

Could protein tertiary structure influence mammary transgene expression more than tissue specific codon usage?

Zuyong He · Yiqiang Zhao · Gui Mei ·
Ning Li · Yaosheng Chen

Received: 11 August 2009 / Accepted: 19 May 2010 / Published online: 19 June 2010
© The Author(s) 2010. This article is published with open access at Springerlink.com

Abstract Animal mammary glands have been successfully employed to produce therapeutic recombinant human proteins. However, considerable variation in animal mammary transgene expression efficiency has been reported. We now consider whether aspects of codon usage and/or protein tertiary structure underlie this variation in mammary transgene expression.

Keywords Transgenic animal bioreactor · Codon usage · Mammary gland · Tertiary structure

Zuyong He and Yiqiang Zhao contribute equally to this work.

Electronic supplementary material The online version of this article (doi:[10.1007/s11248-010-9411-8](https://doi.org/10.1007/s11248-010-9411-8)) contains supplementary material, which is available to authorized users.

Z. He · Y. Chen (✉)
State Key Laboratory of Biocontrol, School of Life Sciences, Sun Yat-Sen University, 510006 Guangzhou, People's Republic of China
e-mail: chyaosh@mail.sysu.edu.cn

Y. Zhao
Buck Institute for Age Research, California, USA

G. Mei
Institute of Animal Sciences, Guangdong Academy of Agricultural Sciences, 510640 Guangzhou, People's Republic of China

N. Li
State Key Laboratory for Agrobiotechnology, China Agricultural University, 100193 Beijing, People's Republic of China

Introduction

The expression of functional proteins in heterologous hosts is a cornerstone of modern biopharming. However, many human proteins are often difficult to express in unicellular organisms such as bacterium and yeast. The underlying problem is that codon bias has a profound impact on the heterologous expression of human proteins in these organisms. Codon usage has been found to be the single most important factor in prokaryotic gene expression. Therefore laborious and time consuming codon optimization is often necessary to achieve a successful expression in unicellular organisms (Gustafsson et al. 2004). Contrary to unicellular organisms, no obvious codon bias has been observed among human and several other mammals in previous studies. Thus animal mammary gland is considered an ideal bioreactor for producing functional human proteins without codon optimization. With this concept in mind, a number of transgenic livestock have been created to produce different recombinant human proteins in their milk, and in 2006 a recombinant human protein purified from a transgenic goat was approved for clinical use in Europe by the European Medicine Agency (Houdebine 2009).

However, considerable variation in expression efficiency has been found in the heterologous expressions of human proteins in the milk of transgenic animals produced in our lab and other groups. Generally a number of proteins such as serpin peptidase inhibitors and immunoglobulins which are abundant in human tissues other than mammary gland tend to exhibit a

higher expression level in transgenic milk, while many other human proteins such as interleukin-2, coagulation factor 8 and catalase are difficult to express in milk (Wright et al. 1991; Tang et al. 2008; Buhler et al. 1990; Niemann et al. 1999; He et al. 2008a, b). These examples cause us to question whether tissue specific codon usage pattern affects translational regulation during the heterologous expression of human proteins in animal mammary gland?

Recently several independent studies concluded that codon bias might be a factor involved in translation regulation in humans. One study concluded that genes selectively expressed in one human tissue can often be discriminated from genes expressed in another tissue on the basis of their synonymous codon usage (Plotkin et al. 2004) while another study reported that the amount of tRNA varies widely among different human tissues based on microarray results, further more they showed that the relative tRNA abundance significantly correlates with codon usage of tissue specific genes (Dittmar et al. 2006). The effective number of codons (ENC) is a most common index for measurement of codon bias. ENC is analogous to the effective number of alleles in population genetics. However, ENC can not reveal which codons are more frequent than others but rather indicates the overall departure from random synonymous codon choice. As a result, two genes may exhibit the same degree of overall bias, but differ dramatically in their particular choice of synonymous codons. Thus in this study, we used a two-tailed Fisher exact test to measure the distance of synonymous codon usage between two genes (Plotkin et al. 2004). Unlike metrics such as ‘‘relative synonymous codon usage’’, which are noisy when applied to individual genes, the Fisher exact test for small sample sizes based codon usage measure can be applied to genes that contain only a few examples of each amino acid.

We propose two scenarios: first, the different expression levels of recombinant human proteins in the milk of transgenic animals is due to the variation in synonymous codon usage patterns between mammary gland and other human tissues; or, second, protein tertiary structure may influence mammary transgene expression.

Human and mouse tissue specific codon usage

We have made pair-wise comparisons of codon usage among seven human tissues. When comparing heart

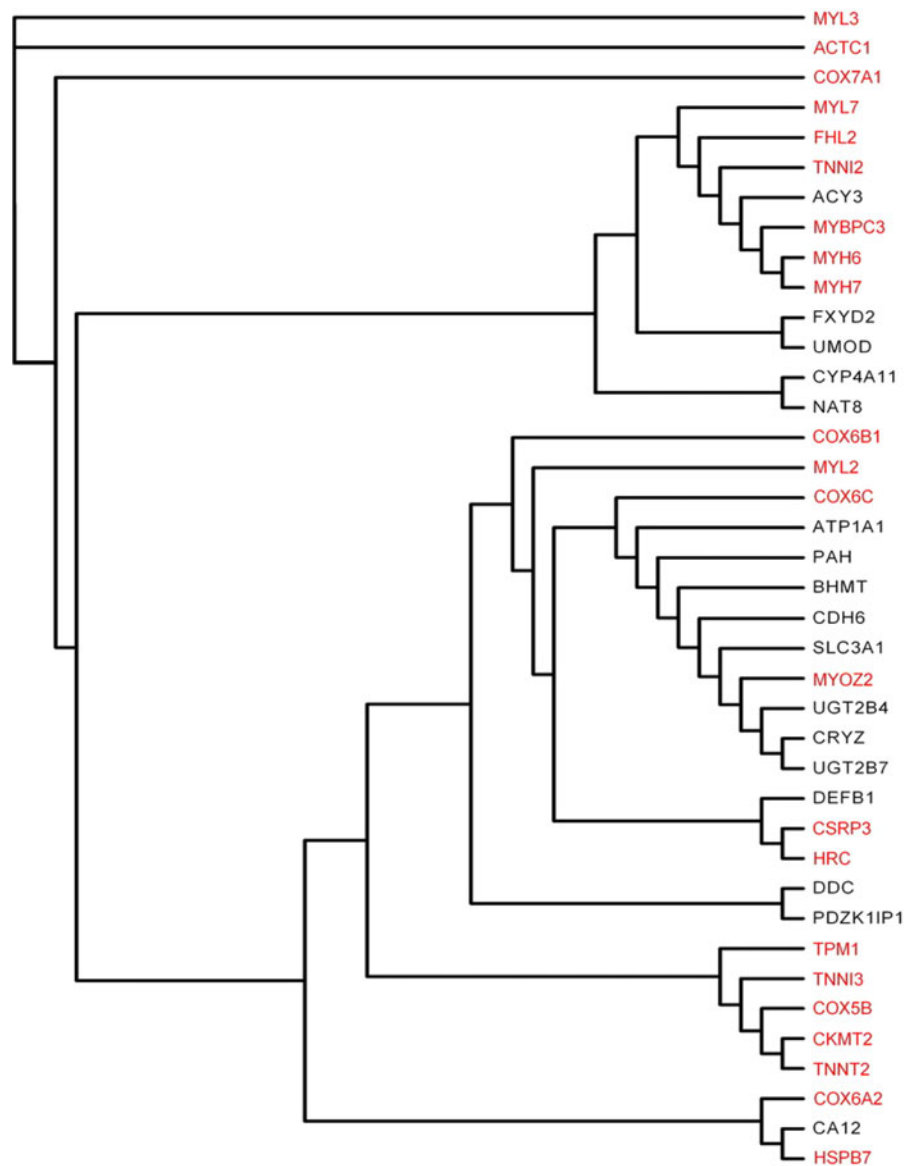
to kidney (Fig. 1), virtually all kidney associated genes are clustered in a separate middle clade from the heart associated genes. The observed separation between these two classes of genes would not have occurred by random chance ($P < 0.001$). The observed clustering is the result of systematic differential codon usage between heart and kidney specific genes. Fig. 1 indicates that we can generally discriminate between heart and kidney expressed genes on the basis of their codon usage alone. Similarly, kidney specific genes can be discriminated from lung and pancreas specific genes (supplementary Fig. 1). However, many pairs of tissue specific gene sets do not exhibit significantly different codon usage patterns (e.g., brain versus pancreas, $P = 0.384$; supplementary Fig. 2). Unexpectedly, in the tested six mouse tissues, we can not observe any pair of tissues that can be separated from each other on the basis of their codon usage with a statistical significant test. Only heart specific genes can be nearly discriminated from liver specific genes ($P = 0.072$; supplementary Fig. 3).

When comparing human mammary gland to six other tissues, only heart, lung and pancreas specific genes can be discriminated from mammary gland specific genes on the basis of codon usage (heart versus mammary gland, $P < 0.001$; lung versus mammary gland, $P = 0.002$; pancreas versus mammary gland, $P < 0.001$; supplementary Fig. 4), and the other three tissues in the test can not be distinguished from mammary gland (brain versus mammary gland, $P = 0.368$; kidney versus mammary gland, $P = 0.368$; liver versus mammary gland, $P = 0.536$; supplementary Fig. 5). In the tested six mouse tissues, only genes selectively expressed in heart and pancreas can be distinguished from mouse mammary gland specific genes (heart versus mammary gland, $P = 0.002$; pancreas versus mammary gland, $P = 0.038$; supplementary Fig. 6). Thus there does appear to be codon usage differences between mammalian tissues.

Does expression levels of recombinant human proteins in transgenic milk correlate with mammary gland specific codon usage?

Successful examples of expressing recombinant human proteins in transgenic animals to date are summarized in Table 1 and the greatest expression

Fig. 1 A dendrogram reflecting the codon usage of 22 genes selectively expressed in heart (*red*) and 17 genes selectively expressed in human kidney (*black*). The pairwise distances underlying this tree reflect the degree to which the genes differ in their codon usage. As this tree demonstrates, heart-specific genes can generally be distinguished from heart-specific genes purely on the basis of their synonymous codon usage. The observed separation between these two classes of genes would not have occurred by random chance ($P < 0.001$)



level of the 31 recombinant proteins reported in Table 1 were showed in Fig. 2. When comparing codon usage of the 31 recombinant proteins expressed in the milk of transgenic animals to human mammary gland specific genes and milk proteins, we found several most efficiently expressed recombinant human proteins (SERPINC1, ATCD20IgL, REG3A, LTF, FGA, FGB and FGG) were clustered close to the mammary gland specific genes and milk proteins. However, we also found the moderately expressed FIX and those less efficiently expressed recombinant proteins (LYZ, CAT, IL2 and mCol18a1) were

clustered close to most human mammary gland specific genes and milk proteins. The observed three classes of genes can not be discriminated from each other (mammary gland, $P = 0.574$; milk proteins, $P = 0.900$; supplementary Fig. 7). Similar results can be observed in the comparing of recombinant proteins to mouse mammary gland specific genes and milk proteins (mammary gland, $P = 0.952$; milk proteins, $P = 0.974$; supplementary Fig. 8). Further comparing codon usage of milk proteins among 19 different mammals showed that most milk proteins were prone to use similar codon usage patterns

Table 1 Expression of recombinant human proteins in the milk of transgenic animals

Protein	Promoter/vector	Gene structure	Transgenic animal	Expression level	Company	Reference
Anti-CD20 mAb	Goat beta-casein	cDNA	Mice	17.0 mg/mL		Tang et al. 2008
SERPINA1	Goat beta-casein		Goats	14.0 mg/mL	GTC biotherapeutics	GTC literature (1996)
SERPINA1	Goat beta-casein		Rabbits	4.0 mg/mL	GTC biotherapeutics	GTC literature (1996)
SERPINA1	Ovine BLG	Genomic DNA	Sheep	35.0 mg/mL		Wright et al. 1991
AFP	Goat beta-casein	Genomic DNA	Goats	0.6–1.1 mg/mL	GTC biotherapeutics	Parker et al. 2004
GAA	Bovine aS1-casein	Genomic DNA	Mice	2.0 mg/mL		Bijvoet et al. 1998
GAA	Bovine aS1-casein	Genomic DNA	Rabbits	8.0 mg/mL		Bijvoet et al. 1999
LALBA			Cattle	2.4 mg/mL		Krimpenfort et al. 1991
LALBA	Human a-LA	Genomic DNA	Cattle	1.55 mg/mL		Wang et al. 2008
LALBA	Human a-LA		Cattle	2.4 mg/mL	PPL therapeutics	PPL literature
SERPINC1	Goat beta-casein		Goats	20.0 mg/mL	GTC biotherapeutics	GTC literature(1996)
SERPINC1	Goat beta-casein		Goats	5.8 mg/mL		Baguisi et al. 1999
REG3A	Rabbit WAP	Genomic DNA	Mice	11.2 mg/mL		Christa et al. 2000
BCHE		cDNA	Mice	1.0–5.0 mg/mL		Baldassarre et al. 2008
BCHE		cDNA	Mice	0.1–5.0 mg/mL		Huang et al. 2007
CEL	Mouse WAP	cDNA	Mice	0.5–1.0 mg/mL		Stromqvist et al. 1996
SERPING1	Bovine aS1-casein		Rabbits	12.0 mg/mL	Pharming	Pharming literature online
CAT	Goat beta-casein	cDNA	Mice	145.3 ug/mL		He et al. 2008a, b
SOD3	Murine WAP	cDNA	Rabbits	2.9 mg/mL		Stromqvist et al. 1997
EPO	Bovine BLG	cDNA	Mice	0.3 mg/mL		Korhonen et al. 1997
EPO	Rabbit-WAP	cDNA	Rabbits	50.0 ug/mL		Massoud et al. 1996
EPO	Bovine BLG	cDNA	Rabbits	0.5 mg/mL		Korhonen et al. 1997
EPO	Adenovirus		Goats	2.0 mg/mL		Toledo et al. 2006
FIB	Ovine BLG	Genomic DNA	Mice	2.0 mg/mL		Prunkard et al. 1996
FIB	Ovine BLG		Sheep	5.0 mg/mL	PPL therapeutics	PPL literatur (1998)
FIB	Ovine BLG		Sheep	5.0 mg/mL		Butler et al. 1997
FIB	Bovine aS1-casein		Cattle	3.0 mg/mL	Pharming	
FIX	Mouse WAP	cDNA	Pigs	2.0–3.0 mg/mL		Lindsay et al. 2004
FIX	Ovine BLG		Sheep	25.0 ng/mL		Simons et al. 1987
FIX	Ovine BLG		Sheep	5.0 ng/mL		Clark 1998
FIX	Ovine BLG	Genomic DNA	Sheep	1.0 mg/mL	PPL therapeutics	Schnieke et al. 1997
FVII	Ovine BLG		Sheep	2.0 mg/mL	PPL therapeutics	PPL literature

Table 1 continued

Protein	Promoter/vector	Gene structure	Transgenic animal	Expression level	Company	Reference
FVIII	Mouse WAP	cDNA	Pigs	2.7 ug/mL		Paleyanda et al. 1997
FVIII	Ovine BLG	cDNA	Sheep	6.0 ng/mL		Niemann et al. 1999
GH	Rat beta-casein	Genomic DNA	Mice	19.0–5500.0 ug/mL		Lee et al. 1996
GH	Adenovirus	Genomic DNA	Mice	2.8 mg/mL		Sanchez et al. 2004
GH	Mouse WAP	Genomic DNA	Rabbits	50.0 ug/mL		Limonta et al. 1995
GH	Retrovirus	cDNA	Goats	60.0 ng/mL		Archer et al. 1994
GH	Adenovirus	Genomic DNA	Goats	0.3 mg/mL		Sanchez et al. 2004
GH	Bovine aS1-casein		Cattle	5.0 mg/mL	Bio Sidus SA	Salamone et al. 2006
IGFI	Bovine a-LA	cDNA	Pigs	228.0–1600.0 ug/mL		Monaco et al. 2005
IGFI	Bovine aS1-casein	cDNA	Rabbits	1.0 mg/mL		Brem et al. 1994
IL2	Rabbit beta-casein	Genomic DNA	Rabbits	430.0 ng/mL		Buhler et al. 1990
LTF	Human LT	F genomic DNA	Mice	8.02 mg/mL		Liu et al. 2004
LTF	Bovine aS1-casein	cDNA	Mice	36.0 ug/mL		Platenburg et al. 1994
LTF	s	Genomic DNA	Mice	6.6 mg/mL		Kim et al. 1999
LTF	Goat beta-casein	cDNA	Rabbits	153.0 ug/mL		Li et al. 2006
LTF	Adenovirus	cDNA	Goats	2.6 mg/mL		Han et al. 2007
LTF	Goat beta-casein	cDNA	Goats	0.765 mg/mL		Zhang et al. 2008
LTF	Bovine aS1-casein		Cattle	2.8 mg/mL	Pharming	van Berkel et al. 2002
LTF	Human LF	Genomic DNA	Cattle	3.4 mg/mL		Yang et al. 2008
LTF			Cattle	2.9 mg/mL		Hyvonen et al. 2006
LYZ	Goat beta-casein	Genomic DNA	Mice	1.405 mg/mL		Yu et al. 2006
NGF	Bovine aS1-casein	cDNA	Rabbits	50.0–250.0 ug/mL		Coulbaly et al. 1999
PROC	Ovine BLG		Rabbits	0.7 mg/mL	GTC Biotherapeutics	GTC literature (1996)
PROC	Mouse WAP	Genomic DNA	Rabbits	0.109–0.301ug/mL		Dragin et al. 2005
PROC	Mouse WAP	cDNA	Pigs	1.0 mg/mL	GTC Biotherapeutics	GTC literature (1992)
PROC	Mouse WAP	cDNA	Pigs	40.0–450.0 ug/mL		Van Cott et al. 2001
PROC	Mouse WAP	Genomic DNA	Pigs	160.0–1200.0 ug/mL		Van Cott et al. 2001
PROC	Ovine BLG		Pigs	0.75 mg/mL	PPL therapeutics	PPL literature
PROC	Ovine BLG		Sheep	0.3 mg/mL	PPL therapeutics	PPL literature (1998)
ALB	Bovine BLG	cDNA	Mice	2.5 mg/mL		Shami et al. 1992
tPA	Bovine aS1-casein	cDNA	Rabbits	8.0–50.0 ng/mL		Riego et al. 1993
tPA	Murine WAP	cDNA	Goats	3.0 ug/mL		Ebert et al. 1991

Table 1 continued

Protein	Promoter/vector	Gene structure	Transgenic animal	Expression level	Company	Reference
tPA	Goat beta-casein		Goats	3.0 mg/mL		Ebert et al. 1994
mCol18a1 ^a	Bovine α S1-casein	cDNA	Mice	70.0–300.0 ng/mL		Zavadskaia et al. 2001
Calc1 ^b	Ovine BLG	cDNA	Rabbits	1.0–2.1 mg/mL	PPL Therapeutics	McKee et al. 1998

^a mCol18a1 mouse collagen, type XVIII, alpha 1

^b Calc1 salmon calcitonin; The human calcitonin aggregates; therefore, the piscine equivalent was produced

among different mammalian species (supplementary Fig. 9). Thus we compared the codon usage of recombinant proteins to milk proteins in the main five livestock cow, sheep, goat, rabbit and pig. In each animal, except for certain proteins, the cluster result was quite similar to human and mouse (Fig. 3). Thus we dismiss our proposal that expression levels of recombinant human proteins in the milk of transgenic animals correlate with mammary gland specific codon usage patterns.

Recombinant human proteins with greater expression levels in transgenic milk share similar protein domains with milk proteins

The main domains of recombinant proteins expressed in the milk of transgenic animals and mammalian milk proteins derived from CATH were summarized in Table 2. Casein makes up the main component of milk proteins, as in bovine milk, it reaches as great as 82% of the total milk proteins (Jensen 1995). When we investigated the main domains of the main four caseins [casein alpha S1 (CSN1S1), casein alpha S2 (CSN1S2), casein beta (CSN2), and casein kappa (CSN3)], it can be observed that all four caseins share a similar alpha–beta based major domain (Fig. 4). With the exception of CSN1S2 which is composed of an alpha–beta barrel domain, the other three caseins all are composed of a 3-layer (aba) sandwich shaped major domain. Interestingly we found that within the top 13 high level expressed recombinant proteins in the milk of transgenic animals (expression level ≥ 5 mg/mL), 10 proteins share a similar alpha–beta based major domain with casein proteins, especially those with extra-higher expression levels such as SERPINA1, SERPINC1, SERPING1, REG3A, LTF and BCHE which all possess a 2-layer or 3-layer (aba) sandwich shaped major domain similar to the most abundant milk proteins CSN1S1, CSN2 and CSN3. Beside the 10 high level expressed proteins, there are another two proteins CEL (1.00 mg/mL) and Calc1 (2.10 mg/mL) each has a CSN2 like 3-layer (aba) sandwich domain and mainly alpha beta domain, respectively, though these two recombinant proteins do not get an extra-higher expression levels in the transgenic milk. However, the moderate expression levels are probably due to their cDNA based expression constructs. Because the gene structure used in an expression construct seems to have

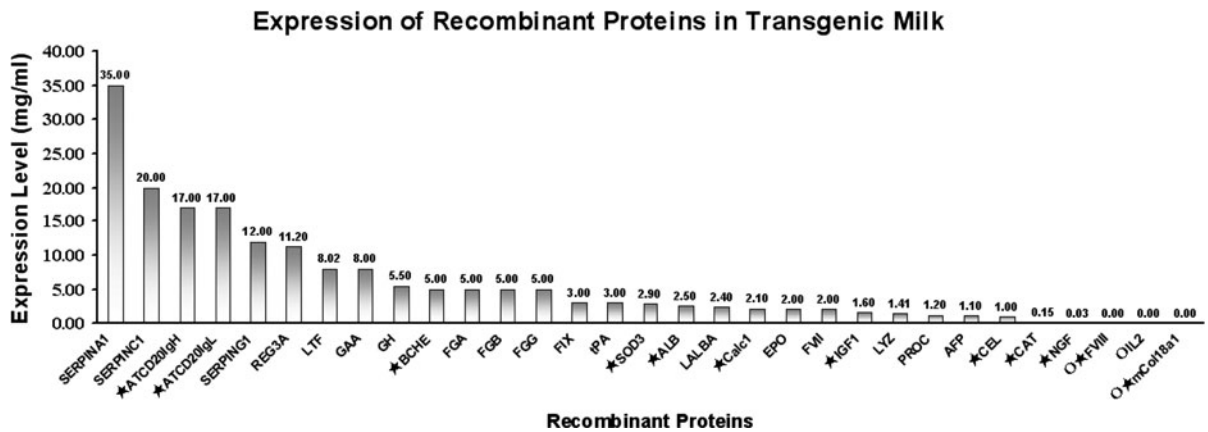


Fig. 2 Expression levels of recombinant human proteins in the milk of transgenic animals. The expression level of each protein presents the highest one among different studies summarized in Table 1; ATCD20 IgH and ATCD20 IgL indicate the heavy and light chain of human anti-CD20 monoclonal antibody; FGA, FGB and FGG note the alpha chain, beta chain and gamma chain of fibrinogen (FIB)

significant impact on expression level, generally a genomic DNA sequence results in several orders of magnitude greater expression level than a cDNA sequence (Whitelaw et al. 1991).

Promoter plays a critical role at transcriptional regulation of transgene expression. Milk protein gene promoters must be used when expressing recombinant proteins in animal mammary gland. Promoters derived from κ -casein and α S2-casein are particularly weak (Houdebine 2000). In Table 1, none of the reported studies used these two kinds of promoter. Bovine α S1-casein, goat β -casein, mouse murine acidic protein, and ovine β -lactoglobulin derived promoters were most popularly used in these studies. Under the regulation of ovine β -lactoglobulin promoter and using genomic DNA of the foreign gene, transgenic sheep expressed high level of SERPINA1 (35 mg/mL) and FIB (5 mg/mL) which both possessing major domains similar to caseins, whereas lower level of FIX (1 mg/mL) which has a major domain distinct with caseins. Similarly, under the control of goat beta-casein promoter, transgenic goats expressed high levels of SERPINC1 (20 mg/mL) and SERPINA1 (14 mg/mL), whereas lower levels of AFP (1.1 mg/mL) and tPA (3 mg/mL); furthermore, under the control of bovine α S2-casein, transgenic rabbits expressed high levels of SERPING1 (12 mg/mL) and GAA (8 mg/mL), whereas lower levels of

respectively; mCol18a1 stands for mouse collagen, type XVIII, alpha 1; *star* notes the expression level of recombinant protein with a cDNA based expression construct; *open circle* indicates the expression level values of hFVIII(0.0027 mg/mL), hIL-2(0.000043 mg/mL) and mCol18a1(0.00003 mg/mL) which all are too low to be fully presented on the top of the bar

tPA (0.05 mg/mL), NGF(0.25 mg/mL) and IGF1 (1 mg/mL). These cases may indicate protein structure appears to be an important factor affecting the expression level of transgene in the mammary gland. Taken together, we suggest that those recombinant proteins share similar major domains to casein proteins may have potential to achieve a greater expression level in the milk of transgenic animals when they are under similar transcriptional regulation.

Materials and methods

Gene sequences analysis

Coding sequences of genes in this study were obtained from GenBank. On the basis of BioGPS (<http://biogps.gnf.org/#goto=welcome>) and several other extensive mRNA expression microarray studies (Warrington et al. 2000; Hsiao et al. 2001; Liang et al. 2006; Shyamsundar et al. 2005; Saito-Hisaminato et al. 2002), we have identified genes which are selectively expressed in human and mouse mammary gland at lactation stage and in six other tissues: human [mammary gland (13 genes), brain (25 genes), heart (22 genes), kidney (17 genes), liver (26 genes), lung (19 genes) and

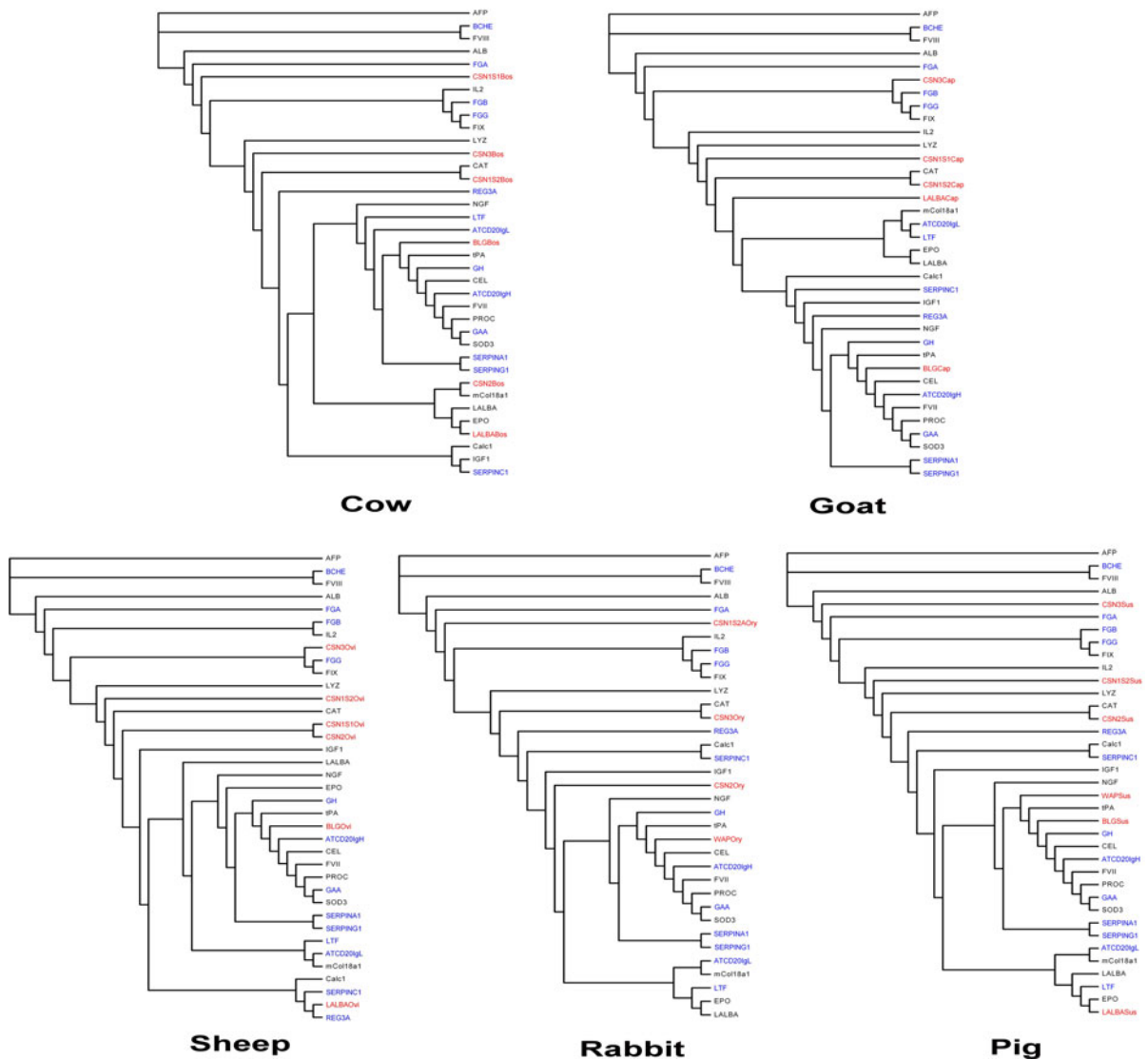


Fig. 3 Dendrograms reflecting the codon usage of milk proteins of the main five livestock (*red*) and recombinant proteins. The top 13 high level expressed recombinant proteins (≥ 5 mg/mL) are showed in *blue*, and moderate and lower expressed recombinant proteins are indicated in *black*. The

three classes of genes (indicated in *three different colour*) can not be clearly discriminated from each other (cow, $P = 0.992$; goat, $P = 0.982$; sheep, $P = 0.992$; rabbit, $P = 0.962$; pig, $P = 0.994$)

pancreas (27 genes; S1)] and mouse [mammary gland (15 genes), brain (24 genes), heart (22 genes), kidney (13 genes), liver (24 genes), lung (17 genes) and pancreas (21 genes; S2)].

The sequences of animal milk protein genes and recombinant proteins were summarized in supporting materials Table S3 and Table S4 respectively.

Codon usage analysis

For this study, the distance of synonymous codon usage between two genes were measured based on a two-tailed Fisher exact test method (Plotkin et al. 2004). Briefly, degree of codon bias in common sense as the departure from random synonymous codon choice is not concerned, but rather the degree to

Table 2 The main domains of recombinant proteins expressed in transgenic milk

Protein	CATH Code	Domain
IGF1	1.10.100.10	Mainly alpha; orthogonal bundle; insulin-like, subunit E; insulin-like, subunit E
AFP	1.10.246.10	Mainly alpha; orthogonal bundle; serum albumin; chain A, dDomain 1
ALB	1.10.246.10	Mainly alpha; orthogonal bundle; serum albumin; chain A, domain 1
LALBA	1.10.530.10	Mainly alpha; orthogonal bundle; lysozyme
LYZ	1.10.530.10	Mainly alpha; orthogonal bundle; lysozyme
EPO	1.20.1250.10	Mainly alpha; up-down bundle; growth hormone; chain: A;
GH ^a	1.20.1250.10	Mainly alpha; up-down bundle; growth hormone; chain A
IL2	1.20.1250.10	Mainly alpha; up-down bundle; growth hormone; chain A
NGF	2.10.90.10	Mainly beta; ribbon; cystine knot cytokines, subunit B; cystine-knot cytokines
FIX	2.40.10.10	Mainly beta; beta barrel; thrombin, subunit H; trypsin-like serine proteases
FVII	2.40.10.10	Mainly beta; beta barrel; thrombin, subunit H; trypsin-like serine proteases
FVIII	2.40.10.10	Mainly beta; beta barrel; thrombin, subunit H; trypsin-like serine proteases
PROC	2.40.10.10	Mainly beta; beta barrel; thrombin, subunit H; trypsin-like serine proteases
tPA	2.40.10.10	Mainly beta; beta barrel; thrombin, subunit H; trypsin-like serine proteases
BLG	2.40.128.20	Mainly beta; beta barrel; lipocalin
CAT	2.40.180.10	Mainly beta; beta barrel; catalase HpII, chain A, domain 1
ATCD20IgH ^a	2.60.40.10	Mainly beta; sandwich; immunoglobulin-like; immunoglobulins
ATCD21IgL ^a	2.60.40.10	Mainly beta; sandwich; immunoglobulin-like; immunoglobulins
SOD3	2.60.40.200	Mainly beta; sandwich; immunoglobulin-like
mCol18a1 ^b	3.10.100.10	Alpha beta; roll; mannose-binding protein A, chain A; mannose-binding protein A, subunit A
REG3A ^{a,b}	3.10.100.10	Alpha beta; roll; mannose-binding protein A, Chain A; mannose-binding protein A, subunit A
CSNIS2	3.20.20.70	Alpha beta; alpha-beta barrel; TIM Barrel; aldolase class I
GAA ^{a,b}	3.20.20.80	Alpha beta; alpha-beta barrel; TIM barrel; glycosidases
SERPINA1 ^{a,b}	3.30.497.10	Alpha beta; 2-layer sandwich; antithrombin; chain I, domain 2; antithrombin, subunit I, domain 2
SERPINC1 ^{a,b}	3.30.497.10	Alpha beta; 2-layer sandwich; antithrombin; chain I, domain 2; antithrombin, subunit I, domain 2
SERPING1 ^{a,b}	3.30.497.10	Alpha beta; 2-layer sandwich; antithrombin; chain I, domain 2; antithrombin, subunit I, domain 2
LTF ^{a,b}	3.40.190.10	Alpha beta; 3-layer(aba) sandwich; D-maltodextrin-binding protein; domain 2; periplasmic binding protein-like II
CSN2	3.40.50.10150	Alpha Beta; 3-Layer(aba) Sandwich; Rossmann fold; Diol Dehydratase; Chain B
BCHE ^{a,b}	3.40.50.1820	Alpha beta; 3-layer(aba) sandwich; rossmann fold
CEL ^{b,c}	3.40.50.1820	Alpha beta; 3-layer(aba) sandwich; rossmann fold
CSN1S1	3.40.630.10	Alpha beta; 3-layer(aba) sandwich; aminopeptidase; Zn peptidases
CSN3	3.40.710.10	Alpha beta; 3-layer(aba) sandwich; beta-lactamase; DD-peptidase/beta-lactamase superfamily
FGA ^{a,b}	3.90.215.10	Alpha beta; alpha-beta complex; gamma fibrinogen; chain A, domain 1
FGB ^{a,b}	3.90.215.10	Alpha beta; alpha-beta complex; gamma fibrinogen; chain A, domain 1
FGG ^{a,b}	3.90.215.10	Alpha beta; alpha-beta complex; gamma fibrinogen; chain A, domain 1
Calc1 ^{b,c}	3.90.320.20	Alpha beta; alpha-beta complex; lambda exonuclease; chain A
WAP	4.10.75.10	Few secondary structures; irregular; R-elafin; R-elafin

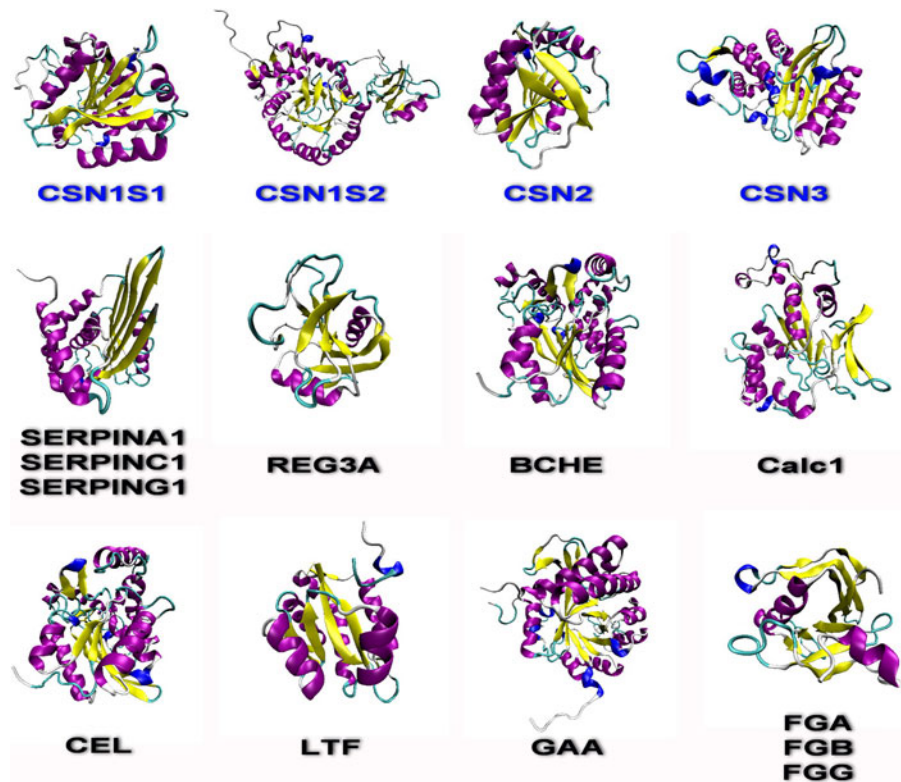
The italic bold words note milk proteins

^a The top 13 high level expressed recombinant proteins in the milk of transgenic animals (expression level ≥ 5 mg/mL)

^b The 13 recombinant proteins share similar domains with casein proteins

^c The expression level was based on a cDNA construct

Fig. 4 The tertiary structures of the main domains of casein proteins, 10 recombinant human proteins highly expressed in the milk of transgenic animal, and two moderately expressed proteins CEL (1.00 mg/mL) and Calc1 (2.10 mg/mL) based on cDNA expression constructs. Figures were made using VMD (<http://www.ks.uiuc.edu/Research/vmd/>) and rendered using Snapshot



which genes differ in their encoding of amino acids is concerned. Given the coding sequences for a pair of genes, absolute frequency of each codon in each gene is first tabulated with *condonW* (<http://codonw.sourceforge.net/>). For each amino acid, a two-tailed Fisher exact test on the $n \times 2$ contingency table given by the frequencies of the amino acid's synonymous codons is calculated (e.g., for Ser $n = 4$: TCA, TCC, TCG, and TCT). As a result, for each amino acid a *P*-value indicating whether or not the genes use significantly different codons to encode that amino acid is obtained. For example, when comparing human brain (25 genes) to heart (22 genes), distance between codon usage of every pair of genes (including pairs from the same tissue) is calculated, thus obtaining a 47-by-47 symmetric matrix of pairwise distances. Distance between two genes is given by the number of amino acids that exhibit significantly different ($P < 0.01$) codon usage, as defined above (detailed SAS program was presented in supplementary materials S5). Then by using the neighbor joining method (PHYLIP v3.66; <http://evolution.gs.washington.edu/phylip.html>), a dendrogram that graphically represents the measured pairwise distances between the codon

usage in the study genes can be produced. To test whether observed clustering of genes in a dendrogram is nonrandom, *P* value is calculated by comparing the observed summed distances along the tree between genes of the same tissue against a null distribution produced by randomly permuting the labels of the leaves.

Protein main domain analysis

Domains of recombinant proteins expressed in the milk of transgenic animals and mammalian milk proteins were evaluated with CATH (<http://www.cathdb.info/>) which is a hierarchical classification of protein domain structures based on clustering proteins at four major levels: Class (C), Architecture (A), Topology (T) and Homologous superfamily (H). Class is determined according to the secondary structure composition and packing within the structure. Three major classes are recognized; mainly-alpha, mainly-beta and alpha-beta. Architecture describes the overall shape of the domain structure as determined by the orientations of the secondary structures but ignores the connectivity between the

secondary structures. It is currently assigned manually using a simple description of the secondary structure arrangement e.g. barrel or 3-layer sandwich. Topology indicates that structure are grouped according to whether they share the same topology or fold in the core of the domain, that is, if they share the same overall shape and connectivity of the secondary structures in the domain core. Homologous superfamily groups together protein domains which are thought to share a common ancestor and can therefore be described as homologous. Similarities are identified either by high sequence identity or structure comparison using SSAP. Boundaries and assignments for each protein domain are determined using a combination of automated and manual procedures which include computational techniques, empirical and statistical evidence, literature review and expert analysis. Generally if a given protein chain has sufficiently high sequence identity and structural similarity (i.e. 80% sequence identity, SSAP score ≥ 80) with a chain that has previously been chopped, the domain boundary assignment is performed automatically by inheriting the boundaries from the other chain (ChopClose). Otherwise, the domain boundaries are assigned manually, based on an analysis of results derived from a range of algorithms which include structure based methods [CATHEDRAL, SSAP, DETECTIVE (Swindells, 1995)], PUU (Holm and Sander, 1994), DOMAK (Siddiqui and Barton, 1995), sequence based methods (Profile HMMs) and relevant literature.

How to improve mammary transgene expression?

We suspect that it is not necessary to consider codon usage optimization when targeting a human gene to express in the animal mammary gland. So we investigated the codon usage pattern between mammary gland and other tissues. Significant differences were found between mammary gland with heart and pancreas tissues both in human and mouse. However, no significant correlation was found between expression levels and codon usage of recombinant human proteins expressed in the milk of transgenic animals. This may indicate that even though a human gene shares a similar codon usage with a mammary gland specific gene, especially milk protein genes, does not

guarantee an effective expression in the mammary gland. In contrast, those proteins which share a similar domain with the four caseins are capable of achieving higher expression levels in animal mammary gland. We suppose that the tertiary structure of a recombinant protein adapts to the synthesis and secretion process of milk proteins especially caseins may permit it to be efficiently expressed in animal mammary gland.

In the secretory epithelial cell of mammary gland, milk protein precursors are assembled on the ribosomes of the highly developed rough endoplasmic reticulum (ER). All milk proteins have conserved secretory signal peptide sequences, which lead the growing nascent peptides to insert into the lumen of the ER. The proteins are then transported to the Golgi apparatus. The caseins then gradually intercalate with each other, calcium and phosphate to form a subcellular structures which lead to the formation of casein micelles, finally secreted by reverse pinocytosis (Farrell et al. 2006). Protein folding in the ER is monitored by ER quality control (ERQC) mechanisms. Proteins that pass ERQC criteria traffic to their final destinations through the secretory pathway, whereas non-native and unassembled subunits of multimeric proteins are degraded by the ER-associated degradation (ERAD) pathway (Vembar and Brodsky 2008). Overall, the environment of the ER lumen would be conducive to the proper casein-casein association, which helps them to escape ERAD and move onto the Golgi for processing and secretion. Thus the recombinant proteins as SERPINA1, SERPINC1, SERPING1 and BCHE, which have compact spherical structures composed of mainly alpha beta, 2-layer or 3-layer sandwich like domains, presumably form self-association or interact with caseins through the conserved structural elements, and escape the ERAD, and are efficiently transported to the Golgi apparatus to assemble micelles, and ultimately efficiently secreted into milk, lead to a high expression level in the transgenic milk. In contrast, AFP, CAT, FIX, IL2, NGF, and tPA all have a mainly alpha or beta barrel based domain which is quite different from casein (Supplementary Fig. 10). They may not interact with caseins properly and are aggregated to evoke the ERAD, thus fail to traffic to the Golgi apparatus, and ultimate lead to less efficient expression in the milk of transgenic animals.

Today it is still difficult to exactly predict the expression efficiency of one transgenic animal bioreactor. Because the expression efficiency of mammary transgene relies on a series of complex molecular regulation and cellular process. And it may be animal species dependent. Protein tertiary structure may be an important factor affecting the expression of animal mammary transgene. We doubt this is an absolute factor but, for example, for those proteins failing to secrete into transgenic milk even after transcriptional optimization, we might need to pay more attention to their tertiary structures. We may need to generate a group of elaborately designed gene targeting mice expressing structural distinct proteins under similar transcriptional regulation to verify this hypothesis. Furthermore, if for mammary transgenes encoding proteins that are structurally distinct from caseins such that they may pass through the ER/Golgi-dependent classical secretion pathway inefficiently, raises engineering the non-classical secretory pathway (He et al. 2008a, b; Nickel 2003) as an intriguing option.

Acknowledgments This work was supported by National S&T Major Project of China (grant no. 2008ZX08006-005).

Open Access This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

References

- Archer JS, Kennan WS, Gould MN, Bremel RD (1994) Human growth hormone (hGH) secretion in milk of goats after direct transfer of the hGH gene into the mammary gland by using replication-defective retrovirus vectors. *Proc Natl Acad Sci USA* 91:6840–6844. doi:10.1073/pnas.91.15.6840
- Baguisi A, Behboodi E, Melican DT, Pollock JS, Destrempes MM, Cammuso C, Williams JL, Nims SD, Porter CA, Midura P, Palacios MJ, Ayres SL, Denniston RS, Hayes ML, Ziomek CA, Meade HM, Godke RA, Gavin WG, Overstrom EW, Echelard Y (1999) Production of goats by somatic cell nuclear transfer. *Nat Biotechnol* 17:456–461. doi:10.1038/8632
- Baldassarre H, Hockley DK, Dore M, Brochu E, Hakier B, Zhao X, Bordignon V (2008) Lactation performance of transgenic goats expressing recombinant human butyrylcholinesterase in the milk. *Transgenic Res* 17:73–84. doi:10.1007/s11248-007-9137-4
- Bijvoet AG, Kroos MA, Pieper FR, Van der Vliet M, De Boer HA, Van der Ploeg AT, Verbeet MP, Reuser AJ (1998) Recombinant human acid alpha-glucosidase: high level production in mouse milk, biochemical characteristics, correction of enzyme deficiency in GSDII KO mice. *Hum Mol Genet* 7:1815–1824. doi:10.1093/hmg/7.11.1815
- Bijvoet AG, Van Hirtum H, Kroos MA, Van de Kamp EH, Schoneveld O, Visser P, Brakenhoff JP, Weggeman M, van Corven EJ, Van der Ploeg AT, Reuser AJ (1999) Human acid alpha-glucosidase from rabbit milk has therapeutic effect in mice with glycogen storage disease type II. *Hum Mol Genet* 8:2145–2153. doi:10.1093/hmg/8.12.2145
- Bosze Z, Baranyi M, Whitelaw CB (2008) Producing recombinant human milk proteins in the milk of livestock species. *Adv Exp Med Biol* 606:357–393
- Brem G, Hartl P, Besenfelder U, Wolf E, Zinovieva N, Pfaller R (1994) Expression of synthetic cDNA sequences encoding human insulin-like growth factor-1 (IGF-1) in the mammary gland of transgenic rabbits. *Gene* 149:351–355. doi:10.1016/0378-1119(94)90175-9
- Buhler TA, Bruyere T, Went DF, Stranzinger G, Burki K (1990) Rabbit beta-casein promoter directs secretion of human interleukin-2 into the milk of transgenic rabbits. *Biotechnology (N Y)* 8:140–143. doi:10.1038/nbt0290-140
- Butler SP, van Cott K, Subrumanian A, Gwazduaskas FC, Velandier WH (1997) Current progress in the production of recombinant human fibrinogen in the milk of transgenic animals. *Thromb Haemost* 78:537–542
- Christa L, Pauloin A, Simon MT, Stinnakre MG, Fontaine ML, Delpal S, Ollivier-Bousquet M, Brechot C, Devinoy E (2000) High expression of the human hepatocarcinoma-intestine-pancreas/pancreatic-associated protein (HIP/PAP) gene in the mammary gland of lactating transgenic mice. Secretion into the milk and purification of the HIP/PAP lectin. *Eur J Biochem* 267:1665–1671. doi:10.1046/j.1432-1327.2000.01159.x
- Clark AJ (1998) The mammary gland as a bioreactor: expression, processing, and production of recombinant proteins. *J Mammary Gland Biol Neoplasia* 3:337–350. doi:10.1023/A:1018723712996
- Coulibaly S, Besenfelder U, Fleischmann M, Zinovieva N, Grossmann A, Wozny M, Bartke I, Togel M, Muller M, Brem G (1999) Human nerve growth factor beta (hNGF-beta): mammary gland specific expression and production in transgenic rabbits. *FEBS Lett* 444:111–116. doi:10.1016/S0014-5793(98)01728-1
- Dittmar KA, Goodenbour JM, Pan T (2006) Tissue-specific differences in human transfer RNA expression. *PLoS Genet* 2:e221
- Dragin S, Chrastinova L, Makarevich A, Chrenek P (2005) Production of recombinant human protein C in the milk of transgenic rabbits from the F3 generation. *Folia Biol (Krakow)* 53:129–132. doi:10.3409/173491605775142684
- Ebert KM, Selgrath JP, DiTullio P, Denman J, Smith TE, Memon MA, Schindler JE, Monastersky GM, Vitale JA, Gordon K (1991) Transgenic production of a variant of human tissue-type plasminogen activator in goat milk: generation of transgenic goats and analysis of expression. *Biotechnology (N Y)* 9:835–838. doi:10.1038/nbt0991-835
- Ebert KM, DiTullio P, Barry CA, Schindler JE, Ayres SL, Smith TE, Pellerin LJ, Meade HM, Denman J, Roberts B (1994) Induction of human tissue plasminogen activator in

- the mammary gland of transgenic goats. *Biotechnology (N Y)* 12:699–702. doi:[10.1038/nbt0794-699](https://doi.org/10.1038/nbt0794-699)
- Farrell HM, Malin EL, Brown EM, Qi PX (2006) Casein micelle structure: what can be learned from milk synthesis and structural biology? *Curr Opin Colloid Inter Sci* 11:135–147. doi:[10.1016/j.cocis.2005.11.005](https://doi.org/10.1016/j.cocis.2005.11.005)
- Gustafsson C, Govindarajan S, Minshull J (2004) Codon bias and heterologous protein expression. *Trends Biotechnol* 22:346–353. doi:[10.1016/j.tibtech.2004.04.006](https://doi.org/10.1016/j.tibtech.2004.04.006)
- Han ZS, Li QW, Zhang ZY, Xiao B, Gao DW, Wu SY, Li J, Zhao HW, Jiang ZL, Hu JH (2007) High-level expression of human lactoferrin in the milk of goats by using replication-defective adenoviral vectors. *Protein Expr Purif* 53:225–231. doi:[10.1016/j.pep.2006.11.019](https://doi.org/10.1016/j.pep.2006.11.019)
- He ZY, Sun XZ, Mei G, Yu SL, Li N (2008a) Nonclassical secretion of human catalase on the surface of CHO cells is more efficient than classical secretion. *Cell Biol Int* 32:367–373. doi:[10.1016/j.cellbi.2007.12.003](https://doi.org/10.1016/j.cellbi.2007.12.003)
- He Z, Yu S, Mei G, Zheng M, Wang M, Dai Y, Tang B, Li N (2008b) Maternally transmitted milk containing recombinant human catalase provides protection against oxidation for mouse offspring during lactation. *Free Radic Biol Med* 45:1135–1142. doi:[10.1016/j.freeradbiomed.2008.07.019](https://doi.org/10.1016/j.freeradbiomed.2008.07.019)
- Holm L, Sander C (1994) Parser for protein folding units. *Proteins* 19:256–268. doi:[10.1002/prot.340190309](https://doi.org/10.1002/prot.340190309)
- Houdebine LM (2000) Transgenic animal bioreactors. *Transgenic Res* 9:305–320
- Houdebine LM (2009) Production of pharmaceutical proteins by transgenic animals. *Comp Immunol Microbiol Infect Dis* 32:107–121. doi:[10.1016/j.cimid.2007.11.005](https://doi.org/10.1016/j.cimid.2007.11.005)
- Hsiao LL, Dangond F, Yoshida T, Hong R, Jensen RV, Misra J, Dillon W, Lee KF, Clark KE, Haverty P, Weng Z, Mutter GL, Frosch MP, Macdonald ME, Milford EL, Crum CP, Bueno R, Pratt RE, Mahadevappa M, Warrington JA, Stephanopoulos G, Stephanopoulos G, Gullans SR (2001) A compendium of gene expression in normal human tissues. *Physiol Genomics* 7:97–104. doi:[10.1152/physiolgenomics.00040.2001](https://doi.org/10.1152/physiolgenomics.00040.2001)
- Huang YJ, Huang Y, Baldassarre H, Wang B, Lazaris A, Leduc M, Bilodeau AS, Bellemare A, Cote M, Herskovits P, Touati M, Turcotte C, Valeanu L, Lemeé N, Wilgus H, Begin I, Bhatia B, Rao K, Neveu N, Brochu E, Pierson J, Hockley DK, Cerasoli DM, Lenz DE, Karatzas CN, Langermann S (2007) Recombinant human butyrylcholinesterase from milk of transgenic animals to protect against organophosphate poisoning. *Proc Natl Acad Sci USA* 104:13603–13608. doi:[10.1073/pnas.0702756104](https://doi.org/10.1073/pnas.0702756104)
- Hyvonen P, Suojala L, Orro T, Haaranen J, Simola O, Rontved C, Pyorala S (2006) Transgenic cows that produce recombinant human lactoferrin in milk are not protected from experimental *Escherichia coli* intramammary infection. *Infect Immun* 74:6206–6212. doi:[10.1128/IAI.00238-06](https://doi.org/10.1128/IAI.00238-06)
- Jensen RG (1995) Handbook of milk composition. Academic press, San Diego
- Kim SJ, Sohn BH, Jeong S, Pak KW, Park JS, Park IY, Lee TH, Choi YH, Lee CS, Han YM, Yu DY, Lee KK (1999) High-level expression of human lactoferrin in milk of transgenic mice using genomic lactoferrin sequence. *J Biochem* 126:320–325
- Korhonen VP, Tolvanen M, Hyttinen JM, Uusi-Oukari M, Sinervirta R, Alhonen L, Jauhainen M, Janne OA, Janne J (1997) Expression of bovine beta-lactoglobulin/human erythropoietin fusion protein in the milk of transgenic mice and rabbits. *Eur J Biochem* 245:482–489. doi:[10.1111/j.1432-1033.1997.00482.x](https://doi.org/10.1111/j.1432-1033.1997.00482.x)
- Krimpenfort P, Rademakers A, Eyestone W, van der Schans A, van den Broek S, Kooiman P, Kootwijk E, Platenburg G, Pieper F, Strijker R et al. (1991) Generation of transgenic dairy cattle using ‘in vitro’ embryo production. *Biotechnology (N Y)* 9:844–847. doi:[10.1038/nbt0991-844](https://doi.org/10.1038/nbt0991-844)
- Lee CS, Kim K, Yu DY, Lee KK (1996) An efficient expression of human growth hormone (hGH) in the milk of transgenic mice using rat beta-casein/hGH fusion genes. *Appl Biochem Biotechnol* 56:211–222
- Li L, Shen W, Min L, Dong H, Sun Y, Pan Q (2006) Human lactoferrin transgenic rabbits produced efficiently using dimethylsulfoxide-sperm-mediated gene transfer. *Reprod Fertil Dev* 18:689–695. doi:[10.1071/RD060001](https://doi.org/10.1071/RD060001)
- Liang S, Li Y, Be X, Howes S, Liu W (2006) Detecting and profiling tissue-selective genes. *Physiol Genomics* 26:158–162. doi:[10.1152/physiolgenomics.00313.2005](https://doi.org/10.1152/physiolgenomics.00313.2005)
- Limonta JM, Castro FO, Martinez R, Puentes P, Ramos B, Aguilar A, Lleonart RL, de la Fuente J (1995) Transgenic rabbits as bioreactors for the production of human growth hormone. *J Biotechnol* 40:49–58. doi:[10.1016/0168-1656\(95\)00026-M](https://doi.org/10.1016/0168-1656(95)00026-M)
- Lindsay M, Gil GC, Cadiz A, Velandar WH, Zhang C, Van Cott KE (2004) Purification of recombinant DNA-derived factor IX produced in transgenic pig milk and fractionation of active and inactive subpopulations. *J Chromatogr A* 1026:149–157. doi:[10.1016/j.chroma.2003.11.006](https://doi.org/10.1016/j.chroma.2003.11.006)
- Liu Z, Zhao C, Fan B, Dai Y, Zhao Z, Wang L, Zheng M, Feng J, Chen Y, Duan Y, Li N (2004) Variable expression of human lactoferrin gene in mice milk driven by its 90 KB upstream flanking sequences. *Anim Biotechnol* 15:21–31. doi:[10.1081/ABIO-120029810](https://doi.org/10.1081/ABIO-120029810)
- Massoud M, Attal J, Thepot D, Pointu H, Stinnakre MG, Theron MC, Lopez C, Houdebine LM (1996) The deleterious effects of human erythropoietin gene driven by the rabbit whey acidic protein gene promoter in transgenic rabbits. *Reprod Nutr Dev* 36:555–563. doi:[10.1051/rnd:19960511](https://doi.org/10.1051/rnd:19960511)
- McKee C, Gibson A, Dalrymple M, Emslie L, Garner I, Cottingham I (1998) Production of biologically active salmon calcitonin in the milk of transgenic rabbits. *Nat Biotechnol* 16:647–651. doi:[10.1038/nbt0798-647](https://doi.org/10.1038/nbt0798-647)
- Monaco MH, Gronlund DE, Bleck GT, Hurley WL, Wheeler MB, Donovan SM (2005) Mammary specific transgenic over-expression of insulin-like growth factor-I (IGF-I) increases pig milk IGF-I and IGF binding proteins, with no effect on milk composition or yield. *Transgenic Res* 14:761–773. doi:[10.1007/s11248-005-7219-8](https://doi.org/10.1007/s11248-005-7219-8)
- Nickel W (2003) The mystery of nonclassical protein secretion. A current view on cargo proteins and potential export routes. *Eur J Biochem* 270:2109–2119. doi:[10.1046/j.1432-1033.2003.03577.x](https://doi.org/10.1046/j.1432-1033.2003.03577.x)
- Niemann H, Halter R, Carnwath JW, Herrmann D, Lemme E, Paul D (1999) Expression of human blood clotting factor VIII in the mammary gland of transgenic sheep. *Transgenic Res* 8:237–247. doi:[10.1023/A:1008999622117](https://doi.org/10.1023/A:1008999622117)

- Paleyanda RK, Velander WH, Lee TK, Scandella DH, Gwazdauskas FC, Knight JW, Hoyer LW, Drohan WN, Lubon H (1997) Transgenic pigs produce functional human factor VIII in milk. *Nat Biotechnol* 15:971–975. doi: [10.1038/nbt1097-971](https://doi.org/10.1038/nbt1097-971)
- Parker MH, Birck-Wilson E, Allard G, Masiello N, Day M, Murphy KP, Paragas V, Silver S, Moody MD (2004) Purification and characterization of a recombinant version of human alpha-fetoprotein expressed in the milk of transgenic goats. *Protein Expr Purif* 38:177–183. doi: [10.1016/j.pep.2004.07.007](https://doi.org/10.1016/j.pep.2004.07.007)
- Platenburg GJ, Kootwijk EP, Kooiman PM, Woloshuk SL, Nuijens JH, Krimpenfort PJ, Pieper FR, de Boer HA, Strijker R (1994) Expression of human lactoferrin in milk of transgenic mice. *Transgenic Res* 3:99–108. doi: [10.1007/BF01974087](https://doi.org/10.1007/BF01974087)
- Plotkin JB, Robins H, Levine AJ (2004) Tissue-specific codon usage and the expression of human genes. *Proc Natl Acad Sci USA* 101:12588–12591. doi: [10.1073/pnas.0404957101](https://doi.org/10.1073/pnas.0404957101)
- Prunkard D, Cottingham I, Garner I, Bruce S, Dalrymple M, Lasser G, Bishop P, Foster D (1996) High-level expression of recombinant human fibrinogen in the milk of transgenic mice. *Nat Biotechnol* 14:867–871. doi: [10.1038/nbt0796-867](https://doi.org/10.1038/nbt0796-867)
- Riego E, Limonta J, Aguilar A, Perez A, de Armas R, Solano R, Ramos B, Castro FO, de la Fuente J (1993) Production of transgenic mice and rabbits that carry and express the human tissue plasminogen activator cDNA under the control of a bovine alpha S1 casein promoter. *Theriogenology* 39:1173–1185. doi: [10.1016/0093-691X\(93\)90015-W](https://doi.org/10.1016/0093-691X(93)90015-W)
- Saito-Hisaminato A, Katagiri T, Kakiuchi S, Nakamura T, Tsunoda T, Nakamura Y (2002) Genome-wide profiling of gene expression in 29 normal human tissues with a cDNA microarray. *DNA Res* 9:35–45. doi: [10.1093/dnares/9.2.35](https://doi.org/10.1093/dnares/9.2.35)
- Salamone D, Baranao L, Santos C, Bussmann L, Artuso J, Werning C, Prync A, Carbonetto C, Dabsys S, Munar C, Salaberry R, Berra G, Berra I, Fernandez N, Papouchado M, Foti M, Judewicz N, Mujica I, Munoz L, Alvarez SF, Gonzalez E, Zimmermann J, Criscuolo M, Melo C (2006) High level expression of bioactive recombinant human growth hormone in the milk of a cloned transgenic cow. *J Biotechnol* 124:469–472. doi: [10.1016/j.jbiotec.2006.01.005](https://doi.org/10.1016/j.jbiotec.2006.01.005)
- Sanchez O, Toledo JR, Rodriguez MP, Castro FO (2004) Adenoviral vector mediates high expression levels of human growth hormone in the milk of mice and goats. *J Biotechnol* 114:89–97. doi: [10.1016/j.jbiotec.2004.06.009](https://doi.org/10.1016/j.jbiotec.2004.06.009)
- Schnieke AE, Kind AJ, Ritchie WA, Mycock K, Scott AR, Ritchie M, Wilmut I, Colman A, Campbell KH (1997) Human factor IX transgenic sheep produced by transfer of nuclei from transfected fetal fibroblasts. *Science* 278:2130–2133. doi: [10.1126/science.278.5346.2130](https://doi.org/10.1126/science.278.5346.2130)
- Shani M, Barash I, Nathan M, Ricca G, Searfoss GH, Dekel I, Faerman A, Givol D, Hurwitz DR (1992) Expression of human serum albumin in the milk of transgenic mice. *Transgenic Res* 1:195–208. doi: [10.1007/BF02524750](https://doi.org/10.1007/BF02524750)
- Shyamsundar R, Kim YH, Higgins JP, Montgomery K, Jorden M, Sethuraman A, van de Rijn M, Botstein D, Brown PO, Pollack JR (2005) A DNA microarray survey of gene expression in normal human tissues. *Genome Biol* 6:R22. doi: [10.1186/gb-2005-6-3-r22](https://doi.org/10.1186/gb-2005-6-3-r22)
- Siddiqui AS, Barton GJ (1995) Continuous and discontinuous domains: an algorithm for the automatic generation of reliable protein domain definitions. *Protein Sci* 4:872–884. doi: [10.1002/pro.5560040507](https://doi.org/10.1002/pro.5560040507)
- Simons JP, McClenaghan M, Clark AJ (1987) Alteration of the quality of milk by expression of sheep beta-lactoglobulin in transgenic mice. *Nature* 328:530–532
- Stromqvist M, Tornell J, Edlund M, Edlund A, Johansson T, Lindgren K, Lundberg L, Hansson L (1996) Recombinant human bile salt-stimulated lipase: an example of defective O-glycosylation of a protein produced in milk of transgenic mice. *Transgenic Res* 5:475–485. doi: [10.1007/BF01980213](https://doi.org/10.1007/BF01980213)
- Stromqvist M, Houdebine M, Andersson JO, Edlund A, Johansson T, Viglietta C, Puissant C, Hansson L (1997) Recombinant human extracellular superoxide dismutase produced in milk of transgenic rabbits. *Transgenic Res* 6:271–278. doi: [10.1023/A:1018406611380](https://doi.org/10.1023/A:1018406611380)
- Swindells MB (1995) A procedure for detecting structural domains in proteins. *Protein Sci* 4:103–112
- Tang B, Yu S, Zheng M, Ding F, Zhao R, Zhao J, Dai Y, Li N (2008) High level expression of a functional human/mouse chimeric anti-CD20 monoclonal antibody in milk of transgenic mice. *Transgenic Res* 17:727–732. doi: [10.1007/s11248-007-9162-3](https://doi.org/10.1007/s11248-007-9162-3)
- Toledo JR, Sanchez O, Segui RM, Garcia G, Montanez M, Zamora PA, Rodriguez MP, Cremata JA (2006) High expression level of recombinant human erythropoietin in the milk of non-transgenic goats. *J Biotechnol* 123:225–235. doi: [10.1016/j.jbiotec.2005.10.019](https://doi.org/10.1016/j.jbiotec.2005.10.019)
- van Berkel PH, Welling MM, Geerts M, van Veen HA, Ravensbergen B, Salaheddine M, Pauwels EK, Pieper F, Nuijens JH, Nibbering PH (2002) Large scale production of recombinant human lactoferrin in the milk of transgenic cows. *Nat Biotechnol* 20:484–487. doi: [10.1038/nbt0502-484](https://doi.org/10.1038/nbt0502-484)
- Van Cott KE, Lubon H, Gwazdauskas FC, Knight J, Drohan WN, Velander WH (2001) Recombinant human protein C expression in the milk of transgenic pigs and the effect on endogenous milk immunoglobulin and transferrin levels. *Transgenic Res* 10:43–51. doi: [10.1023/A:1008963817646](https://doi.org/10.1023/A:1008963817646)
- Vembar SS, Brodsky JL (2008) One step at a time: endoplasmic reticulum-associated degradation. *Nat Rev Mol Cell Biol* 9:944–957. doi: [10.1038/nrm2546](https://doi.org/10.1038/nrm2546)
- Wang J, Yang P, Tang B, Sun X, Zhang R, Guo C, Gong G, Liu Y, Li R, Zhang L, Dai Y, Li N (2008) Expression and characterization of bioactive recombinant human alpha-lactalbumin in the milk of transgenic cloned cows. *J Dairy Sci* 91:4466–4476. doi: [10.3168/jds.2008-1189](https://doi.org/10.3168/jds.2008-1189)
- Warrington JA, Nair A, Mahadevappa M, Tsyganskaya M (2000) Comparison of human adult and fetal expression and identification of 535 housekeeping/maintenance genes. *Physiol Genomics* 2:143–147
- Whitelaw CB, Archibald AL et al (1991) Targeting expression to the mammary gland: intronic sequences can enhance the efficiency of gene expression in transgenic mice. *Transgenic Res* 1(1):3–13. doi: [10.1007/BF02512991](https://doi.org/10.1007/BF02512991)

- Wright G, Carver A, Cottom D, Reeves D, Scott A, Simons P, Wilmut I, Garner I, Colman A (1991) A. High level expression of active human alpha-1-antitrypsin in the milk of transgenic sheep. *Biotechnology (N Y)* 9:830–834. doi:[10.1038/nbt0991-830](https://doi.org/10.1038/nbt0991-830)
- Yang P, Wang J, Gong G, Sun X, Zhang R, Du Z, Liu Y, Li R, Ding F, Tang B, Dai Y, Li N (2008) Cattle mammary bioreactor generated by a novel procedure of transgenic cloning for large-scale production of functional human lactoferrin. *PLoS ONE* 3:e3453. doi:[10.1371/journal.pone.0003453](https://doi.org/10.1371/journal.pone.0003453)
- Yu Z, Meng Q, Yu H, Fan B, Yu S, Fei J, Wang L, Dai Y, Li N (2006) Expression and bioactivity of recombinant human lysozyme in the milk of transgenic mice. *J Dairy Sci* 89: 2911–2918
- Zavadskaja ES, Zakharova ES, Kadulin SG, Kibardin AV, Kiselev SL, Gnuchev NV (2001) Production of recombinant endostatin in milk from transgenic mice. *Genetika* 37:1207–1212. doi:[10.1023/A:1011953312893](https://doi.org/10.1023/A:1011953312893)
- Zhang J, Li L, Cai Y, Xu X, Chen J, Wu Y, Yu H, Yu G, Liu S, Zhang A, Chen J, Cheng G (2008) Expression of active recombinant human lactoferrin in the milk of transgenic goats. *Protein Expr Purif* 57:127–135. doi:[10.1016/j.pep.2007.10.015](https://doi.org/10.1016/j.pep.2007.10.015)