Gene 578 (2016) 7–16

Contents lists available at ScienceDirect

Gene

journal homepage: www.elsevier.com/locate/gene

Research paper

A *colostrum trypsin inhibitor* gene expressed in the Cape fur seal mammary gland during lactation

Elizabeth A. Pharo [a,b,*,1], Kylie N. Cane [a,b], Julia McCoey [c], Ashley M. Buckle [c], W.H. Oosthuizen [d], Christophe Guinet [e], John P.Y. Arnould [a,b,f]

[a] School of BioSciences, The University of Melbourne, Melbourne, VIC 3010, Australia
[b] Cooperative Research Centre for Innovative Dairy Products, Australia
[c] Biomedicine Discovery Institute, Department of Biochemistry and Molecular Biology, Monash University, Clayton, VIC 3800, Australia
[d] Oceans and Coasts, Department of Environmental Affairs, Private Bag X2, Roggebaai 8012, South Africa
[e] Centre d'Etudes Biologiques de Chizé, CNRS, 79360 Villiers en Bois, France
[f] School of Life and Environmental Sciences, Deakin University, Burwood, VIC 3125, Australia

ARTICLE INFO

Article history:
Received 19 July 2015
Received in revised form 13 October 2015
Accepted 25 November 2015
Available online 27 November 2015

Keywords:
Pinniped
Kunitz
BPTI
Protease
Walrus
Milk

ABSTRACT

The *colostrum trypsin inhibitor* (*CTI*) gene and transcript were cloned from the Cape fur seal mammary gland and *CTI* identified by *in silico* analysis of the Pacific walrus and polar bear genomes (Order Carnivora), and in marine and terrestrial mammals of the Orders Cetartiodactyla (yak, whales, camel) and Perissodactyla (white rhinoceros). Unexpectedly, Weddell seal *CTI* was predicted to be a pseudogene. Cape fur seal *CTI* was expressed in the mammary gland of a pregnant multiparous seal, but not in a seal in its first pregnancy. While bovine *CTI* is expressed for 24–48 h postpartum (pp) and secreted in colostrum only, Cape fur seal *CTI* was detected for at least 2–3 months pp while the mother was suckling its young on-shore. Furthermore, *CTI* was expressed in the mammary gland of only one of the lactating seals that was foraging at-sea. The expression of *β-casein* (*CSN2*) and *β-lactoglobulin II* (*LGB2*), but not *CTI* in the second lactating seal foraging at-sea suggested that *CTI* may be intermittently expressed during lactation. Cape fur seal and walrus *CTI* encode putative small, secreted, N-glycosylated proteins with a single Kunitz/bovine pancreatic trypsin inhibitor (BPTI) domain indicative of serine protease inhibition. Mature Cape fur seal CTI shares 92% sequence identity with Pacific walrus CTI, but only 35% identity with BPTI. Structural homology modelling of Cape fur seal CTI and Pacific walrus trypsin based on the model of the second Kunitz domain of human tissue factor pathway inhibitor (TFPI) and porcine trypsin (Protein Data Bank: 1TFX) confirmed that CTI inhibits trypsin in a canonical fashion. Therefore, pinniped CTI may be critical for preventing the proteolytic degradation of immunoglobulins that are passively transferred from mother to young via colostrum and milk.

1. Introduction

Colostrum trypsin inhibitor (*CTI*) is a mammary-specific gene which is expressed and the protein secreted in bovine colostrum for only 24–48 h postpartum (pp) (Laskowski and Laskowski, 1951; Pineiro et al., 1978; Veselsky et al., 1978). CTI is a small ~10–15 kDa, N-glycosylated protein (Klauser et al., 1978; Laskowski and Laskowski 1951; Tschesche et al., 1975) with a single Kunitz/bovine pancreatic trypsin inhibitor (BPTI) domain, characteristic to the family I2 (Kunitz-BPTI) inhibitors of the S1 (chymotrypsin) family of serine endopeptidases (Rawlings et al., 2014). Although CTI inhibits trypsin and plasmin and is a weak inhibitor of α-chymotrypsin *in vitro* (Feeney and Allison, 1969; Laskowski and Laskowski, 1951; Pineiro et al., 1978), neither its target enzyme, nor its function *in vivo* is known.

The expression of *CTI* and the orthologous marsupial *early lactation protein* (*ELP*) gene (Pharo et al., 2012), coincides with the passive transfer of antibodies from the mother to a neonate/young that lacks an

Abbreviations: AMBP, α1-microglobulin/bikunin precursor; APP, amyloid (βA4) precursor protein; CSN2, β-casein; BPTI, bovine pancreatic trypsin inhibitor; cDNA, DNA complementary to RNA; CTI, colostrum trypsin inhibitor; Da, Dalton; ELP, early lactation protein; h, hour; IgG, immunoglobulin G; KD, Kunitz domain; LGB2, β-lactoglobulin II; PIGT, phosphatidyl inositol glycan, class T; pp, postpartum; PRSS1, protease, serine, 1 (trypsin 1); SLPI, secretory leukocyte protease inhibitor; SPINLW1, serine peptidase inhibitor-like, with Kunitz and WAP domains 1; TFPI, tissue factor pathway inhibitor; TKDP, trophoblast Kunitz domain protein; UTR, untranslated region; WAP, whey acidic protein; WFIKKN, WAP, follistatin/kazal, immunoglobulin, Kunitz and netrin domain containing; WFDC2, WAP four disulphide core domain 2.

* Corresponding author at: School of BioSciences, The University of Melbourne, Melbourne, VIC 3010, Australia.
E-mail addresses: e.pharo@unimelb.edu.au (E.A. Pharo), kylie.cane@gmail.com (K.N. Cane), julia.mccoey@gmail.com (J. McCoey), ashley.buckle@monash.edu (A.M. Buckle), oosthuiz@environment.gov.za (W.H. Oosthuizen), guinet@cebc.cnrs.fr (C. Guinet), john.arnould@deakin.edu.au (J.P.Y. Arnould).
[1] Present address.

http://dx.doi.org/10.1016/j.gene.2015.11.042
0378-1119/

adaptive (acquired) immune system and the ability to mount a specific immune response (Brambell 1970; Edwards et al., 2012). Furthermore, during this period, the gut of the young is permeable to intact immunoglobulins and macromolecules and thus these molecules can pass through the intestines and into the circulatory system prior to 'gut closure', i.e., when mucosal enterocytes lose the capacity to absorb macronutrients and immunoglobulins (Kruse, 1983; McFadden et al., 1997). *CTI* expression is brief in eutherians (1–2 days), but ELP expression is extended (for up to 100–125 days pp) in marsupials such as the possum and tammar wallaby (the common and scientific names for mammals described in this study are listed in Supplementary file 1) (Nicholas et al., 1997; Pharo et al., 2012; Piotte and Grigor 1996). Therefore, CTI may prevent the proteolytic degradation of immunoglobulins (Laskowski and Laskowski, 1951), while ELP may protect the marsupial young against pathogens (Piotte and Grigor, 1996).

Neither *CTI*, nor *ELP* is found in birds, fish, reptiles or amphibians and their status in monotremes is inconclusive. They therefore evolved from a common ancestral gene prior to the divergence of marsupials and eutherians (Pharo et al., 2012) ~160 million years ago (Luo et al., 2011). Intriguingly, all marsupials investigated have a functional (putative protein-coding) *ELP* gene, but this is not so for eutherian *CTI*. *CTI* is conserved in species from the orders Carnivora (dog, cat) and Cetartiodactyla (dolphin, cow, pig); but is a pseudogene in the horse (order Perissodactyla), humans and other primates, the elephant, sloth and rodents (Pharo et al., 2012). Gene loss, or loss of function has occurred many times throughout evolution and is often the result of gene duplication (Lynch and Conery, 2000) and/or transposition of genomic DNA fragments within the genome by retro-elements (Cañestro et al., 2013).

Since its evolution over 500 million years ago, the Kunitz domain (KD) has been duplicated many times (Gojobori and Ikeo, 1994; Ikeo et al., 1992) in bacteria, viruses, insects, invertebrates, vertebrates (e.g. fish, birds, amphibians, reptiles (e.g. venoms and dendrotoxins), mammals) and plants (Fry et al., 2009; Jamal et al., 2013; Rawlings et al., 2014). KDs have a diverse range of functions, e.g., serine protease inhibition, antimicrobial, anticoagulant and anti-inflammatory activity; potassium and calcium channel blockers (e.g. neurotoxic venoms), non-neurotoxic venoms, plant protection against herbivores, etc. (Fry et al., 2009; Ranasinghe and McManus, 2013; Shigetomi et al., 2010). Genes encoding as few as one KD, e.g. *BPTI* (also known as *PTI*) (Ascenzi et al., 2003); *serine protease inhibitor Kunitz-type* and *-4* (*SPINT3* and *-4*); *trophoblast Kunitz domain proteins 1*, *2*, *3*, *4* and *5* (*TKDP1-5*); two domains: *SPINT1* and *-2*; three tandemly repeated domains: *tissue factor pathway inhibitor 1* and *-2* (*TFPI1 and -2*), and up to 12 domains, e.g. nematode *Ac*-KPI-1 (*Ancylostoma caninum*)*-Kunitz protease inhibitor-1* have been characterised (Hawdon et al., 2003). Multi-domain type-encoding genes have also been identified, e.g. *serine peptidase inhibitor-like, with Kunitz and WAP domains 1 (eppin) (SPINLW1)*; *WAP, follistatin/kazal, immunoglobulin, Kunitz and netrin domain containing-1* and *-2* (*WFIKKN1* and *-2*); *α1-microglobulin/bikunin precursor (AMBP)* and *amyloid (βA4) precursor protein* (*APP*).

The most extensively structurally characterised KD is BPTI (Ascenzi et al., 2003), which most-likely evolved from bovine *CTI* after the divergence of ruminants from other Cetartiodactyla (Pharo et al., 2012) ~25–35 Myr (Bininda-Emonds et al., 2007). The KD belongs to the α + β fold and is stabilised by three disulphide bonds (Ascenzi et al., 2003; Huber et al., 1972). The P_1 'warhead' residue within the convex exposed 'binding' loop of the inhibitor determines serine protease specificity (Laskowski and Kato, 1980; Laskowski and Qasim, 2000). Kunitz inhibitors with a basic Lys or Arg residue at P_1 inhibit trypsin and trypsin-like enzymes (Ikeo et al., 1992; Laskowski and Kato, 1980), those with Ala or Ser inhibit elastase-like enzymes and a P_1 Met or Leu confers activity against chymotrypsin and chymotrypsin-like enzymes (Laskowski and Kato, 1980). A P_1 Leu is also active against neutrophil elastase (Garcia-Fernandez et al., 2015). Protease inhibition involves the cleavage of the P_1–P_1' peptide bond (inhibitor) and the docking of the convex inhibitor 'binding' loop into the catalytic cleft of the serine protease in a classical 'lock and key' interaction (Marquart et al., 1983; Grzesiak et al., 2000). A new, 1:1, reversible, tight-binding serine protease-Kunitz inhibitor complex is formed, involving P_1 (inhibitor) and the catalytic triad of residues: Ser, His and Asp (protease) (Marquart et al., 1983).

The identification of *CTI* in the dolphin (Cetartiodactyla) (Pharo et al., 2012), suggested that the gene may also be present in other marine mammals, such as the Pinnipedia, 'fin-footed' semi-aquatic mammals (Carnivora). Although monophyletic, the pinnipeds comprise a diverse range of species within three extant families: the Otariidae (eared seals: fur seals, e.g. Cape and subantarctic fur seals; and sea lions, e.g. California seal lion), its sister family, the Odobenidae, with one extant member, the walrus; and the Phocidae (earless or 'true' seals: e.g. Weddell seal, Hooded seal), [Fig. 1; (Higdon et al., 2007; Nyakatura and Bininda-Emonds 2012)]. While all pinnipeds give birth to a pup on land or ice (Atkinson 1997; Oftedal et al., 1987) and produce lipid-rich, high-protein and low-sugar milk for their young, each family has developed unique lactation strategies to ensure the survival of their young in marine environs (Fig. 1). Otariids such as the Cape and subantarctic fur seals fast for the first 4–10 days pp while nursing their pup on land (Bonner, 1984; Oftedal et al., 1987). The mother then commences a pattern of alternating trips to sea to feed interspersed with ~2–3 days on-shore suckling her pup (Bonner, 1984; Gentry and Kooyman, 1986; Oftedal et al., 1987). Each trip to sea usually lasts for up to 25 days (Gamel et al., 2005) and ~23–28 days (Kirkman et al., 2002) for the Cape and subantarctic fur seals respectively, but extreme foraging periods of up to 1–2 months have been reported (Georges and Guinet, 2000; Verrier et al., 2011). Therefore, it is vital that the mother produces milk that will ensure the survival of her pup during these periods.

The aims of this study were to identify the *CTI* gene and characterise its expression in the mammary glands of semi-aquatic members of the order Carnivora (Cape and subantarctic fur seals), identify *CTI* in other species using an *in silico* bioinformatics approach, and to use protein structural homology modelling to investigate whether pinniped CTI, like other KDs has the potential to inhibit trypsin.

2. Materials and methods

2.1. Animals

Mammary gland tissue was collected from six Cape fur seals and one subantarctic fur seal during the reproductive cycle (Table 1), as described previously (Cane, 2005). Approval for this research was obtained from the South African Government (Cape fur seals) and the Ethics Committee of the French Polar Institute (IPEV) and the Polar Environment Committee of Terres Australes et Antarctiques Françaises (subantarctic fur seal).

2.2. Isolation of Cape fur seal genomic DNA and cloning of the CTI gene

Genomic DNA was isolated from the mammary gland of an on-shore lactating Cape fur seal as described (Sambrook and Russell, 2001) and *CTI* amplified by PCR from ~200 ng of genomic DNA with forward 5′-GCCTAGAACATTCAGCTATTGGCACC-3′ and reverse 5′-TGAATGTTTTAT TGACCTAGACCTGGAGG-3′ primers and Platinum *Taq* (Invitrogen) as per the manufacturer's instructions. PCR conditions used were 94 °C for 2 min; 35 cycles of [94 °C for 30 s; 54 °C for 30 s; 68 °C for 4 min] and a final extension at 68 °C for 10 min. The PCR product was cloned into pGEM-T Easy (Promega Corporation) and sequenced.

2.3. In silico identification of CTI

CTI genes were compiled from BLAST searches of the NCBI GenBank nr (http://www.ncbi.nlm.nih.gov/BLAST/) and Ensembl

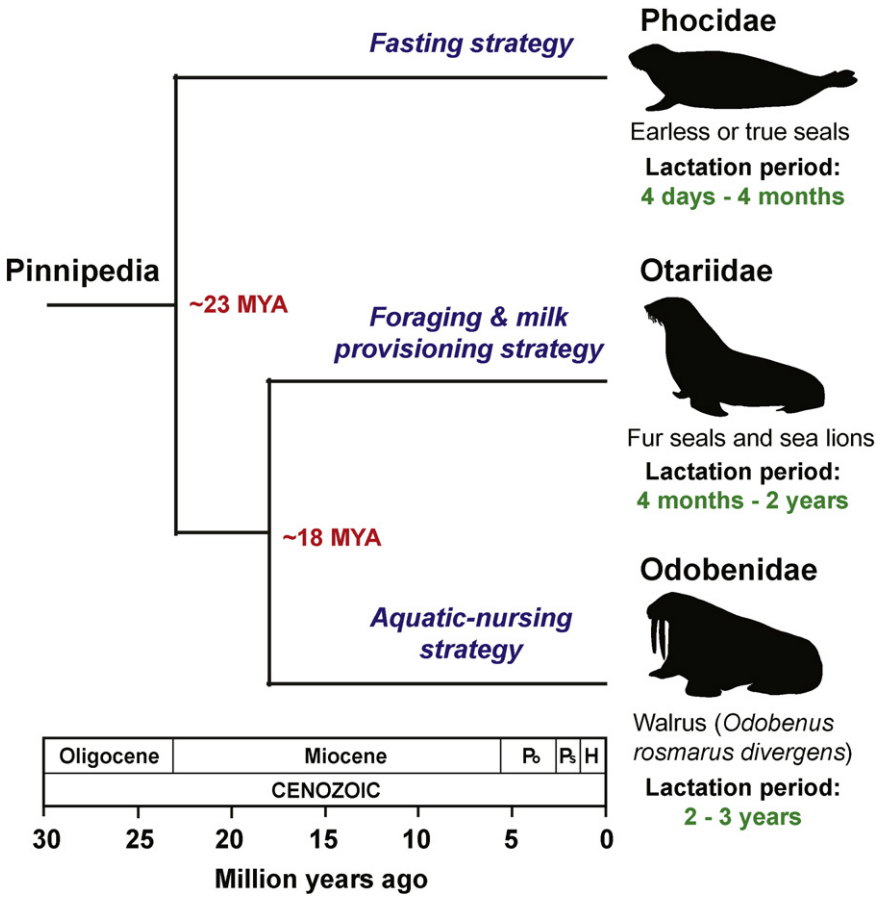

Fig. 1. Evolutionary history of the Pinnipedia. The Pinnipedia consists of three extant families with different reproductive strategies. The Phocidae diverged from the ancestor of the Otariidae and Odobenidae families ~33–23.3 Myr ago and these latter two families then diverged ~26–11 Myr ago (Higdon et al., 2007; Nyakatura and Bininda-Emonds, 2012). The divergence times shown are based on the study by Higdon and colleagues (Higdon et al., 2007). Po, Pliocene; Ps, Pleistocene; H, Holocene.

Release 80, May 2015 (http://www.ensembl.org/) databases or BLAT searches of the UCSC Genome Browser (http://genome.ucsc.edu/) using *CTI* mRNA and protein sequences. Expect-values ≤ 1e − 8 and E-value ≤ 1e − 17 were used as cut-offs for nucleotide and protein comparisons respectively. Putative exons, transcripts and proteins within genomic sequences were predicted using the web-based tools, GENSCAN (http://genes.mit.edu/GENSCAN.html) and Spidey (mRNA-to-genomic alignment) (http://www.ncbi.nlm.nih.gov/spidey/).

2.4. Preparation of total RNA

Total RNA for Northern analysis was isolated from pregnant and lactating Cape fur seal mammary gland tissue using the RNeasy Lipid Tissue Mini kit (Qiagen, Hilden, Germany) as per the manufacturer's instructions.

2.5. RT-PCR cloning of the fur seal CTI transcript

First strand cDNA was generated with Superscript II and total RNA (5 μg) isolated from mammary tissue of an on-shore, lactating Cape fur seal (O3). The Cape fur seal *CTI* transcript was amplified by PCR using the same primers as for the *CTI* gene (Section 2.2), high-fidelity proof-reading Platinum *Taq* (Invitrogen, Carlsbad, CA, USA) and 5% of the first strand reaction as a template. PCR conditions used were 94 °C for 2 min; 35 cycles of [94 °C for 30 s; 55 °C for 30 s; 68 °C for 1 min] and a final extension at 68 °C for 10 min and the PCR product cloned

Table 1
Reproductive stages of Cape and subantarctic fur seals used for this study.

Seal species	Reproductive stage	Label	Location
Cape fur seal	Late-pregnancy, primiparous (first pregnancy) female	P1	Small breeding colony at Robesteen, South Africa (33°37′60S, 18°23′60E)
Cape fur seal	Late-pregnancy, multiparous female	P2	Small breeding colony at Robesteen, South Africa (33°37′′60S, 18°23′′60E)
Cape fur seal	Early-lactation female (on-shore) ~2–3 months pp	O1	Large breeding colony in Kleinsee, South Africa (29°40′0S, 17°4′60E), near the Namibian border
Cape fur seal	Early-lactation female (on-shore) ~2–3 months pp	O2	Large breeding colony in Kleinsee, South Africa (29°40′′0S, 17°4′′60E), near the Namibian border
Subantarctic fur seal	Lactating female (on-shore) ~2 months pp	O*At*	La Mare aux Elephants (46°22′S, 51°40′E), Possession Island (Iles Crozet), southern Indian Ocean
Cape fur seal	Lactating female (at-sea)	S1	West of Cape Town, South Africa, past the continental shelf edge ~100 km from the nearest seal colony
Cape fur seal	Lactating female (at-sea)	S2	West of Cape Town, South Africa, past the continental shelf edge ~ 100 km from the nearest seal colony

into pGEM-T Easy (Promega Corporation) and sequenced using SP6 and/or T7 primers.

2.6. Northern blot analyses

The membrane used to evaluate *CTI* expression was prepared as previously described (Cane, 2005). Briefly, total RNA (10 μg) was electrophoresed at ≥5 V/cm through a 1% agarose, low-formaldehyde (1.1%) gel with 1× MOPS [3(N-Morpholino) Propane Sulfonic Acid] buffer at 4 °C. The RNA was transferred to Zeta-Probe GT Blotting Membrane (Bio-Rad) in 20× SSC (3.0 M sodium chloride, 0.3 M trisodium citrate, pH 7.0) overnight. The blot was rinsed briefly in 2× SSC, cross linked with UV light (120 mJ) and hybridised for 4 h at 42 °C in 25 mL [30% deionised formamide, 5× SSC, 50 mM sodium acetate, herring sperm DNA (100 μg/μL), 5 mL Denhart's 50× stock solution, 0.1% SDS]. The membrane was probed overnight at 42 °C with [α-^{32}P] dCTP-labelled fur seal *CTI* cDNA using the DECAprime II Random Priming DNA Labelling Kit (Ambion, Invitrogen, Carlsbad, CA, USA).

2.7. Sequence analysis

Physical characteristics of CTI were predicted using the web-based tools ProtParam (http://www.expasy.org/tools/protparam.html) and Pepstats (http://www.ebi.ac.uk/Tools/emboss/pepinfo/). Conserved protein motifs and post-translational modifications were predicted using the PROSITE database of protein domains, families and functional sites (http://prosite.expasy.org/) and the Center for Biological Sequence Analysis prediction servers: NetNGlyc 1.0 (N-glycosylation sites based on the Asn-Xaa-Ser/Thr sequon) (http://www.cbs.dtu.dk/services/NetNGlyc/) and NetPhos 2.0 (serine, threonine and tyrosine phosphorylation) (http://www.cbs.dtu.dk/services/NetPhos/). Signal peptides (indicative of secreted proteins) were predicted with SignalP 4.0 (http://www.cbs.dtu.dk/services/SignalP/). Sequences were aligned with Clustal W2 or Clustal Omega available at the EMBL-EBI website (http://www.ebi.ac.uk/services). BoxShade Server version 3.2 (http://www.ch.embnet.org/software/BOX_doc.html) was used to shade multiple sequence alignments. Similarity and identity between either nucleotide, or amino acid sequences were determined with MatGAT (Matrix Global Alignment Tool) 2.03 software (Campanella et al., 2003) using the BLOSUM50 matrix. MatGAT produces pairwise alignments and homology between each sequence pair. Genomic DNA and mRNA sequences and third party annotations were submitted to the GenBank nr database.

2.8. Structural homology modelling

Homology modelling was performed using Chimera (Pettersen et al., 2004) and Modeller (Šali and Blundell, 1993). BLAST searches of the NCBI GenBank nr (http://www.ncbi.nlm.nih.gov/BLAST/) and RSC PDB (Research Collaboratory for Structural Bioinformatics Protein Data Bank), (http://www.rcsb.org/pdb/home/home.do) were used to identify trypsinogen sequences and the latter to find the best KD 3D structure to model CTI, as well as the best model for the CTI-trypsin interaction. Models of Cape fur seal CTI [GenBank: AFZ99004; residues 33–94 inclusive], Pacific walrus trypsin [GenBank: XP_004416603; residues 24–246 inclusive] and a Cape fur seal CTI-Pacific walrus trypsin complex were constructed using the X-ray crystal structure of the second KD of tissue factor pathway inhibitor [GenBank: NP_006278; residues 121–178 inclusive] in complex with porcine trypsin [GenBank: NP_001156363; residues 24–246 inclusive], PDB ID: 1TFX. Pymol Version 1.3r1 (http://www.pymol.org/releases/v1.3r1) was used for model visualisation and preparation of Fig. 5.

3. Results

3.1. The Cape fur seal and Pacific walrus have a putative protein-coding CTI gene, but Weddell seal CTI encodes a pseudogene

To investigate whether pinnipeds have a functional *CTI* gene, we cloned the 2679 bp *CTI* gene from the Cape fur seal mammary gland [GenBank: KC152480] and identified the 2605 bp Pacific walrus *CTI* by BLAST searches of the walrus genome (Oros_1.0). Unexpectedly, Weddell seal *CTI* (LepWed1.0) was predicted to be a pseudogene, which if transcribed and translated would produce a ~6.7 kDa secreted protein lacking a KD (Supplementary file 2). Functional *CTI* genes were also predicted in the polar bear [GenBank: BK009384] (order Carnivora) and in marine mammals from the order Cetartiodactyla: killer whale, Scammon's minke whale and terrestrial species: sheep, goat, alpaca, wild yak, Tibetan antelope, Bactrian and Arabian camels and bison (Accession numbers are listed in Supplementary file 3). While *CTI* was thought to be a pseudogene in Perissodactyla, e.g. thoroughbred horse (Pharo et al., 2012) and Przewalski's horse (this study), another member of this order, the white rhinoceros, has a functional *CTI* gene (this study).

Consistent with other *CTI* genes (Pharo et al., 2012), pinniped *CTI* included three exons and two introns (Fig. 2; Supplementary file 4). However, pinniped *CTI* exon 1 was 12 nt shorter than that of other eutherians and encoded the signal peptide only. *CTI* gene identify was confirmed where possible, by the single-copy flanking genes, *phosphatidyl inositol glycan, class T* (*TPIG*) and *WAP four disulphide core domain 2* (*WFDC2*). A summary of *CTI* and *ELP* gene and *CTI* pseudogene distribution in mammals is provided (Supplementary file 5). An alignment of *CTI* and *ELP* transcripts is also provided (Supplementary file 6), as well as their evolutionary history (Supplementary file 7). Notably, *CTI* and *ELP* phylogeny was consistent with that of mammals.

3.2. Cape and subantarctic fur seal CTI are expressed in the mammary gland for at least 2–3 months postpartum

In order to investigate whether *CTI* is expressed in otariids, we amplified a 449 bp *CTI* transcript from the mammary gland of a pregnant Cape fur seal (P2) and used Northern analysis to characterise *CTI* expression in the mammary glands of pregnant, lactating at-sea, and lactating on-shore Cape fur seals, and from an on-shore lactating subantarctic fur seal (Fig. 3; Supplementary file 8). The expression of *LGB2* and *CSN2*, both of which encode major milk proteins of the whey and casein milk fractions respectively, were used as markers for prolactin-responsive whey protein gene expression and functional differentiation of the mammary gland (Rijnkels et al., 2013) respectively. *CTI* was detected in the mammary gland of the multiparous (P2), but not the primiparous (P1), pregnant seal, similar to the pattern of *CSN2* expression, while *LGB2* (Cane 2005; Sharp et al., 2005; Sharp et al., 2006a; Sharp et al., 2006b) was detected in both pregnant seals. *CTI*, *LGB2* and *CSN2* were expressed in the mammary glands of the lactating on-shore Cape (O3, O4) and subantarctic (O*At*) fur seals, which were nursing pups that were at least 2–3 months old. However, *CTI* transcripts were only detected in mammary tissue from one of the lactating females foraging at-sea. In contrast, *LGB2* and *CSN2* transcripts were detected in both lactating females at-sea.

3.3. Pinniped CTI, putative N-glycosylated Kunitz serine protease inhibitors

Next, we translated the pinniped *CTI* transcripts and determined that Cape fur seal and walrus *CTI* encode a putative 21 aa signal peptide and a 75 aa mature protein of ~8.47 and 8.59 kDa respectively, similar to other CTI and ELP peptides, 8.56–9.68 kDa (Supplementary file 9). As expected, mature Cape fur seal CTI was most similar to walrus CTI (93.3%), but less similar to Carnivora (78.5–83.5%), rhinoceros

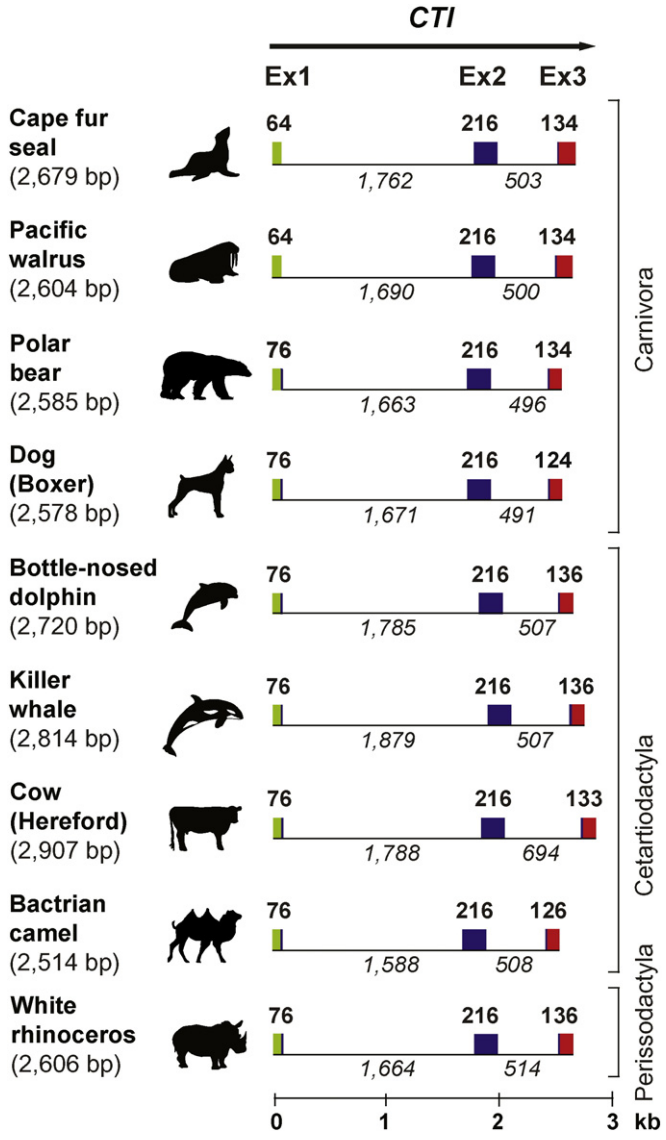

Fig. 2. Structure of pinniped *CTI* and other eutherian *CTI* genes. The conserved 3-exon structure of selected *CTI* genes is shown. Cape fur seal *CTI* [GenBank: KC152480], Pacific walrus [GenBank: LOC101362989], Polar bear [GenBank: BK009384; LOC103669871], dog [GenBank: BK008082], bottle-nosed dolphin [GenBank: BK008085; LOC101340157], killer whale [GenBank: LOC101273120], cow [GenBank: LOC104973896], Bactrian camel [GenBank: LOC105087824], white rhinoceros [GenBank: LOC101393639]. The *CTI* transcript encodes a signal peptide (green rectangle) and a mature secreted protein (blue), with a short non-coding 3′ UTR (red). Exon 2 which encodes the Kunitz domain is 216 bp. Exon size is shown in bold text and intron sizes are italicised. Gene size (within brackets) represents the number of nucleotides from the putative translation start (ATG, exon 1) to the polyadenylation signal (AATAAA, exon 3) inclusive.

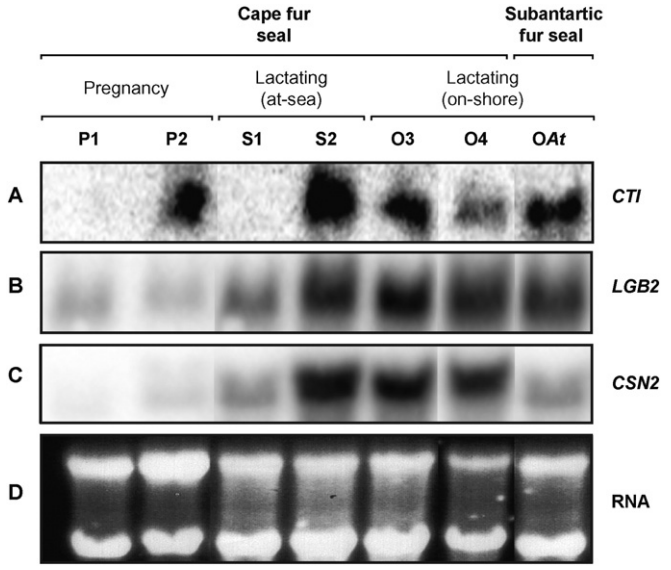

Fig. 3. Expression of *CTI* in the fur seal mammary gland. Northern analysis of total RNA (10 μg) extracted from the mammary glands of the Cape fur seal during late-pregnancy (~1 month pre-partum) and Cape and subantarctic fur seals during early-mid lactation (~2–3 months pp). A. *CTI* was not detected in the seal in its first pregnancy (P1), nor in one of the lactating seals foraging at-sea (S1). In contrast, *CTI* was expressed in the mammary glands of the multiparous seal during pregnancy (P2), in one of the lactating at-sea females (S2) and in the on-shore lactating Cape and subantarctic fur seals ~2–3 months pp (O3, O4, O*At*). B. *LGB2* was detected at all reproductive stages, with expression highest in the mammary glands of the on-shore lactating seals. C. *CSN2*, indicative of mammary gland differentiation, was detected in all animals except the primarous seal, with strong expression in the on-shore lactating Cape fur seals and one of the lactating seals at-sea, D. Ethidium bromide-stained ribosomal RNA bands show total RNA integrity and loading.

(Perissodactyla; 73.5%) and Cetartiodactyla CTI (67.5–71.1%), and marsupial ELP (49.4–56.8%) (Supplementary file 10).

Alignment of the pinniped CTI precursor proteins with CTI of selected species, plus the ubiquitously expressed BPTI (shown as cow PTI, Fig 4) and human TFPI KD2, upon which Cape fur seal CTI was modelled (Section 3.4), highlighted regions of similarity and difference (Fig 4). An alignment of all CTI and ELP precursor proteins is also provided (Supplementary file 11). Characteristic to the Kunitz/BPTI domain [BPTI KUNITZ 2 motif, Prosite: PS50279 and BPTI KUNITZ 1 motif: Prosite: PS00280, F-x(2)-{I}-G-C-x(6)-[FY]-x(5)-C; residues within square brackets are permitted and those within curly brackets are not], pinniped CTI shares six conserved cysteine residues which form three disulphide bonds: C1–C6, C2–C4 and C3–C5, corresponding to Cys19–Cys69, Cys28–Cys52, Cys44–Cys65 (pinniped protein numbering, Fig. 4; Supplementary files 11 and 12). The 51 aa CTI KD is flanked by non-structured, low complexity regions of 18–22 aa and 3–10 aa at the N- and C-termini respectively.

The putative interaction of Cape fur seal CTI with a serine protease, e.g. trypsin, is most likely mediated by two protease recognition loops (blue box, 29–37; green box, 52–59; alignment numbering Fig. 4), which are stabilised by the C2-C4 disulphide bond (Xu et al., 1998). Like BPTI (Hanson et al., 2003), CTI is predicted to be stabilised by the internal hydrophobic core residues: Tyr39–Phe40–Tyr41, although the camel has Ser40 (Fig. 4). Notably, a basic P_1 reactive site residue (Arg/Lys) was conserved in all sequences (#, residue 33, Fig 4, Supplementary file 11). Although Carnivora CTI and human TFPI KD2 P_1–P_1' (Arg–Gly) and Cetartiodactyla CTI and bovine PTI P_1–P_1' (Lys–Ala) were conserved, the rhinoceros CTI P_1–P_1' (Arg-Gly) were a combination of the two.

In contrast to the non-glycosylated BPTI (Creighton and Charles, 1987), CTI has a conserved putative N-glycosylation site at Asn42 (alignment numbering Fig. 4, Supplementary files 11 and 12) [ASN glycosylation motif, Prosite: PS00001, N-{P}-[ST]-{P}, where N is the sugar attachment site]. The N-glycosylation site has been confirmed for bovine CTI and is located at the 'base' of the pear-shaped protein (Klauser et al., 1978). TFPI KD2 is also N-glycosylated, but at Asn117 (Nakahara et al., 1996) (position 43 alignment numbering, Fig. 4), adjacent to the CTI sugar-attachment site.

All CTI peptides were predicted to be phosphorylated at Ser43 (ϕ, alignment numbering Fig. 4, Supplementary files 11 and 12) and all Carnivora CTI peptides phosphorylated by casein kinase II at Thr44 (alignment numbering Fig. 4, Supplementary files 11 and 12), [CK2 PHOSPHO SITE, Prosite: PS00006, [ST]-x(2)-[DE], residues 44–47, phosphorylation at position 1 (Thr44)].

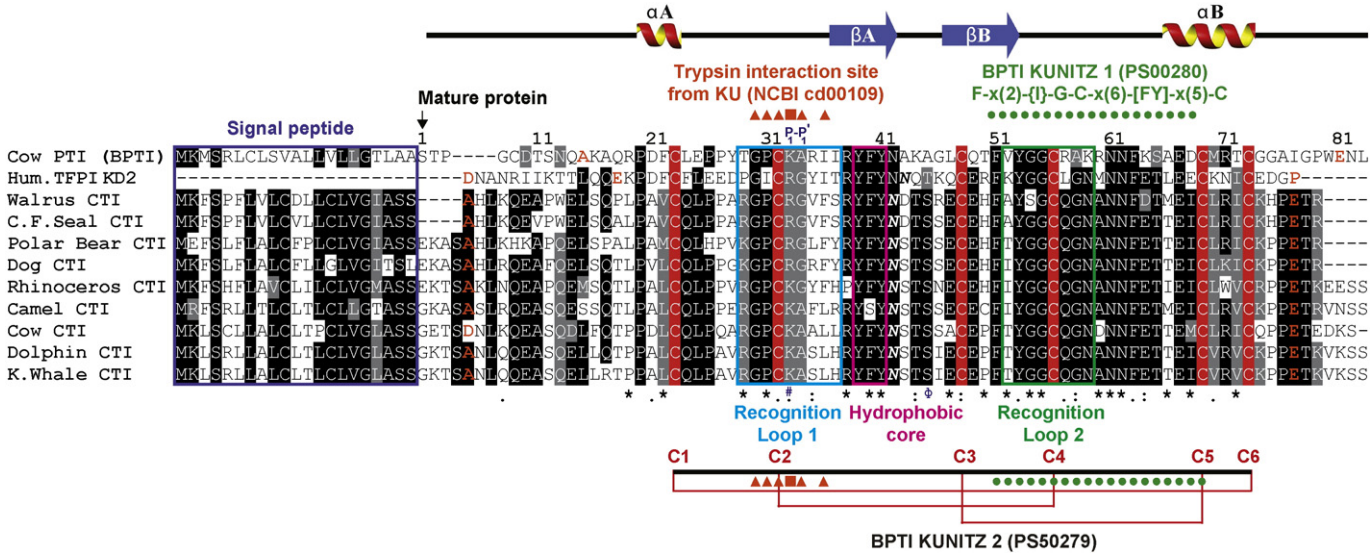

Fig. 4. Alignment of CTI, BPTI and the second Kunitz domain of human TFPI. Putative CTI and bovine PTI (BPTI) precursor proteins and KD2 of human TFPI were aligned with Clustal Omega. BPTI (cow PTI) [GenBank: AAI49369], human TFPI KD2 (includes part of the flanking regions, amino acids 79–150 inclusive shown, mature protein numbering) [GenBank: NP_006278] and CTI of the Pacific walrus [GenBank: XP_012422725], Cape fur seal CTI [GenBank: AFZ99004], Polar bear [GenBank: BK009384], dog [GenBank: AEN62470], white rhinoceros [GenBank: XP_004430613], Bactrian camel [GenBank: XP_006193597]; cow [GenBank: AEN62469], bottle-nosed dolphin [GenBank: BK008085; DAA35188] and killer whale [GenBank: XP_004272954]. Predicted signal peptides (dark blue border) and residues that span splice sites (orange text) are shown. The P_1 reactive site residue [#, the putative trypsin interaction site from the KU motif (NCBI cd00109)] and the BPTI_KUNITZ_1 and 2 motifs and are indicated. The two putative serine protease recognition loops (blue and green boxes respectively), the predicted hydrophobic core (pink box), N-glycosylation sites (bold, italicised text) and conserved serine phosphorylation site (ϕ) are shown. The predicted 3D-structure of CTI is indicated above the alignment: coiled regions, black bar; α-helices, red and yellow; and β-sheets, blue arrows. Conserved sites were shaded with BoxShade. Residues common to all species are indicated (*) and those shared by at least 19 of 27 species (>70%) are shaded black. Grey residues indicate conserved residue type. Conservation between groups of amino acids with strongly similar properties, i.e., score > 0.5 in the Gonnet PAM 250 matrix is indicated (:). Conservation between groups of amino acids with weakly similar properties (score < 0.5 in the Gonnet PAM 250 matrix) is also noted (.). Gaps within the alignment are indicated (−).

3.4. Structural homology modelling of Cape fur seal CTI suggests that it inhibits trypsin

In order to further investigate the protease inhibitory potential of Cape fur seal CTI, we constructed a homology model. High (54%) sequence identity with the second KD of human TFPI, for which an X-ray crystal structure exists [PDB ID: 1TFX, Chain D (Burgering et al., 1997)], predicted a high degree of structural conservation throughout the molecule and, in particular, in the protease-recognition loops (Fig. 5A). Apart from 3-residue and 2-residue insertions at the N- and C-termini respectively (both located distal to the inhibitory loop), there are no insertions or deletions.

We next constructed a model of the complex between Cape fur seal CTI and Pacific walrus trypsin, using the X-ray crystal structure of the second KD of human tissue factor pathway inhibitor in complex with porcine trypsin [PDB ID: 1TFX (Burgering et al., 1997)]. Neither Cape fur seal, nor any other otariid trypsin sequence was available and so we used the trypsin sequence of its closest relative, the walrus. As these species are from sister families, we expected minimal sequence difference. Notably, the active trypsin enzymes of the walrus (odobenid) and Weddell seal (phocid) share 89.7% identity, but less than with other Carnivora peptides (79.8–89.2%), Cetartiodactyla trypsin (69.5–77.1%), and human trypsin (68.3–71.4%) (Supplementary file 13).

Alignment of pinniped trypsinogen with the bovine, human, dog, polar bear and killer whale sequences shows the trypsin motif [TRYPSIN_DOM; Prosite: PS50240 (residues 9–229, alignment numbering; Fig. 6)], conservation of eight cysteine residues which form four disulphide bonds and stabilise the enzyme, as well as the conservation of the catalytic triad of residues (Ser185, His48 and Asp92; nucleophile-base–acid) which are necessary for protease-inhibitor binding (Vandermarliere et al., 2013). Like porcine trypsin (PDB ID: 1TFX, Chains A, B) and bovine trypsin (PDB ID: 4GUX, chain A), walrus trypsin is predicted to form two six-stranded beta-barrels (α-trypsin chain 1 and α-trypsin chain 2) separated by an interface comprising the catalytic residues (Vandermarliere et al., 2013).

Consistent with Kunitz inhibitor-serine protease complex formation (Burgering et al., 1997; Vandermarliere et al., 2013), homology modelling suggests that Cape fur seal CTI possesses the necessary structural and chemical features in order to inhibit trypsin (Fig. 5B). Strikingly, there were very few residue differences between the homology inhibitor-trypsin templates and the seal/walrus proteins in the vicinity of the protease-inhibitor interface, with no actual differences in the interface itself (Fig. 5B). The resulting features of the modelled walrus trypsin–Cape fur seal CTI complex, such as interface hydrogen bonding (Fig. 5C) and structural complementarity (Fig. 5D) were highly conserved with the X-ray crystallographically-determined mammalian complex (PDB ID: 1TFX). Interactions between the inhibitor P_1 residue (Arg in this case) and the protease S_1 specificity pocket typically make dominant energetic contributions to the complex stability and thus the inhibitory activity. Taken together, the modelling strongly suggests that seal CTI most likely possesses inhibitory activity against trypsin-like serine proteases in a highly similar manner to canonical inhibitors such as BPTI.

4. Discussion

A functional *colostrum trypsin inhibitor* gene was thought to be restricted to all species of the eutherian orders Carnivora and Cetartiodactyla (Pharo et al., 2012), but we have shown that although *CTI* is present in semi-aquatic marine mammals such as the Cape fur seal (Otariidae) and walrus (Odobenidae), it is a pseudogene in the Weddell seal (Phocidae), the only marine mammal with a disrupted *CTI* gene thus far. Furthermore, the identification of a putative protein-coding *CTI* gene in the white rhinoceros (family Rhinocerotidae, order Perissodactyla), suggests that *CTI* may also be present in other Perissodactyls, e.g. tapirs (family Tapiridae), with a disrupted *CTI* gene limited to members of the Equidae (this study; Pharo et al., 2012). The

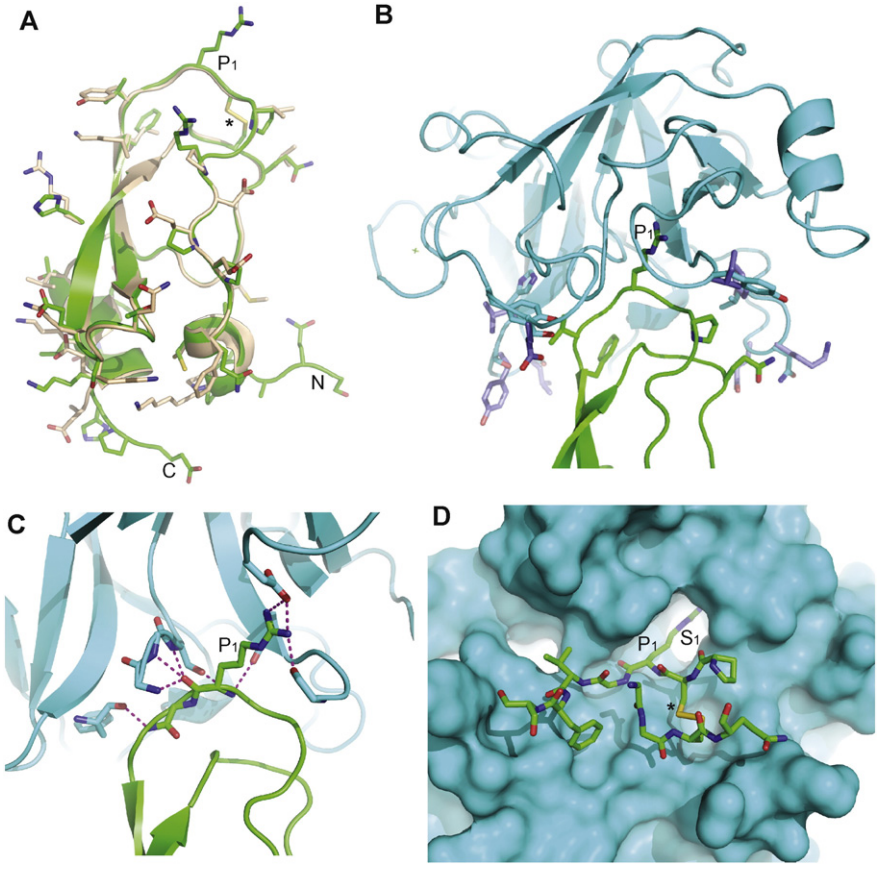

Fig. 5. Panel showing salient features of the modelled Cape fur seal CTI (green) and its complex with Pacific walrus trypsin (cyan). Orientations of the protease inhibitor are kept approximately the same throughout. A. Cartoon of Cape fur seal CTI superimposed on the second Kunitz domain of human tissue factor pathway inhibitor I (PDB ID: 1TFX chain D, wheat) on which it was modelled. The side chains of residues that differ in the sequence alignment are shown as sticks. N- and C-termini, showing 3- and 2-residue insertions, respectively, in Cape fur seal CTI are indicated; B. cartoon of model of Cape fur seal CTI (green) in complex with walrus trypsin (cyan). Interface residues that differ between the models and the homologues are shown as sticks (cyan = porcine trypsin, blue = walrus trypsin); C. Close-up cartoon of the walrus trypsin-Cape fur seal CTI complex showing hydrogen bonds (magenta) at the complex interface. A cross-section is shown where some residues have been removed for clarity; D. walrus trypsin-Cape fur seal CTI complex model showing trypsin molecular surface (cyan) and interacting residues of CTI (sticks). P_1 arginine of CTI fits snugly into the trypsin S_1 specificity pocket (both labelled). Panels B, C and D show the high degree of chemical and structural complementarity at the complex interface consistent with the proposed trypsin-inhibitory function of Cape fur seal CTI. The conserved disulphide bond in the protease inhibitor that anchors the inhibitory loop is shown as sticks (yellow S atoms) and marked with an asterisk. The P_1 residue is labelled in all panels.

prolonged period of *CTI* expression in the mammary gland of the Cape and subantarcic fur seals suggests that CTI is essential in these otariids. Additionally, the conservation of a basic P_1 residue (Lys/Arg) in all eutherian CTI peptides, unlike in present-day marsupial ELP, suggests that eutherian CTI acquired a basic P_1 residue after the divergence of the ancestral marsupials and eutherians ~160 million years ago (Luo et al., 2011). Structural homology modelling suggests that pinniped CTI binds to and inhibits trypsin in a canonical fashion, which may play an important role in neonate health.

4.1. Conservation of pinniped CTI gene structure

While the flanking genes of Cape fur seal *CTI* are unknown, the greater than 93% nucleotide identity shared been the Cape fur seal and walrus *CTI* genes, the latter of which is flanked by the single-copy *PIGT* and *WFDC2* genes, confirmed the identity of Cape fur seal *CTI*. The conservation of the 3-exon, 2-intron *CTI* structure and in particular, the 216 bp exon 2, which encodes a low-complexity 18 aa coiled region at its 5′ end (CTI N-terminus), a 51 aa Kunitz domain and a short 4 aa 3′ end suggests that conservation of exon 2 in its entirety may be important for CTI function. Furthermore, the eutherian Order-specific conservation of the 3–7 aa encoding CTI C-terminus suggests that this region may have an Order/species-specific function. The loss of *CTI* in the Weddell seal and other eutherians implies that either CTI function is provided by an alternate protein, or that CTI is not essential in these species and hence there is no evolutionary disadvantage to its loss.

4.2. The unique expression pattern of Cape fur seal CTI

The expression of *CTI* in the mammary glands of the on-shore lactating Cape and subantarctic fur seals 2–3 months pp, but the lack of *CTI* transcripts in one of the two lactating Cape fur seals foraging at-sea is consistent with reduced milk synthesis at-sea (>80%) in Antarctic (Arnould and Boyd, 1995) and Cape fur seals; and the resumption of milk production as the mother returns to shore to suckle her pup (Cane 2005; Sharp et al., 2005; Sharp et al., 2006a). Both the absence of lactose (the major osmole in most mammalian milks which directly affects milk volume) (Brew 2003; Neville et al., 1983) and a 'switch' in energy partitioning in the otariid mother, from the mobilisation of her body reserves for milk production on-shore, to replenishing her bodily stores while foraging at-sea (Costa 1991) may facilitate the 'suspension' of lactation at-sea without the gland involuting (Cane 2005; Sharp et al., 2005; Sharp et al., 2006a). In contrast, the cessation of sucking stimulus for ~48–72 h triggers mammary gland involution in most mammals (Watson 2006). Furthermore, the detection of *CSN2* and *LGB2* transcripts, but not *CTI* in one of the at-sea females suggests that *CTI* expression may be intermittent during lactation, i.e., up-regulated on-shore and down-regulated at-sea. Unlike other eutherians, otariid *CTI* may be expressed throughout lactation, rather than for only the first 24–

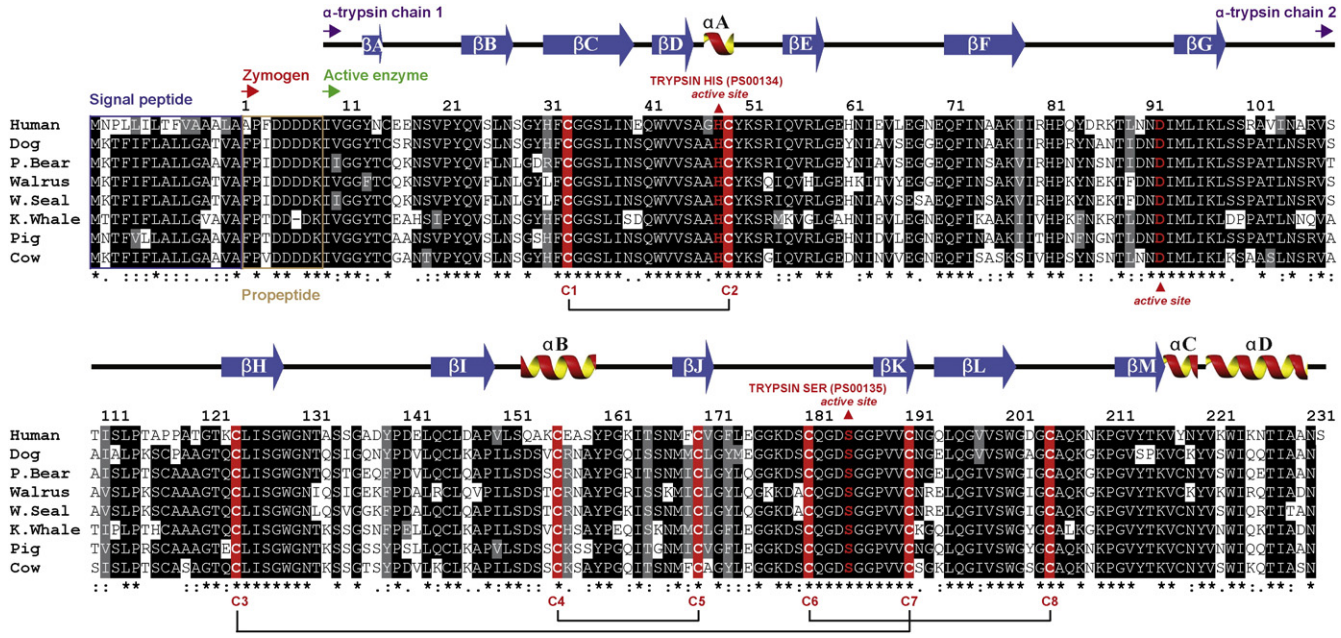

Fig. 6. Alignment of selected eutherian trypsinogen precursor proteins. Eutherian trypsinogen precursor proteins were aligned with Clustal Omega. Human [GenBank: NP_002760], dog [GenBank: P06871], Polar bear [GenBank: XP_008695541], Pacific walrus [GenBank: XP_004416603], Weddell seal [GenBank: XP_006743617], killer whale [GenBank: XP_004272168], pig [GenBank: NP_001156363], and cow [GenBank: NP_00110719]. The signal peptide and propeptide are boxed (blue and brown respectively) and the start site of the active trypsin enzyme shown (green arrow). The four conserved disulphide bonds are indicated, as well as the catalytic triad (bold red text). The 3D structure of trypsin [based on bovine trypsin (PDB ID: 1TFX chain A) and porcine trypsin (PDB ID: 1TFX chain A), is shown above the alignment: coiled regions, black bar; α-helices, red and yellow; and β-sheets, blue arrows. Conserved sites are shaded with BoxShade. Residues common to all species (*) are indicated and those shared by at least 6 of 8 species (>70%) are shaded black. Grey residues indicate conserved residue type. Conservation between groups of amino acids with strongly similar properties (score > 0.5, Gonnet PAM 250 matrix) is shown (:) and between those with weakly similar properties (score < 0.5, Gonnet PAM 250 matrix) also noted (.). Gaps within the alignment are indicated (—).

48 h pp. However, no conclusions could be drawn due to the limited number of animals available for this study. The lack of *CTI* expression in the pregnant, primiparous Cape fur seal mammary gland, in contrast to the multiparous seal at the same stage of gestation, may reflect the limited development of the primiparous gland (Cane, 2005), and its reduced secretory capacity (Lang et al., 2012).

4.3. Cape fur seal CTI — a putative trypsin inhibitor

Structural homology modelling based on human TFPI KD2 confirmed that Cape fur seal CTI forms a disulphide rich alpha + beta fold protein stabilised by 3 disulphide bonds. Although the Cape fur seal trypsin sequence is unknown, the high conservation between the active eutherian trypsin enzymes suggested that walrus trypsin was an excellent substitute for otariid trypsin for protease-inhibitor structural homology modelling using the human TFPI KD2-porcine trypsin template (PDB ID: 1TFX). Furthermore, the conservation of a basic P_1 Arg residue in Cape fur seal CTI and the catalytic triad of residues (Ser185, His48 and Asp92) in walrus trypsin strongly supported the inhibition of walrus trypsin by Cape fur seal CTI in a canonical mechanism of interaction characteristic to Kunitz inhibitors such as BPTI (Krowarsch et al., 2003; Laskowski and Kato 1980).

The precise *in vivo* target of eutherian CTI is unknown, but pinniped CTI may have similar anti-protease activity to bovine CTI. Hence, Cape fur seal and walrus CTI may inhibit trypsin, α-chymotrypsin and plasmin, but be inactive against β-chymotrypsin (Wu and Laskowski 1955), kallikrein, thrombin and renin (Pineiro et al., 1978). Similar to human TFPI KD2, CTI may also inhibit Factor Xa, a key component in the blood coagulation pathway (Burgering et al., 1997). In addition, the resistance of bovine CTI to pepsin digestion (Kassell and Laskowski, 1956) suggests that pinniped CTI may remain intact on its passage through the gastrointestinal tract and so may enter the circulatory system of the young pup. Therefore, CTI may have multiple sites of action against one or more serine peptidases in the maternal mammary gland and milk, as well as in the gastrointestinal tract and/or circulatory system/other organs of the young. Future studies of Cape fur seal CTI activity could involve the preparation and testing of recombinant CTI against serine endopeptidases.

4.4. Putative CTI functions in the mammary gland and neonate

Eutherian CTI may have multiple functions in the mother and young. In particular, as a putative serine protease inhibitor, it may prevent the digestion of milk proteins in the mammary gland and milk before their ingestion by the infant. The proteases in pinniped milk have yet to be characterised, but human milk has four major active proteases (Khaldi et al., 2014). These include a serum-derived trypsin-like enzyme and plasmin, plus elastase and cathepsin D, which are produced by the mammary gland. With the exception of cathepsin D, all are serine proteases, which may be inhibited by CTI. Notably, bovine milk contains neither trypsin, nor chymotrypsin, but does contain plasmin, one of the most active proteases in milk (Dallas et al., 2015). Hence, CTI may protect the maternal mammary gland and neonatal tissues from self-digestion by plasmin which is involved in tissue remodelling. CTI may also safeguard the neonate intestines and other tissues from maternal immune cells. These cells are absorbed across the intestinal epithelium and enter the newborn's circulatory system but release trypsin and elastase (Reber et al., 2008; Salmon, 2000). CTI may also influence cell-cell interactions by inhibiting trypsin and other enzymes that alter cell surface proteins.

The positive correlation of both CTI and IgG (immunoglobulin G) content in bovine colostrum (Laskowski and Laskowski, 1951; Pineiro et al., 1978; Quigley et al., 1995) and the passive transfer of immunoglobulins from mother to young via colostrum for the first 24–48 h pp (Veselsky et al., 1978) suggests that CTI protects IgG against proteolytic degradation in the neonate intestines (Brambell 1970; Quigley et al., 1995). Although the newborn calf has innate immunity, it has insufficient lymphocytes to mount an adaptive (specific) immune

response to protect its systemic and mucosal compartments (Brambell 1970; Salmon, 2000). Therefore, IgG transfer prior to gut 'closure' (cessation of passage of macromolecules) is essential for the survival of the young calf (Brambell, 1970; Quigley et al., 1995) and by analogy, the Cape fur seal pup.

4.5. Potential role of oligosaccharides in CTI function

The conservation of a sugar attachment site at Asn42 in all CTI and ELP peptides; with human TFPI KD2 glycosylated at an adjacent asparagine, in contrast to the non-glycosylated BPTI suggests that N-linked oligosaccharides are important for CTI function. The glycan attachment site at the protein's 'base' (Klauser et al., 1978) is unlikely to interfere with the P_1 protease reactive site at the 'opposite' end of the protein and so both 'ends' of CTI may have different biological roles. Oligosaccharides have anti-inflammatory and anti-adhesion properties acting as prebiotics, or as chemical messengers in cell-cell communication and may also bind directly to immune cells and influence their migration (Jost et al., 2015; Urashima et al., 2012; Zivkovic et al., 2011). Most importantly, oligosaccharides can act as soluble receptor analogues, or decoys for epithelial cell surface receptors which bind bacterial and viral pathogens (Jost et al., 2015; Kunz et al., 2014; Urashima et al., 2012). Hence microorganisms bind to these decoys, preventing infections in the young (Pacheco et al., 2015).

The cell- and species-specific nature of glycan production (Pacheco et al., 2015; Urashima et al., 2012) suggests that milk oligosaccharides may have unique functions. For example, milk oligosaccharides influence gut microbiota and gut maturation (Jost et al., 2015; Zivkovic et al., 2011). Many of these sugars are resistant to digestion (Pacheco et al., 2015) and so may pass intact through the intestines into the circulatory system and be transported to other tissues and organs (Kunz et al., 2014). Similarly, attached sugars may also increase the longevity of glycolipids and glycoproteins by preventing their non-specific degradation in the gut, as well as improving bioactivity, structure, stability and solubility (Roth et al., 2010). Hence N-glycosylated CTI may persist in the gut of the pup and enter the circulatory system. As the half-life of maternal IgG in newborn serum ranges from 12 to 20 days (Murphy et al., 2014), the extended period of *CTI* expression may provide prolonged protection for IgG and other milk proteins in the pup. This would be advantageous, particularly during the long periods of fasting that the pup experiences while its mother forages at-sea.

In summary, there appears to be no consistent set of characteristics that distinguish species that have retained the *CTI* gene from those that have a *CTI* pseudogene. However, *CTI* has not been reported in species that rely solely on the placental transfer of maternal immunoglobulins (humans, chimpanzees, rabbits). In contrast, *CTI* or the orthologous marsupial *ELP* gene are present in mammals that transfer immunoglobulins exclusively via colostrum (cow, pig, grey short-tailed opossum), or those that utilise both methods of transfer (dog, cat, pinnipeds, tammar wallaby) (Brambell 1970; Cavagnolo, 1979; Cavagnolo and Vedros, 1979; Deane et al., 1990). It would be of interest in the future to identify the target protease(s) of CTI and/or constituent(s) with which CTI interacts.

5. Conclusions

Colostrum trypsin inhibitor is expressed in the mammary gland of the lactating Cape and subantarctic fur seals (Otariidae) during the first 2–3 months pp, but unlike for other eutherians, *CTI* expression may be cyclical; up-regulated on-shore and down-regulated at-sea. The retention of a putative protein-coding *CTI* gene in the Pacific walrus (Odobenidae), but its disruption in the Weddell seal (Phocidae), was unexpected. Structural homology modelling of Cape fur seal CTI suggests that it, like other members of the BPTI/Kunitz family, inhibits trypsin-like serine proteases in a reversible, canonical mechanism. Therefore, pinniped CTI may prevent the proteolytic degradation of immunoglobulins and other milk proteins in the gastrointestinal tract and circulatory system of the young. Furthermore, glycosylated CTI may bind pathogenic bacteria in the pinniped pup gut, thereby preventing systemic infections and enhancing the survival of the young.

Acknowledgements

This research was funded by the Cooperative Research Centre for Innovative Dairy Products (CRC-IDP) and the Department of Zoology, The University of Melbourne. EAP was the recipient of an Australian Postgraduate Award (APA) and KNC a Dairy Australia Postgraduate Student Award and both EAP and KNC received a scholarship top-up from the CRC-IDP. AMB is a National Health and Medical Research Senior Research Fellow (1022688). The assistance of Mike Meyer, Chris Wilke, Deon Kotze, and Stefan Swanson in collecting Cape fur seal mammary tissue samples is gratefully acknowledged. We also thank the captain and crew of the *MV Sardinops* for logistical support without which this work could not have been conducted. Professor Marilyn Renfree provided insightful comments on this manuscript. This research was conducted under permits from the Department of Environmental Affairs and Tourism, Republic of South Africa.

Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.gene.2015.11.042.

References

Arnould, J.P.Y., Boyd, I.L., 1995. Temporal patterns of milk production in Antarctic fur seals (*Arctocephalus gazella*). J. Zool. (Lond.) 237, 1–12.
Ascenzi, P., Bocedi, A., Bolognesi, M., Spallarossa, A., Coletta, M., De Cristofaro, R., Menegatti, E., 2003. The bovine basic pancreatic trypsin inhibitor (Kunitz inhibitor): a milestone protein. Curr. Protein Pept. Sci. 4, 231–251.
Atkinson, S., 1997. Reproductive biology of seals. Rev. Reprod. 2, 175–194.
Bininda-Emonds, O.R., Cardillo, M., Jones, K.E., MacPhee, R.D., Beck, R.M., Grenyer, R., Price, S.A., Vos, R.A., Gittleman, J.L., Purvis, A., 2007. The delayed rise of present-day mammals. Nature 446, 507–512.
Bonner, W.N., 1984. Lactation strategies in pinnipeds: problems for a marine mammalian group. Symp. Zool. Soc. Lond. 51, 253–272.
Brambell, F.W.R., 1970. The Transmission of Passive Immunity From Mother to Young, Vol 18. North-Holland Publishing Company, Amsterdam.
Brew, K., 2003. α-Lactalbumin. In: Fox, P.F., McSweeney, P.L.H. (Eds.), Advanced Dairy ChemistryProteins vol. 1. Kluwer Academic/Plenum Publishers, New York, pp. 387–419.
Burgering, M.J., Orbons, L.P., van der Doelen, A., Mulders, J., Theunissen, H.J., Grootenhuis, P.D., Bode, W., Huber, R., Stubbs, M.T., 1997. The second Kunitz domain of human tissue factor pathway inhibitor: cloning, structure determination and interaction with factor Xa. J. Mol. Biol. 269, 395–407.
Campanella, J.J., Bitincka, L., Smalley, J., 2003. MatGAT: an application that generates similarity/identity matrices using protein or DNA sequences. BMC Bioinf. 4, 29.
Cane, K.N., 2005. The Physiological and Molecular Regulation of Lactation in Fur Seals. PhD Thesis, Department of Zoology. University of Melbourne, Melbourne.
Cañestro, C., Albalat, R., Irimia, M., Garcia-Fernàndez, J., 2013. Impact of gene gains, losses and duplication modes on the origin and diversification of vertebrates. Semin. Cell Dev. Biol. 24, 83–94.
Cavagnolo, R.Z., 1979. The immunology of marine mammals. Dev. Comp. Immunol. 3, 245–257.
Cavagnolo, R.Z., Vedros, N.A., 1979. Serum and colostrium immunoglobulin levels in the northern fur seal *Callorhinus ursinus*. Dev. Comp. Immunol. 3, 139–146.
Costa, D.P., 1991. Reproductive and foraging energetics of pinnipeds: implications for life history patterns. In: Renouf, D. (Ed.), The Behaviour of Pinnipeds. Springer, Dordrecht, pp. 300–344.
Creighton, T.E., Charles, I.G., 1987. Sequences of the genes and polypeptide precursors for two bovine protease inhibitors. J. Mol. Biol. 194, 11–22.
Dallas, D.C., Murray, N.M., Gan, J., 2015. Proteolytic systems in milk: perspectives on the evolutionary function within the mammary gland and the infant. J Mammary Gland Biol. Neoplasia 20, 133–147.
Deane, E.M., Cooper, D.W., Renfree, M.B., 1990. Immunoglobulin G levels in fetal and newborn tammar wallabies (*Macropus eugenii*). Reprod. Fertil. Dev. 2, 369–375.
Edwards, M.J., Hinds, L.A., Deane, E.M., Deakin, J.E., 2012. A review of complementary mechanisms which protect the developing marsupial pouch young. Dev. Comp. Immunol. 37, 213–220.
Feeney, R.E., Allison, R.G., 1969. Evolutionary Biochemistry of Proteins : Homologous and Analogous Proteins from Avian egg Whites, Blood Sera, Milk, and Other Substances. Wiley, New York.

Fry, B.G., Roelants, K., Champagne, D.E., Scheib, H., Tyndall, J.D., King, G.F., Nevalainen, T.J., Norman, J.A., Lewis, R.J., Norton, R.S., et al., 2009. The toxicogenomic multiverse: convergent recruitment of proteins into animal venoms. Annu. Rev. Genomics Hum. Genet. 10, 483–511.
Gamel, C.M., Davis, R.W., David, J.H., Meÿer, M.A., Brandon, E., 2005. Reproductive energetics and female attendance patterns of Cape fur seals (*Arctocephalus pusillus pusillus*) during early lactation. Am. Midl. Nat. 153, 152–170.
Garcia-Fernandez, R., Perbandt, M., Rehders, D., Ziegelmuller, P., Piganeau, N., Hahn, U., Betzel, C., Chavez Mde, L., Redecke, L., 2015. Three-dimensional structure of a Kunitz-type inhibitor in complex with an elastase-like enzyme. J. Biol. Chem. 290, 14154–14165.
Gentry, R.L., Kooyman, G.L., 1986. Fur Seals: Maternal Strategies on Land and at Sea. Princeton University Press, Princeton N. J.
Georges, J.-Y., Guinet, C., 2000. Maternal care in the subantarctic fur seals on Amsterdam Island. Ecology 81, 295–308.
Gojobori, T., Ikeo, K., 1994. Molecular evolution of serine protease and its inhibitor with special reference to domain evolution. Philos. Trans. R. Soc. Lond. Ser. B Biol. Sci. 344, 411–415.
Grzesiak, A., Helland, R., Smalas, A.O., Krowarsch, D., Dadlez, M., Otlewski, J., 2000. Substitutions at the P1′ position in BPTI strongly affect the association energy with serine proteinases. J. Mol. Biol. 301, 205–217.
Hanson, W.M., Beeser, S.A., Oas, T.G., Goldenberg, D.P., 2003. Identification of a residue critical for maintaining the functional conformation of BPTI. J. Mol. Biol. 333, 425–441.
Hawdon, J.M., Datu, B., Crowell, M., 2003. Molecular cloning of a novel multidomain Kunitz-type proteinase inhibitor from the hookworm *Ancylostoma caninum*. J. Parasitol. 89, 402–407.
Higdon, J.W., Bininda-Emonds, O.R., Beck, R.M., Ferguson, S.H., 2007. Phylogeny and divergence of the pinnipeds (Carnivora: Mammalia) assessed using a multigene dataset. BMC Evol. Biol. 7, 216.
Huber, R., Kukla, D., Ruhlmann, A., Steigemann, W., 1972. Pancreatic trypsin inhibitor (Kunitz). I. Structure and function. Cold Spring Harb. Symp. Quant. Biol. 36, 141–148.
Ikeo, K., Takahashi, K., Gojobori, T., 1992. Evolutionary origin of a Kunitz-type trypsin inhibitor domain inserted in the amyloid beta precursor protein of Alzheimer's disease. J. Mol. Evol. 34, 536–543.
Jamal, F., Pandey, P., Singh, D., Khan, M.Y., 2013. Serine protease inhibitors in plants: nature's arsenal crafted for insect predators. Phytochem. Rev. 12, 1–34.
Jost, T., Lacroix, C., Braegger, C., Chassard, C., 2015. Impact of human milk bacteria and oligosaccharides on neonatal gut microbiota establishment and gut health. Nutr. Rev. 73, 426–437.
Kassell, B., Laskowski, M., 1956. The comparative resistance to pepsin of six naturally occurring trypsin inhibitors. J. Biol. Chem. 219, 203–210.
Khaldi, N., Vijayakumar, V., Dallas, D.C., Guerrero, A., Wickramasinghe, S., Smilowitz, J.T., Medrano, J.F., Lebrilla, C.B., Shields, D.C., German, J.B., 2014. Predicting the important enzymes in human breast milk digestion. J. Agric. Food Chem. 62, 7225–7232.
Kirkman, S.P., Bester, M.N., Hofmeyr, G.J.G., Pistorius, P.A., Makhado, A.B., 2002. Pup growth and maternal attendance patterns in subantarctic fur seals. Afr. Zool. 37, 13–19.
Klauser, R., Cechova, D., Tschesche, H., 1978. The carbohydrate linkage site of cow colostrum trypsin inhibitor. Hoppe Seylers Z. Physiol. Chem. 359, 173–180.
Krowarsch, D., Cierpicki, T., Jelen, F., Otlewski, J., 2003. Canonical protein inhibitors of serine proteases. Cell. Mol. Life Sci. 60, 2427–2444.
Kruse, P.E., 1983. The importance of colostral immunoglobulins and their absorption from the intestine of the newborn animals. Ann. Rech. Vet. 14, 349–353.
Kunz, C., Kuntz, S., Rudloff, S., 2014. Bioactivity of human milk oligosaccharides. In: Moreno, J.F., Sanz, M.L. (Eds.), Food Oligosaccharides: Production, Analysis and Bioactivity. John Wiley & Sons, Ltd., pp. 1–20.
Lang, S.L., Iverson, S.J., Bowen, W.D., 2012. Primiparous and multiparous females differ in mammary gland alveolar development: implications for milk production. J. Exp. Biol. 215, 2904–2911.
Laskowski, M., Kato, I., 1980. Protein inhibitors of proteinases. Annu. Rev. Biochem. 49, 593–626.
Laskowski Jr., M., Laskowski, M., 1951. Crystalline trypsin inhibitor from colostrum. J. Biol. Chem. 190, 563–573.
Laskowski, M., Qasim, M.A., 2000. What can the structures of enzyme-inhibitor complexes tell us about the structures of enzyme substrate Complexes? Biochim. Biophys. Acta 1477, 324–337.
Luo, Z.X., Yuan, C.X., Meng, Q.J., Ji, Q., 2011. A Jurassic eutherian mammal and divergence of marsupials and placentals. Nature 476, 442–445.
Lynch, M., Conery, J.S., 2000. The evolutionary fate and consequences of duplicate genes. Science 290, 1151–1155.
Marquart, M., Walter, J., Deisenhofer, J., Bode, W., Huber, R., 1983. The geometry of the reactive site and of the peptide groups in trypsin, trypsinogen and its complexes with inhibitors. Acta Crystallogr. B 39, 480–490.
McFadden, T.B., Besser, T.E., Barrington, G.M., 1997. Regulation of immunoglobulin transfer into mammary secretions of ruminants. In: Welch, R.A.S., Burns, D.J.W., Davis, S.R., Popay, A.I., Prosser, C.G. (Eds.), Milk Composition, Production and Biotechnology. CAB International, New York, pp. 133–152.
Murphy, J.M., Hagey, J.V., Chigerwe, M., 2014. Comparison of serum immunoglobulin G half-life in dairy calves fed colostrum, colostrum replacer or administered with intravenous bovine plasma. Vet. Immunol. Immunopathol. 158, 233–237.
Nakahara, Y., Miyata, T., Hamuro, T., Funatsu, A., Miyagi, M., Tsunasawa, S., Kato, H., 1996. Amino acid sequence and carbohydrate structure of a recombinant human tissue factor pathway inhibitor expressed in Chinese hamster ovary cells: one N- and two O-linked carbohydrate chains are located between Kunitz domains 2 and 3 and one N-linked carbohydrate chain is in Kunitz domain 2. Biochemistry 35, 6450–6459.
Neville, M.C., Allen, J.C., Watters, C., 1983. The mechanisms of milk secretion. In: Neville, M.C., Neifert, M.R. (Eds.), Lactation: Physiology, Nutrition, and Breast-Feeding. New York, Plenum Publishing Co, pp. 49–92.
Nicholas, K.R., Simpson, K.J., Wilson, M., Trott, J., Shaw, D., 1997. The tammar wallaby: a model to study putative autocrine-induced changes in milk composition. J. Mammary Gland Biol. Neoplasia 2, 299–310.
Nyakatura, K., Bininda-Emonds, O.R., 2012. Updating the evolutionary history of Carnivora (Mammalia): a new species-level supertree complete with divergence time estimates. BMC Biol. 10, 12.
Oftedal, O.T., Boness, D.J., Tedman, R.A., 1987. The behavior, physiology, and anatomy of lactation in the pinnipedia. Curr. Mamm. 1, 175–245.
Pacheco, A.R., Barile, D., Underwood, M.A., Mills, D.A., 2015. The impact of the milk glycobiome on the neonate gut microbiota. Annu. Rev. Anim. Biosci. 3, 419–445.
Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., Ferrin, T.E., 2004. UCSF Chimera — a visualization system for exploratory research and analysis. J. Comput. Chem. 25, 1605–1612.
Pharo, E.A., De Leo, A.A., Renfree, M.B., Thomson, P.C., Lefevre, C.M., Nicholas, K.R., 2012. The mammary gland-specific marsupial *ELP* and eutherian *CTI* share a common ancestral gene. BMC Evol. Biol. 12, 8.
Pineiro, A., Brock, J.H., Esparza, I., 1978. Isolation and properties of bovine colostral trypsin inhibitor. Ann. Rech. Vet. 9, 281–286.
Piotte, C.P., Grigor, M.R., 1996. A novel marsupial protein expressed by the mammary gland only during the early lactation and related to the Kunitz proteinase inhibitors. Arch. Biochem. Biophys. 330, 59–64.
Quigley, J.D., Martin, K.R., Dowlen, H.H., 1995. Concentrations of trypsin inhibitor and immunoglobulins in colostrum of Jersey cows. J. Dairy Sci. 78, 1573–1577.
Ranasinghe, S., McManus, D.P., 2013. Structure and function of invertebrate Kunitz serine protease inhibitors. Dev. Comp. Immunol. 39, 219–227.
Rawlings, N.D., Waller, M., Barrett, A.J., Bateman, A., 2014. MEROPS: the database of proteolytic enzymes, their substrates and inhibitors. Nucleic Acids Res. 42, D503–D509.
Reber, A.J., Donovan, D.C., Gabbard, J., Galland, K., Aceves-Avila, M., Holbert, K.A., Marshall, L., Hurley, D.J., 2008. Transfer of maternal colostral leukocytes promotes development of the neonatal immune system part II. Effects on neonatal lymphocytes. Vet. Immunol. Immunopathol. 123, 305–313.
Rijnkels, M., Freeman-Zadrowski, C., Hernandez, J., Potluri, V., Wang, L., Li, W., Lemay, D.G., 2013. Epigenetic modifications unlock the milk protein gene loci during mouse mammary gland development and differentiation. PLoS One 8, e53270.
Roth, J., Zuber, C., Park, S., Jang, I., Lee, Y., Kysela, K.G., Le Fourn, V., Santimaria, R., Guhl, B., Cho, J.W., 2010. Protein N-glycosylation, protein folding, and protein quality control. Mol. Cells 30, 497–506.
Šali, A., Blundell, T.L., 1993. Comparative protein modelling by satisfaction of spatial restraints. J. Mol. Biol. 234, 779–815.
Salmon, H., 2000. Mammary gland immunology and neonate protection in pigs. Homing of lymphocytes into the MG. Adv. Exp. Med. Biol. 480, 279–286.
Sambrook, J., Russell, D.W., 2001. Molecular Cloning: A Laboratory Manual, Third edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.
Sharp, J.A., Cane, K.N., Lefevre, C., Arnould, J.P., Nicholas, K.R., 2006a. Fur seal adaptations to lactation: insights into mammary gland function. Curr. Top. Dev. Biol. 72, 275–308.
Sharp, J.A., Cane, K.N., Mailer, S.L., Oosthuizen, W.H., Arnould, J.P., Nicholas, K.R., 2006b. Species-specific cell–matrix interactions are essential for differentiation of alveoli like structures and milk gene expression in primary mammary cells of the Cape fur seal (*Arctocephalus pusillus pusillus*). Matrix Biol. 25, 430–442.
Sharp, J.A., Cane, K., Arnould, J.P., Nicholas, K.R., 2005. The lactation cycle of the fur seal. J. Dairy Res. 72, 81–89.
Shigetomi, H., Onogi, A., Kajiwara, H., Yoshida, S., Furukawa, N., Haruta, S., Tanase, Y., Kanayama, S., Noguchi, T., Yamada, Y., et al., 2010. Anti-inflammatory actions of serine protease inhibitors containing the Kunitz domain. Inflamm. Res. 59, 679–687.
Tschesche, H., Klauser, R., Cechova, D., Jonakova, V., 1975. On the carbohydrate composition of bovine colostrum trypsin inhibitor. Hoppe Seylers Z. Physiol. Chem. 356, 1759–1764.
Urashima, T., Fukuda, K., Messer, M., 2012. Evolution of milk oligosaccharides and lactose: a hypothesis. Animal 6, 369–374.
Vandermarliere, E., Mueller, M., Martens, L., 2013. Getting intimate with trypsin, the leading protease in proteomics. Mass Spectrom. Rev. 32, 453–465.
Verrier, D., Groscolas, R., Guinet, C., Arnould, J.P.Y., 2011. Development of fasting abilities in subantarctic fur seal pups: balancing the demands of growth under extreme nutritional restrictions. Funct. Ecol. 25, 704–717.
Veselsky, L., Cechova, D., Jonakova, V., 1978. Secretion and immunochemical properties of the trypsin inhibitor from bovine colostrum. Hoppe Seylers Z. Physiol. Chem. 359, 873–878.
Watson, C., 2006. Involution: apoptosis and tissue remodelling that convert the mammary gland from milk factory to a quiescent organ. Breast Cancer Res. 8, 203.
Wu, F.C., Laskowski, M., 1955. Action of the naturally occurring trypsin inhibitors against chymotrypsins α and β. J. Biol. Chem. 213, 609–619.
Xu, Y., Carr, P.D., Guss, J.M., Ollis, D.L., 1998. The crystal structure of bikunin from the inter-α-inhibitor complex: a serine protease inhibitor with two Kunitz domains. J. Mol. Biol. 276, 955–966.
Zivkovic, A.M., German, J.B., Lebrilla, C.B., Mills, D.A., 2011. Human milk glycobiome and its impact on the infant gastrointestinal microbiota. Proc. Natl. Acad. Sci. 108, 4653–4658.