161 SSCAWT I VOVE I NRCL FFVFOLTGKL SKOGMEP   193     Sos mutus   161 SSCAWT I VOVE I NRCL FFVFOLTGKL SKOGME F   193     Sos mutus   161 SSCAWT I VOVE I NRCL FFVFOLTGKL SKOGME F   193     Sos mutus   161 SSCAWT I VOVE I NRCL FLVFRL TRKL SEQGME T   193     Sos mutus   161 SSCAWT I VOVE I NRCL FLVFRL TRKL SEQGME T   193     Gots arise   161 SSCAWT I VOVE I NRCL FLVFRL TRKL SEQGME T   193     Gots arise   161 SSCAWT I VOVE I NRCL FLVFRL TRKL SEQGME T   193     Macaca mulatta   161 SSCAWT I VOVE I NRCL FFVFQL TGKL STOGMEP   193     Gots arise   161 SSCAWT I VOVE I NRCL FFVFQL TGKL STOGMEP   193     Macaca amulatta	Macaca mulatta	81	LHEMI		FSI	FRAN	I SL	D GWE	ENH	MEK	FLIQ	LHQ	<b>ALE</b>	YLE	ALMG	ILEA	EKL	SGT	LGSD	NLR	LQV	<b>(MY</b> F	RRI	HDYL	ENQDY	160
Homo sapiens   81   L HEML QQ I F SL F RAN I SL D GWEEN HTEK FL I QL HOQL EYL EALMGL EAEKL SGTL G SDNL RL QVKMY F R I HDYL ENQDY 160     170   180   190   200     Camelus dromedarius   161   SS C AWT I VQ I E I NRCL F FMI QL T GKL SKQ GMD P   193     Camelus ferus   161   SS C AWT I VQ I E I NRCL F FMI QL T GKL SKQ GMD P   193     Scaneous appens   161   SS C AWT I VQ I E I NRCL F FVI QL T GKL SKQ GMD P   193     Scaneous appens   161   SS C AWT I VQ I E I NRCL F FVI QL T GKL SKQ GMD P   193     Scalaenoptera acutorostrata   161   SS C AWT I VQ VE I NRCL F FVF RL T GKL SKQ GME P   193     Scalaenoptera acutorostrata   161   SS C AWT I VQ VE I NRCL F FV F RL T GKL SKQ GME P   193     Scas scrofa   161   SS C AWT I VQ VE I NRCL F FV F RL T GKL SKQ GME P   193     Sos staurus   161   SS C AWT I VQ VE I NRCL F LV F RL T RK L SE Q GME T   193     Sos staurus   161   SS C AWT I VQ VE I NRCL F LV F RL T RK L SE Q GME T   193     Sos taurus   161   SS C AWT I VQ VE I NRCL F FV F Q L T GK L S K Q G ME T   193     Sos taurus   161   SS C AWT I VQ VE I NRCL F FV F Q L T GK L S K Q G ME T   193     Macaca mulatata	Felis catus	80	LHEMI		FNI	FRAN	TSS	D GWE	ESH	VEK	FLTE	LHQ	<b>RE</b>	YLE	ELTG	IPEA	EQD	SCI	LGSE	NVR		<b>MY</b> F	QRI	HDYI	ESQEY	159
170180190200Camelus dromedarius161SS C AWT I VQ I E I NRCL F FMI QL T GK L SKQ GMD P193Camelus ferus161SS C AWT I VQ I E I NRCL F FMI QL T GK L SKQ GMD P193Vicurga pacos161SS C AWT I VQ I E I NRCL F FV I QL T GK L SKQ GMD P193Salaenoptera acutorostrata161SS C AWT I VQ VE I NRCL F FV FQL T GK L SKQ GME P193Sus scrofa161SS C AWT I VQ VE I NRCL F FV FQL T GK L SKQ GME P193Sos staurus161SS C AWT I VQ VE I NRCL F FV FQL T GK L SKQ GME T193Jois aries161SS C AWT I VQ VE I NRCL FLV FRL T RK L SE Q GME T193Equus caballus161SS C AWT I VQ VE I NRCL FLV FRL T RK L SE Q GME T193Gacaa mulatta161SS C AWT I VQ VE I NRCL FLV FRL T RK L SE Q GME T193Felis catus161SS C AWT I VQ VE I NRCL FLV FRL T RK L SE Q GME T193Gacaa mulatta161SS C AWT I VQ VE I NRCL FFV FS L T EK L SKQ GT DP193Game as pienes161SS C AWA I VQ VE I NRCL FFV FS L T EK L SKQ GT DP193Felis catus160SS C AWA I VQ VE I NRCL FFV FS L T EK L SKQ GT DP193Geno sapienes161ST C AWA I VQ VE I NRCL FFV FS L T EK L SKQ G RP L NDMKQE L T T E FRS PR208	Homo sapiens	81	LHEMI		FSI	FRAN	I SL	D GWE	ENH	TEK	FLIQ	LHQ		YLE	ALMG	LEA	EKL	SGT	LGSD	NLR	LQV	MY F	RRI	HDYI	ENQDY	160
170180190200Camelus dromedarius161SS CAWT I VO I E I NRCL FFMI QL TGKL SKOGMD P193Camelus ferus161SS CAWT I VO I E I NRCL FFMI QL TGKL SKOGMD P193Vicurga pacos161SS CAWT I VO I E I NRCL FFVI QL TGKL SKOGMD P193Salaenoptera acutorostrata161SS CAWT I VO VE I NRCL FFVFQL TGKL SKOGMD P193Sus scrofa161SS CAWT I VO VE I NRCL FFVFQL TGKL SKOGME P193Sos staurus161SS CAWT I VO VE I NRCL FFVFQL TGKL SKOGME P193Jois aries161SS CAWT I VO VE I NRCL FLVFRL TRKL SE QGME T193Equus caballus161SS CAWT I VO VE I NRCL FFVFQL TGKL SKOGMD P193Macaca mulatta161SS CAWT I VO VE I NRCL FFVFQL TGKL ST QGMD P193Gausa camulatta161SS CAWT I VO VE I NRCL FFVFQL TGKL ST QGMD P193Equus caballus161SS CAWT I VO VE I NRCL FFVFQL TGKL ST QGMD P193Gausa camulatta161SS CAWT I VO VE I NRCL FFVFSL TEKL SK QG GME T193Gausa caballus161SS CAWT I VO VE I NRCL FFVFSL TGKL ST QGMD P193Gausa caballus161SS CAWA I VO VE I NRCL FFVFSL TEKL SK QG GMD P193Gausa caballus161SS CAWA I VO VE I NRCL FFVFSL TEKL SK QG GMD P100Gausa caballas161SS CAWA I VO VE I NRCL FFVFSL TEKL SK QG P100Gausa caballas161SS CAWA I VO VE I NRCL FFVFSL TEKL SK QG P100Gausa caballas161SS CAWA I VO VE I NRCL FFVFSL TEKL SK QG P100Gausa																										
Camelus dromedarius   161   SS C AWT I VOIE INRCLFFMIOLTGKL SKOGMDP   193     Camelus ferus   161   SS C AWT I VOIE INRCLFFMIOLTGKL SKOGMDP   193     Vicugna pacos   161   SS C AWT I VOIE INRCLFFVIOLTGKL SKOGMDP   193     Salaenoptera acutorostrata   161   SS C AWT I VOVE INRCLFFVFQLTGKL SKOGMDP   193     Sus scrofa   161   SS C AWT I VOVE INRCLFFVFQLTGKL SKOGMEP   193     Sus scrofa   161   SS C AWT I VOVE INRCLFFVFQLTGKL SKOGMEP   193     Sos staurus   161   SS C AWT I VOVE INRCLFLVFRLTRKL SE Q GME T   193     Sos taurus   161   SS C AWT I VOVE INRCLFLVFRLTRKL SE Q GME T   193     Jois aries   161   SS C AWT I VOVE INRCLFLVFRLTRKL SE Q GME T   193     Jois aries   161   SS C AWT I VOVE INRCLFLVFRLTRKL SE Q GME T   193     Jois aries   161   SS C AWT I VOVE INRCLFFVFRLTRKL SE Q GME T   193     Gausa abulus   161   SS C AWT I VOVE I NRCLFFVFRLTRKL SE Q GME T   193     Macaca mulatta   161   SS C AWA I VOVE I NRCLFFVFSLTEKL SK Q G   100     Macaca mulatta   161   SS C AWA I VOVE I NRCLFFVFSLTEKL SK Q G   100     Macaca mulatta					17	0		180			190			200												
Cumelus Submedials   161   SS C AWT I VOIE INRCL FFMIOLTGKL SKOGMD P   193     Vicugna pacos   161   SS C AWT I VOIE INRCL FFVIOLTGKL SKOGMD P   193     Salaenoptera acutorostrata   161   SS C AWT I VOVE INRCL FFVFRL TGKL SKOGMD P   193     Sus scróp   161   SS C AWT I VOVE INRCL FFVFRL TGKL SKOGME P   193     Sus scróp   161   SS C AWT I VOVE INRCL FFVFRL TGKL SKOGME P   193     Sos mutus   161   SS C AWT I VOVE INRCL FLVFRL TKL SE OGME T   193     Sos mutus   161   SS C AWT I VOVE INRCL FLVFRL TKL SE OGME T   193     Sos mutus   161   SS C AWT I VOVE INRCL FLVFRL TKL SE OGME T   193     Sos mutus   161   SS C AWT I VOVE INRCL FLVFRL TKL SE OGME T   193     Gois aries   161   SS C AWT I VOVE INRCL FLVFRL TKL SE OGME T   193     Guus caballus   161   SS C AWT I VOVE I NRCL FFLVFRL TKL SE OGME T   193     Macaca mulatta   161   SS C AWA I VOVE I NRCL FFVFS TEKL SKOG G   193     Macaca mulatta   161   SS C AWA I VOVE I NRCL FFVFS TEKL SKOG G   193     Felis catus   160   SS C AWT I VOVE I NRCL FFVFS TEKL SKOG G   193     Macaca mulatta	Camalus decom adarius	161	SSCAV				LEE	MIOL	TGK	SK	GMD	D				0.0000000	7225									103
Construint   161   SSCAWT   IVQIE INRCLFFVIQLTGKLSKQGMDP   193     Salaenoptera acutorostrata   161   SSCAWT   IVQVE INRCLFFVFRLTGKLSKQGMDP   193     Sus scrofa   161   SSCAWT   IVQVE INRCLFFVFRLTGKLSKQGMEP   193     Sos mutus   161   SSCAWT   IVQVE INRCLFFVFRLTGKLSKQGMEP   193     Sos mutus   161   SSCAWT   IVQVE INRCLFFVFRLTRKLSEQGME T   193     Sos taurus   161   SSCAWT   IVQVE INRCLFFVFRLTRKLSEQGME T   193     Ovis aries   161   SSCAWT IVQVE INRCLFFVFRLTRKLSEQGME T   193     Gacaca mulatta   161   SSCAWT IVQVE INRCLFFVFRLTRKLSEQGME T   193     Macaca mulatta   161   SSCAWT IVQVE INRCLFFVFRLTRKLSEQGME T   193     Felis catus   161   SSCAWT IVQVE INRCLFFVFSLTEKLSKQGMDP   193     Macaca mulatta   161   SSCAWT IVQVE INRCLFFVFSLTEKLSKQG   100     Homo sapiens   161   STCAWA IVQVE INRCLFFVFSLTEKLSKQGRPLNDMKQELTTEFRSPR   208	Camelus forus	161	SSCAV	VT I V			LEF	MIOI	TGK	SK	GMD	P														193
Anagina puess   Anagina puess<	Vicuana pacos	161	SSCAV	VT I V	01	INRC	LEE	VIOI	TGK	SK	GMD	P					-									193
Susservig   161   SS C AWT I V R VE I NRCL FFV FQL T GKL SKQ GME P   193     Sos suutus   161   SS C AWT I V Q VE I NRCL FLV FRL T RKL SE Q GME T   193     Sos suutus   161   SS C AWT I V Q VE I NRCL FLV FRL T RKL SE Q GME T   193     Jois aries   161   SS C AWT I V Q VE I NRCL FLV FRL T RKL SE Q GME T   193     Jois aries   161   SS C AWT I V Q VE I NRCL FLV FRL T RKL SE Q GME T   193     Equus caballus   161   SS C AWT I V Q VE I NRCL FFLV FRL T RKL SE Q GME T   193     Macaca mulatta   161   SS C AWT I V Q VE I NRCL FFLV FRL T RKL SE Q GME P   193     Macaca mulatta   161   SS C AWT I V Q VE I NRCL FFLQL I RK I SK Q G   10P     Felis catus   160   SS C AWT I V Q VE I NRCL FFLQL I RK I SK Q G   10P     Macaca mulatta   161   SS C AWT I V Q VE I NRCL FFLQL I RK I SK Q G   10P   193     Macaca mulatta   160   SS C AWT I V Q VE I NRCL FFLQL I RK I SK Q G   10P   193     Macaca mulatta   161   ST C AWA I V Q VE I SR CL FFV FSL T E KL SK Q G R PL NDMK Q EL T T E FR SP R   208     Momo sapienes   161   ST C AWA I V Q VE I SR CL FFV FSL T E KL SK Q G R PL NDMK Q EL T T E FR SP R   208	Ralaenontera acutorostrata	161	STCAV	VT I V	vov		LFF	VERL	TGK	SK	GME	T					-									193
Main of the second s	Sus scrofa	161	SSCAV	VT I V			LFF	VFQL	TGK	LSK	GME	P					-									193
Sos taurus   161   SS C AWT I VQVE I NRCL FLVFRL TRKL SEQ GME T   193     Ovis aries   161   SS C AWT I VKVE I NRCL FLVFRL TRKL SEQ GME T   193     Equus caballus   161   SS C AWT I VKVE I NRCL FFVFQ T TRKL SEQ GME T   193     Iguar caballus   161   SS C AWT I VKVE I NRCL FFVFQ T TRKL SEQ GME T   193     Macaca mulatta   161   SS C AWT I VQVE I NRCL FFVFQ T TEKL SKQ G   193     Felis catus   160   SS C AWT I VQVE I NRCL FFVFS T TEKL SKQ G   TDP   193     Homo sapiens   161   ST C AWA I VQVE I NRCL FFVFS T TEKL SKQ GR PLNDMKQEL T TE FRSPR   208	Bosmutus	161	SSCAV	VT I V	vov	INRC	LFL	VFRL	TRK	LSE	GME	T					-									193
Ovis aries   161   S C AWT I V K V E I NR CL FL V FRL TR KL S E Q GME T   193     Equus caballus   161   S C ART I V Q V E I NR CL F F V F Q L TG KL S T Q GM P   193     Macaca mulatta   161   S C AWT I V Q V E I NR CL F F V F S L TE KL S K Q G   109     Macaca mulatta   161   S C AWT I V Q V E I NR CL F F V F S L TE KL S K Q G   109     Felis catus   160   S C AWT I V R V E I NR CL F F ALQ L I R K I S KR GMH S S K NV E H E P R AD F R S I G   208     Homo sapiens   161   S T C AWA I V Q V E I S R CL F F V F S L TE KL S K Q G R P L NDM KQ E L T T E F R S P R   208	Bostaurus	161	SSCAV	VT I V	vov	INRC	LFL	VFRL	TRK	LSEC	GME	Τ					-									193
Equus caballus   161   SSCARTIVQVEINRCLFFVFQLTGKLSTQGMDP   193     Macaca mulatta   161   SSCAWAIVQVEINRCLFFVFSLTEKLSKQG   193     Felis catus   160   SSCAWTIVRVEINRCLFFALQLIRKISKRGMHSSKNVEHEPRADFRSIG   208     Homo sapiens   161   STCAWAIVQVEISRCLFFVFSLTEKLSKQGRPLNDMKQELTTEFRSPR   208	Ovis aries	161	SSCAV	VT I V	KV	INRC	LFL	VFRL	TRK	LSE	GME	T 1					-									193
Macaca mulatta   161   SSCAWA I VQVE I NRCL FFVFSL TEKL SKQG	Equus caballus	161	SSCAR	RTIV	QV	INRC	LFF	VFQL	TGK	LST	GMD	P					-									193
Felis catus   160   SSCAWT I VRVE I NRCL F FALQL I RK I SKRGMHSSKNVEHE PRADFRSIG   208     Homo sapiens   161   STCAWA I VOVE I SRCL F FVF SLTEKLSKOGRPL NDMKQEL T T E FRSPR -   208	Macaca mulatta	161	SSCAV	VA I V	QV.	INRC	LFF	VFSL	T E KI	LSK	QG				T	DP-	-									193
Homo sapiens 161 STCAWA LVQVE I SRCL FFVFSLTE KLSKQGRPLNDMKQELTTEFRSPR - 208	Felis catus	160	SSCAV	VI I V		INRC	LFF	ALQL	IRK	I SKI	RGMH	SSK	NVE	HEP	RADF	RSI	G									208
	Homo sapiens	161	STCAV	VA I V		ISRC	LFF	VFSL	TE K	LSK	GRP	LND	MKQ	ELT	TEFR	SPR	2 -									208

Fig 2. Alignment of the deduced amino acid sequence of C. dromedrius IFNE with IFNE from other species.

https://doi.org/10.1371/journal.pone.0213880.g002

#### Expression, solubilization, and refolding of camel IFNE protein

The Arabian camel IFNE cDNA was expressed in E. coli BL21(DE3) pLysS as a 6-histidine fusion protein under the control of the T7 promoter of the pET28a (+) vector. The recombinant protein was found to be overexpressed when E. coli cells were induced with either 1.0 mM IPTG or 2.0 g/L lactose in the fermentation medium (Fig 6A and 6B). Surprisingly, the recombinant protein was found as insoluble inclusion bodies that were precipitated in the form of submicron spherical proteinaceous particles upon cell disruption by sonication and after centrifugation at 12,000 rpm for 10 min at 4°C, leaving behind a supernatant devoid of the recombinant IFNe protein (Fig 6B). The transmission electron micrograph (Fig 7A) showed that the *E. coli* cells becomes to form dark, dense spot areas in the cytoplasm when induced to express the recombinant IFNe protein either by IPTG or lactose. The recombinant camel IFNe inclusion bodies appeared as homogeneous spherical particles of the diameter ranging from 0.5 to 1.0  $\mu$ m under SEM (Fig 7). It is well documented that the expression of a foreign gene in the E. coli cells results in the accumulation of recombinant proteins in the form of inactive, insoluble aggregates of inclusion bodies. Thus, the biggest challenge remaining is the recovery of soluble and functional active recombinant protein from inclusion bodies; this requires standardization protocols for solubilization, re-folding, and subsequent purification [39]. Interestingly, the camel IFN $\varepsilon$  inclusion bodies are localized preferentially in the polar region of the E. coli cells, as well as in the mid-cell region. This polar distribution is mainly attributed to macromolecular crowding in the nucleoid region that is rich in nucleic acids and



Fig 3. Phylogenetic relationship of *C. dromedarius* interferon epsilon and sequences from other species. Maximum likelihood tree based on complete coding sequences deposited in GenBank. Values at nodes are bootstrap  $\geq$ 50%, obtained from 1000 re-samplings of the data.

https://doi.org/10.1371/journal.pone.0213880.g003

other macromolecules, which might prevent the accumulation of large protein aggregates [39]. In most cases, urea at a high concentration (4–8 M) or guanidine hydrochloride was used to solubilize and refold inclusion bodies. Our attempt to solubilize and refold the camel IFN¢ inclusion bodies was failed (data not shown). Thus, the alternative solubilization and refolding protocol was applied based on a strong anionic detergent SDS, which can be easily removed by precipitation with KCl. The recombinant camel IFN¢ inclusion bodies were collected, solubilized and refolded by the SDS/KCl method (Fig 8A, Lane 3). The solubilized and refolded inclusion bodies were subjected to nickel-affinity chromatography. The recombinant camel

#### GAATTC ATGATTAACAAGCCTTTCTT

1	ATG	ATT	AAC	AAG	CCT	TTC	TTT	GAA	ATT	GTG	TTG	GTG	CTG	CTG	GCT	45
1	Met	Ile	Asn	Lys	Pro	Phe	Phe	Glu	Ile	Val	Leu	Val	Leu	Leu	Ala	15
46	TTT	TCC	ACC	ATC	TTC	TCC	CGA	GAG	TTG	AAA	CCG	ATT	CTT	TTC	CAA	90
16	Phe	Ser	Thr	Ile	Phe	Ser	Arg	Glu	Leu	Lys	Pro	Ile	Leu	Phe	(Gln)	30
91	CAA	AGA	AGA	GTA	AAC	AGA	GAG	AGT	TTA	AAA	CTC	CTG	AAT	AAA	TTG	135
31	Gln	Arg	Arg	Val	Asn	Arg	(Glu)	Ser	Leu	Lys	Leu	Leu	(Asn)	Lys	Leu	45
136	CGG	ACC	TCA	TCA	ATT	CAG	CAG	TGT	CTA	CCA	CAT	AGG	AAA	AAC	TTC	180
46	Arg	Thr	Ser	Ser	Ile	Gln	Gln	Cys	Leu	Pro	His	Arg	Lys	Asn	Phe	60
181	TTG	CTT	CCC	CAG	AAG	TCT	ATG	GAT	CCT	CAC	CAG	TAT	CAG	AAA	GGA	225
61	Leu	Leu	Pro	Gln	Lys	Ser	Met	Asp	Pro	His	Gln	Tyr	Gln	Lys	Gly	75
226	CAC	ATA	CTG	GCC	ATT	CTT	CAT	GAG	ATG	CTT	CAG	CAG	ATT	TTC	AAC	270
76	His	Ile	Leu	Ala	Ile	Leu	His	Glu	Met	Leu	Gln	Gln	Ile	Phe	Asn	90
271	CTC	TTC	AGG	GCA	GTT	ATT	TCT	CTG	GAT	GGT	TGG	GAA	GAA	ATC	CAA	315
91	Leu	Phe	Arg	Ala	Val	Ile	Ser	Leu	Asp	Gly	Trp	Glu	Glu	Ile	Gln	105
316	ATG	GAT	AGA	TTC	CTC	TCT	GAA	CTT	CAT	CAA	CAG	CTG	GAA	TAC	CTA	360
106	Met	Asp	Arg	Phe	Leu	Ser	Glu	Leu	His	Gln	Gln	Leu	Glu	Tyr	Leu	120
361	GAA	ACA	CTC	ATA	CGA	CTG	CAA	GCT	GAA	CAG	AGA	AGT	GGC	ATC	TTG	405
121	Glu	Thr	Leu	Ile	Arg	Leu	Gln	Ala	Glu	Gln	Arg	Ser	Gly	Ile	Leu	135
406	GGT	AGT	GAG	AAC	CTT	AGG	TTA	CAG	GTT	AAA	AGT	TAC	TTC	CAA	AGG	450
136	Gly	Ser	Glu	Asn	Leu	Arg	Leu	Gln	Val	Lys	Ser	Tyr	Phe	Gln	Arg	150
451	ATC	CAT	GAT	TAC	CTG	GAA	AGT	CAG	GAA	TAC	AGC	AGC	TGT	GCC	TGG	495
151	Ile	His	Asp	Tyr	Leu	Glu	Ser	Gln	Glu	Tyr	Ser	Ser	Cys	Ala	Trp	165
496	ACC	ATT	GTC	CAG	ATA	GAA	ATC	AAC	CGG	TGT	CTG	TTC	TTT	ATG	ATC	540
166	Thr	Ile	Val	Gln	Ile	Glu	Ile	Asn	Arg	Cys	Leu	Phe	Phe	Met	Ile	180
541	CAA	CTC	ACA	GGA	AAG	CTG	AGC	AAA	CAA	GGA	ATG	GAT	CCT	TGA	582	
181	Gln	Leu	Thr	Gly	Lys	Leu	Ser	Lys	Gln	Gly	Met	Asp	Pro	End		

#### CGTTTGTTCCTTACCTAGGATTCGAA

**Fig 4.** Nucleotide and deduced amino acid encoding region of *C. dromedarius* IFN $\epsilon$ . Important amino acid residues and regions include: residues contact to N-Acetyl-2-Deoxy- are in box; residues contact to SO<sub>4</sub> ion are in bold underline; residue contact to Zn<sup>+2</sup> are bold double underline, conserved amino acid residues in IFN $\epsilon$  protein is in bold dashed underline, residues involved in IFNAR-1 binding are in circle and residues involved in IFNAR-2 binding are in bold dashed box. Arrows indicates the location of the forward and reverse primers with restriction enzyme sites are in bold underline italics.

https://doi.org/10.1371/journal.pone.0213880.g004

IFNE was bound to the affinity matrix and eluted using imidazole at a concentration of 500 mM (Fig 8B, Peak 2). The purified protein showed a specific, unique protein band at 24,970 kDa as shown in Fig 8C and 8D. The concentration of *C. dromedarius* recombinant IFNE proteins were monitored after each purification steps as shown in Table 3. Western blotting analysis for recombinant *C. dromedarius* IFNE protein with 6x-His-Tag monoclonal antibody showed that both crude and affinity purified proteins were interacted and gave a unique protein band corresponding to 24.97 KDa(S1 Fig, Panel A, Lanes 3–8 and Panel B, Lanes 2–7). On the other hand, un-induced culture showed no cross reactivity with 6x-His-Tag monoclonal antibody (S1 Fig, Panel A, Lane 1).

#### C. dromedarius IFNE inhibits survival of breast cancer cells

A growing body of evidence demonstrates the antitumor effect of type I interferons [40] however, the effects of recombinant IFN $\epsilon$  on human cancer cells have not been fully elucidated. In order to study the effects of the Arabian camel IFN $\epsilon$  on human cancer cells, MDA-MB-231 and MCF-7 breast cancer cells were treated with different concentrations of recombinant IFN $\epsilon$  protein. After 48 h of treatment, morphological changes were observed starting from 2.6  $\mu$ M of the recombinant protein. Cells rounded up and were more easily detached. The cells



**Fig 5.** (a) Sequence annotations for *C. dromedarius* IFN $\epsilon$  showing the location of  $\alpha$ -helices and residues contact to ligand and ions. Secondary structure by homology active sites residues from PDB site record ( $\mathbf{v}$ ); residues contacts to ligand (\*) and to ions (\*). (b) Predicted 3D structure of *C. dromedarius* IFN $\epsilon$  protein shows the overall secondary structure in cartoon form; ribbon form (c) and DNA binding form (d). Components of secondary structure are  $\alpha$ -helices (blue), coils (green) and turns (red). Alpha helices are labelled from A to F. (e) Model-template alignment of amino acid residues of *C. dromedarius* IFN $\epsilon$  and *H. sapiens* IFN $\alpha$ 2. Components of the secondary structure are shown in blue ( $\alpha$ -helices) and brown (coils). Identical amino acid residues are in bold black. (f) Predicted 3D structure model of *C. dromedarius* based on this model template alignment.

https://doi.org/10.1371/journal.pone.0213880.g005

exhibited shrinkage and reduction in size compared to the control cells, suggesting inhibition of cell viability (Fig 9A). To investigate the effect of recombinant IFN $\epsilon$  on cell viability, MTT assays were performed. Results demonstrate that IFN $\epsilon$  inhibits the viability of both cell lines in a dose dependent manner. IC<sub>50</sub> was calculated revealing concentrations of 5.65±0.2µM and 3.91±0.6 µM for MDA-MB-231 and MCF-7 cells, respectively (Fig 9B). Evasion of regulated modes of cell death has been well established as a hallmark of cancer [41]. To understand the

Annotation features	Amino acid residues
Contact(s) to ligands - N-Acetyl-2-Deoxy-2-Amino- Galactose	Arg <sup>131</sup> , Ser <sup>132</sup>
Sulfate ion (SO <sub>4</sub> )	Arg <sup>36</sup> , Glu <sup>37</sup> , Lys <sup>40</sup> , Lys <sup>44</sup> , Lys <sup>188</sup>
- Beta-D-Glucose, 6-Deoxy-Alpha- D-Glucose	Lys <sup>44</sup> , Glu <sup>102</sup> , Ile <sup>104</sup> , Gln <sup>105</sup> , Arg <sup>108</sup>
- 1,2-Ethanediol	Asn <sup>35</sup> , Asn <sup>43</sup> , Arg <sup>46</sup> , Leu <sup>85</sup> , Gln <sup>116</sup> , Leu <sup>176</sup>
- 4-(2-Hydroxyethyl)-1-Piperazine	Arg <sup>46</sup> , Ser <sup>49</sup>
Contact(s) to metals -Zinc ion	Gln <sup>143</sup>
Contact(s) to nucleic acids	Gln <sup>30</sup> , Arg <sup>33</sup> , Val <sup>34</sup> , Arg <sup>36</sup> , Glu <sup>37</sup> , Leu <sup>39</sup> , Lys <sup>40</sup> , Asp <sup>107</sup> , Ser <sup>111</sup> , Glu <sup>112</sup> , Gln <sup>115</sup> , Glu <sup>118</sup> , Tyr <sup>119</sup> , Phe <sup>177</sup> , Gln <sup>181</sup> , Gly <sup>184</sup>

Table 2. Conserved amino acid residues of C. dromedarius interferon  $\varepsilon$  involved in different ligands and metal ions binding.

https://doi.org/10.1371/journal.pone.0213880.t002

mechanism underlying IFN $\varepsilon$ -induced inhibition of cancer cell survival, MDA-MB-231 and MCF-7 cells were incubated with IFN $\varepsilon$ , and apoptosis assays were performed. Results reveal that, IFN $\varepsilon$  induces early and late apoptosis in both cell lines (Fig 9C). Taken together interferon epsilon induces morphological changes and inhibits the survival of cancer cells in a dose dependent manner via the induction of apoptosis. Cancer is considered an aberrant tissue/organ comprising a hierarchical composition of heterogeneous cell populations. The tumor microenvironment and related cytokines, such as interferons, play a crucial role during tumor development and regulation of cancer cell survival and tumor progression [40]. Type I INFs, such as IFN $\varepsilon$ , signal through interferon  $\alpha/\beta$  receptor (IFNAR) which is composed of two subunits, INFAR1 and IFNAR2. Studies have reported that mice with an impaired Type 1 interferon signaling (*Ifnar1*<sup>-/-</sup>) are more tumor-prone compared with wild type mice when exposed to the carcinogen methylcholanthrene [42] and mice lacking



Fig 6. (a) SDS-PAGE (12%) for un-induced *E. coli* DE3 (BL21) pLysS pET28-a (+) harboring *C. dromedarius* IFNɛ cDNA (Lanes 2 and 3) and lactose induced culture (Lanes 4–7). (b) SDS-PAGE (12%) for un-induced *E. coli* DE3 (BL21) pLysS pET28-a (+) harboring *C. dromedarius* IFNɛ cDNA (Lane 2), IPTG induced culture supernatant (Lane 3), IPTG induced culture inclusion bodies (Lane 4), lactose induced culture supernatant (Lane 5) and lactose induced inclusion bodies (Lane 6). Lane 1 represents pre-stained protein molecular weight markers. Induction was carried out for 5 h at 1 mM IPTG and 2 g/L lactose in the fermentation medium. Arrow indicates the location of inclusion bodies.

https://doi.org/10.1371/journal.pone.0213880.g006



**Fig 7.** (a) Transmission electron microscope micrograph for normal *E. coli* BL21 (DE3) pLysS harboring pET28a (+) carrying *C. dromedarius* IFNε gene becomes to form inclusion bodies, dark spots when induced to overexpress the recombinant protein. Direct magnification was 10,000 x. (b), (c) and (d) Scanning electron micrograph for the inclusion body showing a spherical particle of a diameter ranging from 0.5 to 1.0 µm. Direct magnification was 35,000 x for b and c and 50,000 x for d.

https://doi.org/10.1371/journal.pone.0213880.g007

functional Type I IFN signaling have shown enhanced susceptibility of for *v*-*Abl*-induced leukemia/lymphoma [43]. IFNAR1-deficient tumors are rejected when transplanted into wild type mice, however, tumors grow when transplanted in *Ifnar1<sup>-/-</sup>* mice, demonstrating the role of type I IFNs in carcinogenesis and tumor progression [42]. IFN- $\alpha/\beta$  has direct effects in tumor cells, inducing growth arrest and apoptosis via activating the JAK-STAT pathway and the expression of genes whose promoters contain the IFN-stimulated response element, such as the apoptosis mediators FAS and TRAIL [44, 40]. The effects of type I IFNs on cancer cells vary depending on the type of tumor, and not all tumor cells are susceptible to the apoptotic effects of IFNs. Similar to orthologs in other species, recombinant canine IFN $\epsilon$  has shown to be capable of activating the JAK-STAT pathway and inhibiting the proliferation of canine cell lines [45]. To complement what has been investigated in the study, the expression level of Caspase-3 was determined to evaluate the cytotoxicity strength and the effectiveness of the potential camel IFN $\epsilon$  protein. Caspase-3 expression has been directly correlated with apoptosis because of its location in the protease cascade pathway as it is



**Fig 8.** (a) SDS-PAGE of *C. dromedarius* IFNε inclusion bodies (Lane 2) and solubilized inclusion bodies (Lane 3). (b) Elution profile of *C. dromedarius* recombinant IFNε after nickel affinity chromatography. Column flow rate was adjusted to be 3 mL/5 min. Arrow indicates the fraction at which buffer was changed to contain imidazole at a concentration of 500 mM as eluent. (c) SDS-PAGE (12%) electrophoresis of nickel affinity purified refolded *C. dromedarius* IFNε, fraction # 21 (Lane 2). (d) SDS-PAGE (12%) for nickel affinity purified recombinant *C. dromedarius* IFNε (Lanes 2–4, 5–15 µg purified protein was loaded into each well). Lane 1 represents pre-stained protein molecular weight markers.

https://doi.org/10.1371/journal.pone.0213880.g008

Purification steps	Protein concentration mg/mL
Crude protein after sonication.	5.54
Refolded dialyzed protein.	0.6
Nickel affinity purified protein.	0.114

eps.
i

https://doi.org/10.1371/journal.pone.0213880.t003

activated by diverse death-inducing signal such as chemotherapeutic agents [46, 47]. Our results showed that caspase-3 expression level was increased in MDA treated cells and the fold of induction was found to be 168.03% and 157.8% at a protein concentration of 3 and 6  $\mu$ M, respectively compared to untreated control cells (Fig 9D). This finding has important clinical implication and in conjunction with other studies suggest that IFNE can be considered as a chemotherapeutic agent that may help in improving the response of adjuvant



C

Fig 9. (a) Recombinant Arabian camel IFN $\epsilon$  alters the morphology of breast cancer cell lines MDA-MB-231 (upper) and MCF-7 (lower). (b) Interferon epsilon inhibits the survival of breast cancer cells. Cells were treated with different concentrations of IFN $\epsilon$  for 48 h. MTT assay was performed and percentage cell viability was calculated compared to control cells. GraphPad Prism 6 was used to calculate the IC<sub>50</sub> of IFN $\epsilon$ : 5.65±0.2 µM and 3.91±0.6 µM for MDA-MB-231 and MCF-7 cells respectively. Experiments were repeated at least 3 times in triplicate. (c) Interferon epsilon induces apoptosis in breast cancer cells. Cells were treated with 5 µM IFN $\epsilon$  protein for 48 h. Apoptosis assay was performed, and the percentage cell viability was calculated (\*p<0.5, \*\*p<0.1 and \*\*\*p<0.01). (d) Expression of caspase-3 in MDA-MB-231 cell line untreated and recombinant IFN $\epsilon$  treated cells at a concentration of 3 and 6 µM.

https://doi.org/10.1371/journal.pone.0213880.g009

therapy for breast cancer. One potential concern relating to the cytotoxic and apoptotic effect of recombinant *C. dromedarius* IFN $\epsilon$  is that some or all of these effects could be accounted for endotoxin contamination from the host *E. coli* cells. Therefore, we measured the Gram-negative endotoxins in the affinity purified recombinant protein and it was clearly observed that, there was no endotoxin was detected in the purified protein (S1 Table). This result indicated that contamination from *E. coli* host cells was extremely unlikely to account for the data produced using nickel affinity purified recombinant *C. dromedarius* IFN $\epsilon$  protein.

In conclusion, we presented here cloning, expression, refolding, and characterization of a novel gene encoding the Arabian camel IFNE. Moreover, this study does underpin the Arabian camel recombinant IFNE as a possible anti-cancer.

#### **Supporting information**

**S1 Fig. Supplementary 1 (S1).** Western blotting analysis of *C. dromedarius* recombinant IFNε protein with 6x-His-Tag monoclonal antibody (1:1000 dilution). Panel (A): Lane 2 represents un-induced *E. coli* harboring pET28a(+) carrying the full-length cDNA; Lane 3, 50 µg of crude sonicated extract from lactose induced culture; Lanes 4–8 represent nickel-affinity purified recombinant protein at 5 to 25 µg concentration. Panel (B), Lanes 2–7 represent nickel affinity purified recombinant protein at 2 to 12 µg concentration. Lanes 1 Panel A and B represent pre-stained protein molecular weight markers. (TIF)

S1 Table. Detection of gram-negative bacterial endotoxins using LAL chromogenic endotoxin quantitation kit.

(DOCX)

#### **Author Contributions**

Data curation: Hesham Saeed, Lamiaa El-Shennawy, Farid Ataya.

Formal analysis: Manal Shalaby.

Investigation: Hesham Saeed, Farid Ataya.

Methodology: Manal Abdel-Fattah, Hesham Saeed, Farid Ataya.

Project administration: Hesham Saeed, Amira Embaby.

Supervision: Hesham Saeed, Hoda Mahmoud.

Writing - original draft: Ahmed Hussein.

#### References

- Baldo BA. Side effects of cytokines approved for therapy. Drug Saf. 2014; 37: 921–943. https://doi.org/ 10.1007/s40264-014-0226-z PMID: 25270293
- Borish LC, Steinke JW. Cytokines and chemokines. J Allergy Clin Immunol. 2003; 111: S460–S475. https://doi.org/10.1067/mai.2003.108 PMID: 12592293
- Vacchelli E, Eggermont A, Fridman WH, Galon J, Zitvogel L, Kroemer G, et.al. Trial Watch Immunostimulatory cytokines. Oncoimmunology. 2013; 2(7): E24850. https://doi.org/10.4161/onci.24850 PMID: 24073369
- 4. Peng FW, Duan ZJ, Zheng LS, Xie ZP, Gao HC, Zhang H, et al. Purification of recombinant human interferon-epsilon and oligonucleotide microarray analysis of interferon-epsilon-regulated genes. Protein Expr Purif. 2007; 53: 356–362. https://doi.org/10.1016/j.pep.2006.12.013 PMID: 17287131

- Tayal V, Kalra BS. Cytokines and anti-cytokines as therapeutics—an update. Trial Watch: Immunostimulatory cytokines. Eur J Pharmacol. 2008; 579: 1–12. https://doi.org/10.1016/j.ejphar.2007.10.049 PMID: 18021769
- Vacchelli E, Galluzzi L, Eggermont A, Galon J, Tartour E, Zitvogel L, et.al. Trial Watch: Immunostimulatory cytokines. Oncoimmunology. 2012; 1:493–506. https://doi.org/10.4161/onci.20459 PMID: 22754768
- Fisher CD, Wachoski-Dark GL, Grant DM, Bramer SA, Klein C. Interferon epsilon is constitutively expressed in equine endometrium and up-regulated during the luteal phase. Anim Reprod Sci. 2018; 195:38–43. https://doi.org/10.1016/j.anireprosci.2018.05.003 PMID: 29807828
- Guo Y, Gao M, Bao J, Luo X, Liu Y, An D, et al. Molecular cloning and characterization of a novel bovine IFN-ε. Gene. 2015; 558 (1): 25–30. https://doi.org/10.1016/j.gene.2014.12.031 PMID: 25523095
- Krause CD, Pestka S. Evolution of the Class 2 cytokines and receptors, and discovery of new friends and relatives. Pharmacol Ther. 2005; 106: 299–346. <u>https://doi.org/10.1016/j.pharmthera.2004.12.002</u> PMID: 15922016
- Hardy HP, Owczarek CM, Jermiin LS, Ejdebäck M, Hertzog PJ. Characterization of the type I interferon locus and identification of novel gene. Genomics.2004; 84(2):331–345. <u>https://doi.org/10.1016/j.ygeno.</u> 2004.03.003 PMID: 15233997
- Kontsek P, Karayianni-Vasconcelos G, Kontsekove E. The human interferon system: characterization and classification after discovery of novel member S47. Acta Virol. 2003; 201–215. PMID: 15068375
- Green DS, Nunes AT, Annunziata CM, Zoon KC. Monocyte and interferon based therapy for the treatment of ovarian cancer. Cytok Growth Factor Rev. 2016; 29: 109–115.
- 13. Cooksley WG. The role of interferon therapy in hepatitis B. Med Gen Med. 2004; 6 (1):16.
- Noël N, Béatrice BJ, Huot N, Goujard C, Lambotte O, Müller-Trutwin M. Interferon-associated therapies toward HIV control; The black and forth. Cytok Growth Factor Rev. 2018; 40: 99–112.
- Shepherd J, Waugh N, Hewitson P. Combination therapy (interferon alfa and ribavirin) in the treatment of chronic hepatitis C: a rapid and systematic review. Health Technol Assess. 2000; 4(33):1–67. PMID: 11134916
- Day SL, Ramshaw IA, Ramsay AJ, Ranasinghe C. Differential effects of the type I interferons alpha4, beta, and epsilon on antiviral activity and vaccine efficacy. J Immunol. 2008; 180: 7158–7166. <u>https://doi.org/10.4049/jimmunol.180.11.7158</u> PMID: 18490714
- Demers A, Kang G, Ma F, Lu W, Yuan Z, Li Y, et al. The mucosal expression pattern of interferon-ε in rhesus macaques. J Leukoc Biol.2014; 96: 1101–1107. <u>https://doi.org/10.1189/jlb.3A0214-088RRR</u> PMID: 25139290
- Fung KY, Mangan NE, Cumming H, Horvat JC, Mayall JR, Stifter SA, et al. Interferon-epsilon protects the female reproductive tract from viral and bacterial infection. Science.2013; 339: 1088–1092. https:// doi.org/10.1126/science.1233321 PMID: 23449591
- Hermant P, Francius C, Clotman F, Michiels T. IFN-ε Is Constitutively Expressed by Cells of the Reproductive Tract and Is Inefficiently Secreted by Fibroblasts and Cell Lines. PloSOne. 2013; 8(8): e71320. https://doi.org/10.1371/journal.pone.0071320 PMID: 23951133
- Altmann SM, Mellon MT, Distel DL, Kin CH. Molecular and functional analysis of an interferon gene from the zebrafish, Danio rerio. J Virol. 2003; 77:1992–2002. https://doi.org/10.1128/JVI.77.3.1992-2002.2003 PMID: 12525633
- Li H, Ma B, Jin H, Wang J. Cloning, *in vitro* expression and bioactivity of goose interferon-α. Cytok. 2006; 34 (3–4): 177–183.
- Robertsen B, Bergan V, Rokenes T, Larsen R, Albuquerque A. Atlantic Salmon Interferon Genes: Cloning, Sequence Analysis, Expression, and Biological Activity. J Interf Cytok Res. 2003; 23: 601–612.
- Suresh M, Karaca K, Foster D, Sharma JM. Molecular and Functional Characterization of Turkey Interferon. J Virol. 1995; 69: 8159–8163. PMID: 7494342
- Al-Swailem AM, Shehara MM, Adu-Duhier FM, Al-Yamani EJ, Al-Busadah KA, Al-Arawi MS, et al. Sequencing, Analysis, and Annotation of Expressed Sequence Tags for *Camelus dromedaries*. PLo-SOne. 2010; 5(5):e10720. https://doi.org/10.1371/journal.pone.0010720
- Ataya FS, Al-Jafari AA, Daoud MS, Al-Hazzani AA, Shehata A, Saeed HM, et al. Genomics, phylogeny and *in silico* analysis of mitochondrial glutathione Stransferase-kappa from the camel *Camelus dromedaries*. Res Vet Sci. 2014; 97(1): 46–54. https://doi.org/10.1016/j.rvsc.2014.04.004 PMID: 24810173
- Sambrook J, Frisch E, Maniatis T. Molecular Cloning: A Laboratory Manual, 2nd ed. Cold Spring Harbor Laboratory Press, New York; 1989.
- Sanger F, Nicklen S, Coulson AR. DNA sequencing with chain-terminating inhibitors. Proc Natl Acad Sci. USA. 1977; 74 (12): 5463–5467. https://doi.org/10.1073/pnas.74.12.5463 PMID: 271968

- Dereeper A, Guignon V, Blanc G, Audic S, Buffet S, Chevenet F, et al. Phylogeny.fr: robust phylogenetic analysis for the non-specialist. J Nucleic Acids Res. 2008; 36: W465–W469.
- Ortiz AR, Strauss CE, Olmea O. MAMMOTH (Matching molecular models obtained from theory): An automated method for model comparison. Protein Sci. 2002; 11: 2606–2621. <u>https://doi.org/10.1110/ ps.0215902</u> PMID: 12381844
- Roy A, Kucukural A, Zhang Y. I-TASSER: a unified platform for automated protein structure and function prediction. Nat Protoc. 2010; 5(4):725–738. https://doi.org/10.1038/nprot.2010.5 PMID: 20360767
- Bradford MM. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal Biochem. 1976; 72 (1):248–254.
- Laemmli UK. Cleavage of Structural Proteins during the Assembly of the Head of Bacteriophage T4. Nature. 1970; 227: 680–685. https://doi.org/10.1038/227680a0 PMID: 5432063
- Towbin H, Staehelin T, Gordon J. Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. Proc Natl Acad Sci USA. 1979; 76:4350–4354. https://doi.org/10.1073/pnas.76.9.4350 PMID: 388439
- He C, Ohnishi K. Efficient renaturation of inclusion body proteins denatured by SDS. Biochem Biophys Res Commun. 2017; 490(4):1250–1253. https://doi.org/10.1016/j.bbrc.2017.07.003 PMID: 28684315
- **35.** Bornhorst JA, Falke JJ. [16] Purification of Proteins Using Polyhistidine Affinity Tags. Methods Enzymol. 2000; 326:245–254. https://doi.org/10.1016/s0076-6879(00)26058-8 PMID: 11036646
- 36. van de Loosdrech AA, Nennie E, Ossenkoppele GP, Beelen RHJ, Langenhuijsen MAC. Cell mediated cytotoxicity against U 937 cells by human monocytes and macrophages in a modified colorimetric MTT assay: A methodological study. J Immunol Met. 1991; 141(1): 15–22.
- 37. Gad HH, Hamming OH, Hartmann R. The structure of human interferon lambda and what it has taught us. J Interf Cytok Res. 2010; 30: 565–571.
- Karpussas M, Nolte M, Benton CB, Meier W, Lipscomb WN, Goelz S. The crystal structure of human interferon beta at 2.2-A resolution. Proc Natl Acad Sci USA. 1997; 94: 11813–11818. <u>https://doi.org/10. 1073/pnas.94.22.11813</u> PMID: 9342320
- Samudzi CT, Burton LE, Rubin JR. Crystal structure of recombinant rabbit interferon-gamma at 2±7-AI resolution. J Biol Chem. 1991; 266: 21791–21797. PMID: 1939201
- Vallejo LF, Rinas U. Strategies for the recovery of active proteins through refolding of bacterial inclusion body proteins. Microb Cell Fact. 2004; 3: 11. https://doi.org/10.1186/1475-2859-3-11 PMID: 15345063
- Ruan MFV, Aline H, Medrano MRFV, Hunger A, Mendonca SA, Barbuto JAM, et. al. Immunomodulatory and antitumor effects of type I interferons and their application in cancer therapy. Oncotarget. 2017; 8:71249–71284. https://doi.org/10.18632/oncotarget.19531 PMID: 29050360
- Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. Cell. 2011; 144: 646–674. <a href="https://doi.org/10.1016/j.cell.2011.02.013">https://doi.org/10.1016/j.cell.2011.02.013</a> PMID: 21376230
- **43.** Dunn GP, Bruce AT, Sheehan KCF, Shankaran V, Uppaluri R, Bui JD, et al. A critical function for type I interferons in cancer immunoediting. Nat Immunol. 2005; 6:722–729. <u>https://doi.org/10.1038/ni1213</u> PMID: 15951814
- 44. Herranz S, Través PG, Luque A, Hortelano S. Role of the tumor suppressor ARF in macrophage polarization.Enhancement of the M2 phenotype in ARF-deficient mice. Oncoimmunology.2012; 1(8): 1227– 1238. https://doi.org/10.4161/onci.21207 PMID: 23243586
- 45. Shen CJ, Chan TF, Chen CC, Hsu YC, Long CY, Lai CS. Human umbilical cord matrix-derived stem cells expressing interferon-β gene inhibit breast cancer cells via apoptosis. Oncotarget.2016; 7: 34172– 34179. https://doi.org/10.18632/oncotarget.8997 PMID: 27129156
- Klotz D, Baumgärtner W, Gerhauser I. Type I interferons in the pathogenesis and treatment of canine diseases. Vet Immunol Immunopathol. 2017; 191: 80–93. https://doi.org/10.1016/j.vetimm.2017.08. 006 PMID: 28895871
- Bonfoco E, Krainc D, Ankarcrona M, Nicotera P, Lipton SA. Apoptosis and necrosis: two distinct events induced, respectively, by mild and intense insults with N-methyl-D-aspartate or nitric oxide/superoxide in cortical cell cultures. Proc Natl Acad Sci. USA. 1995; 92(16):7162–7166. https://doi.org/10.1073/ pnas.92.16.7162 PMID: 7638161