Variability in immune-active human milk components

Mohèb Elwakiel

Propositions

1. Variation in human milk oligosaccharide composition among mothers makes it difficult to design a targeted infant formula enriched with these components for all babies.
   (this thesis)

2. The extent of human milk protein degradation *in vitro* is not directly affected by protease inhibitors.
   (this thesis)

3. PhD research should be based on a proof of concept.

4. Functional foods should be tested as drugs.

5. Secondments need to be made mandatory for doctorate students.

6. Although chemists and chefs are both using step-by-step guided recipes, a good chemist is not necessary a good chef.

Propositions belonging to the thesis, entitled
Variability in immune-active human milk components

Mohèb Elwakiel
Wageningen, 3 June 2020

Variability in immune-active human milk components

Mohèb Elwakiel

Thesis Committee

Promotors
Prof. Dr H.A. Schols
Personal Chair at the Laboratory of Food Chemistry
Wageningen University & Research

Dr K.A. Hettinga
Associate Professor, Food Quality and Design Group
Wageningen University & Research

Other members
Prof. Dr H.F.J. Savelkoul, Wageningen University & Research
Dr S.S. van Leeuwen, University Medical Center Groningen
Dr J.J. Mes, Wageningen University & Research
Dr B. Stahl, Utrecht University

This research was conducted under the auspices of the Graduate School VLAG (Advanced Studies in Food Technology, Agrobiotechnology, Nutrition and Health Sciences).

Variability in immune-active human milk components

Mohèb Elwakiel

Thesis

submitted in fulfilment of the requirements for the degree of doctor
at Wageningen University
by the authority of the Rector Magnificus
Prof. Dr A.P.J. Mol,
in the presence of the
Thesis Committee appointed by the Academic Board
to be defended in public
on Wednesday 3 June 2020
at 4 p.m. in the Aula

Mohèb Elwakiel

Variability in immune-active human milk components
168 Pages

PhD Thesis, Wageningen University, Wageningen, The Netherlands (2020)
With references, with summary in English

ISBN: 978-94-6395-344-3
DOI: https://doi.org/10.18174/517435

Table of Contents

Chapter 1

General introduction

Human milk is the best source of nutrition for infants

The World Health Organization advices mothers to exclusively breastfeed babies for the first 6 months of life[1]. In addition to its nutritional advantage, breastfeeding stimulates the infant's physical and mental state, improves maternal bonding, and significantly decrease the risk of having food allergies and osteoporosis[2-4]. Comparison studies between breast and formula-fed infants have shown that breastfed infants are less likely to develop infections, including bacterial meningitis, urinary tract infection, and diarrhoea[5]. Although human milk is superior to infant formula, only 41% of the breastfed infants exclusively receive milk for the entire first 6 months of life[6]. Among the world's regions, Europe has the lowest rates (circa 25%) for exclusive breastfeeding at month 6[7].

The composition of human milk cannot be duplicated easily. Human milk is a complex mixture of nutrients (e.g. lipids, lactose, proteins) and bioactive constituents, contributing to the infant's growth, development and health[8,9]. There are numerous immune modulating components in human milk, including polyunsaturated fatty acids, leucocytes, serum proteins such as lactoferrin and immunoglobulins, and components like hormones, growth factors, oligosaccharides, and peptides[10]. This PhD thesis will focus on serum proteins, serum protein *N*-glycans, and human milk oligosaccharides (HMOs), and addresses the enzymatic hydrolysis of human milk proteins in an *in vitro* infant (0–3 months) digestion model.

Composition of human colostrum and mature milk

Human milk is a complex food containing water, macronutrients and micronutrients, including lactose, lipids, HMOs, proteins, vitamins and minerals[9]. Lipids comprise the second largest component of human milk after lactose (Figure 1), providing infants with energy and essential vitamins, and bioactive components like triacylglycerides and phospolipids[10]. HMOs are the third largest part and exceed the protein content in human milk (Figure 1).

Human milk composition varies among mothers, between populations and over lactation[2,8]. The milk produced by mothers in the first two weeks after delivery is colostrum, which differs in volume, appearance and composition with milk from other timepoints further in lactation[9]. When the lactose concentration increases, the production of transition milk begins, which shares some of the characteristics of colostrum. Human milk is considered fully mature around the first month after birth[9]. The level of lactose is lower in colostrum (between 0–2 weeks), and increases in transition milk, and then remains constant in mature milk (> 4 weeks postpartum)[9]. The lipid concentrations in human stay constant over a 6 month lactation period[10]. In early life, infants have an immature intestinal immune system, making them more vulnerable to infection by opportunistic pathogens[10]. The higher levels of HMOs and serum proteins in colostrum and transition milk in comparison to mature milk might provide additional protection to the infant in this sensitive stage of its development.

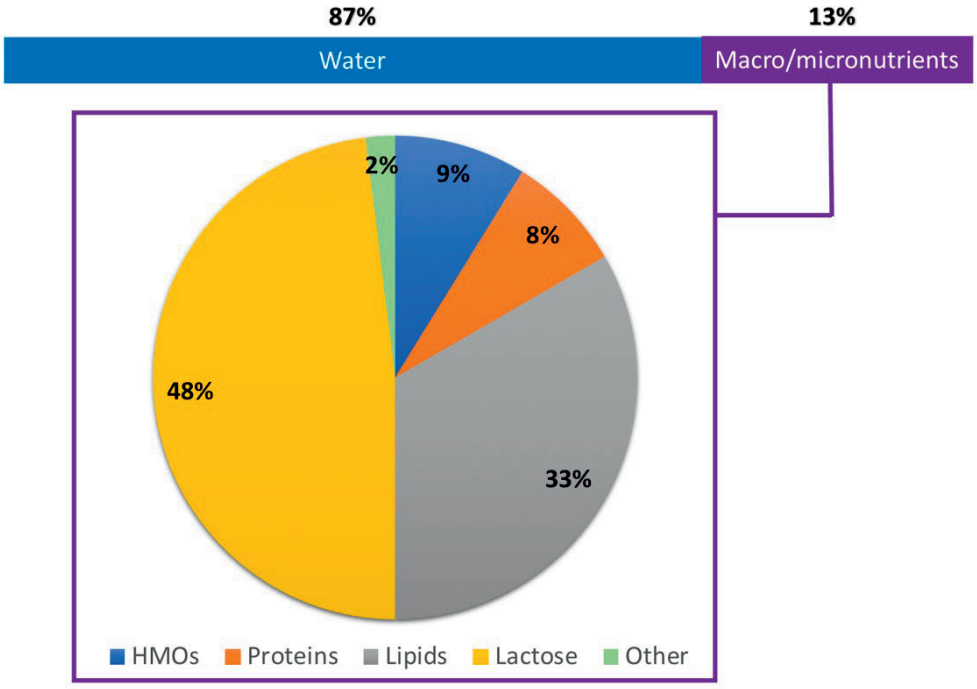


Figure 1. An approximate of the human milk composition of colostrum[3]. Major nutritional components of human milk are carbohydrates, lipids and proteins[10]. Other includes vitamins and minerals.

Human milk proteins

Human milk consists of 2 distinct types of proteins, serum proteins (e.g. α-lactalbumin, lactoferrin, immunoglobulins) and caseins (β-, α_{S1}- and κ-casein)[11]. The ratio between serum proteins and caseins in colostrum is circa 70:30, while in mature milk the ratio is circa 50:50. The total protein concentrations in colostrum range from 14 to 16 g/L[11], and in mature milk from 7 to 10 g/L. The most abundant proteins in Dutch human milk can be found in Table 1. The glycosylation status of these proteins will be further discussed later in this introduction. Caseins are generally described as transport proteins and many bioactive peptides are encrypted within their primary structure[12-14]. Caseins are relatively easy digested in peptides and amino acids by proteases and peptidases in the infant's digestive tract, which facilitates the uptake of relatively small peptides, minerals and essential amino acids[15,16].

Table 1. The 15 most abundant proteins in milk of Dutch mothers, as based on literature[11].

Function	Protein Name	Glycosylation status *
Enzyme	α-lactalbumin (P00709)	N (2)
	Bile salt-activated lipase (P19835)	N (1), O (10)
Immunity	Lactoferrin (P02788)	N (3), O (1)
	Immunoglobulins	
	Ig α_1-chain c-region (P01876)	N (2), O (5)
	Ig λ_2-chain c-region (P0DOY2)	X
	Ig κ-chain c-region (P01834)	X
	Polymeric immunoglobulin receptor (P01833)	N (7)
	Clusterin (P10909)	N (6)
	Osteopontin (P10451)	O (5)
	β_2-microglobulin (P61769)	X
Transport	β-casein (P05814)	X
	α_{S1}-casein (P47710)	X
	Serum albumin (P02768)	N (2)
	Fatty acid-binding protein (P05413)	X
	κ-casein (P07498)	O (8)

* Information on the glycosylation status of the proteins was gathered from the Uniprot database (https://www.uniprot.org, accessed September 2016). The number indicates the number of glycosites. N, known *N*-linked glycoprotein; O, known *O*-linked glycoprotein; X, no *N*- and *O*-linked glycans.

Next to the caseins, more than 200 serum proteins can be found in human milk. These serum proteins have a broad range of functionalities. The most abundant human milk serum protein α-lactalbumin is essential for the synthesis of lactose, supplies infants with large amounts of tryptophan, and facilitates the absorption of essential minerals when α-lactalbumin is digested[12]. Several other human milk serum proteins, like lactoferrin and immunoglobulins, protect infants against pathogens and decrease the risk of having acute or chronic diseases[15]. Lactoferrin, a globular glycoprotein of the transferrin family, partially ends up intact in the infant's faeces, and was shown to influence the microbiota composition of neonates[17]. Immunoglobulins are transferred via human milk from mother to child, and especially high quantities of secretory immunoglobulin A (sIgA) are present in human milk in the first days of lactation[16]. These immunoglobulins are also resistant to proteolysis and can therefore be partially found intact in the stool of infants[17]. Next to immunoglobulins, innate immune-active proteins, including complement proteins (complement factors 3, 4, 7, 9, B, and I), antibacterial proteins (e.g. lactoferrin, lactadherin and lysozyme) are highly abundant in colostrum[11]. The newborn infant is known to be deficient in CD14, which is part of the Toll-

like receptor complex. This Toll-like receptor complex can detect lipopolysaccharides on gram-negative bacteria and subsequently activate the innate immune system[11]. These innate immune-active proteins in colostrum may be needed against broad groups of pathogens in the infant's gastrointestinal tract[11]. The relatively high amount of both innate and adaptive immune-active proteins in colostrum suggests the importance of both the innate and adaptive immune system in the protection of the neonate.

Polymeric immunoglobulin receptor (PIGR) assists in the transportation of sIgA[18]. PIGR in human milk is mainly described as immune-active protein. Likewise, clusterin, β_2-microglobulin, and osteopontin are highly abundant in human milk and important for the development of the infant's immune system[19]. These latter proteins are also more abundant in colostrum than in mature milk[12]. Lysozyme also plays an important role in the mucosal immune system, and the concentration of this antibacterial protein is much higher in colostrum than in mature milk[18]. Serum albumin is transferred from the maternal blood circulation to milk, and is mainly involved in the transportation of a broad range of ligands (e.g. fatty acids, calcium, hormones)[13]. The digestion and uptake of lipids in infants is aided by bile salt-activated lipase[13], whereas fatty acid-binding protein is important for transporting fatty acids and other lipophilic substances like eicosanoids and retinoids[2]. The digestive enzyme, lipoprotein lipase hydolyzes triglycerides and is involved in promoting the cellular uptake of chylomicron remnants, colesterol-rich lipoproteins, and free fatty acids.

Overall, there are many serum proteins in human milk contributing to the infant's growth and development, with 15 proteins covering >90% of the total protein content[11]. In addition, transport, enzymes, and immune-active proteins are the most abundant in human milk[11]. The levels of the immune-active proteins generally decrease over lactation[11]. Given the various potential benefits of milk serum proteins, especially immune-active proteins, it would be of interest to obtain insight in the variability of serum proteins in milk from mothers from different geographical and ethnic origin, as individual differences in type and level of milk serum proteins between mothers and over time have been already reported[11].

The differences between human and bovine milk proteins have been already extensively studied[12,13]. The protein levels in bovine milk range between 30 and 35 g/L, and the casein content represents 80% of the bovine milk proteins over lactation[12]. In contrast to human milk, β-lactoglobulin is the most abundant serum protein in bovine milk, followed by α-lactalbumin and serum albumin[12]. The proteins, α_{S2}-casein and β-lactoglobulin can only be found in bovine milk[13]. The concentration of lactoferrin and lysozyme is much lower in bovine than in human milk[11]. Bovine colostrum was found to contain similar amounts of immunoglobulins as human colostrum, but their concentrations in bovine milk declines faster after the first days of lactation than human milk[12]. The differences in the dominant digestive enzymes found in human milk (bile salt-activated lipase) and bovine milk (pancreatic ribonuclease) may reflect the differences in the developmental needs of the

digestion system between human and cows[12]. Pancreatic ribonuclease plays a major role in the digestion of nucleic acids of microorganisms in the rumen of calves and thereby promote the nutrient uptake in the intestine[12]. Lipoprotein lipase was found in both human and bovine milk, which is important for nutrient digestion and adsorption[12].

The digestion of proteins in human milk and in the infant's gastrointestinal tract

Some of the major proteins (caseins, osteopontin, clusterin, PIGR) are predigested by proteases, resulting in the existence of peptides in human milk[14]. Some endogenous peptides have bioactivities in which they e.g. prevent infections and pathogen colonisation[14]. Most milk proteins are degraded into peptides and amino acids in the digestive tract[20]. Some of the peptides formed during digestion are known for their bio-active/immune-active properties[21,22]. Besides peptides of β-casein, studies have shown that lactoferrin contains unique antibacterial peptides, such as lactoferricin and lactoferrampin[21,22].

Around 3-10% of the proteins in human milk are not digested in the gastrointestinal tract, resulting in the visibility of proteins in the stool of the newborn[17]. The mechanism how this happens is still unclear. Protease inhibitors are present in human milk, but present studies do not show what functions they exert during infant digestion[11,23]. This is where a gap in knowledge exists which leads us to the question, whether protease inhibitors can reduce the extent of protein digestion in the infant's gastrointestinal tract. A possible mechanism behind the survival of colostral human milk serum proteins can be found in Figure 2.

Protease inhibitors in human milk and their possible role against proteolytic enzymes in the gastrointestinal tract

Human milk contains a wide range of proteases, protease inhibitors, protease activators, and zymogens[23,24]. These constituents together determine the proteolytic activity of human milk[23]. Several enzymes activated by the proteolytic system are plasmin, trypsin, thrombin, amino-, and carboxypeptidase, cathepsin, elastase, and kalikrein[23]. The most abundant protease inhibitors found in human milk are α_1-antichymotrypsin, α_1-antitrypsin, cystatin C[11]. Milk also contains antithrombin III, but this protease inhibitor is only found in trace amounts[11]. Antithrombin III protects the proteins from proteolysis caused by thrombin. Thrombin is not only a protease, but also an activator of trypsin[11]. By inhibition of thrombin, performed by antithrombin III and α_1-antichymotrypsin, the activation of trypsin is declined[23]. The protease inhibitor α_1-antitrypsin inhibits multiple proteases (elastase, chymotrypsin, trypsin, plasmin and thrombin)[21].

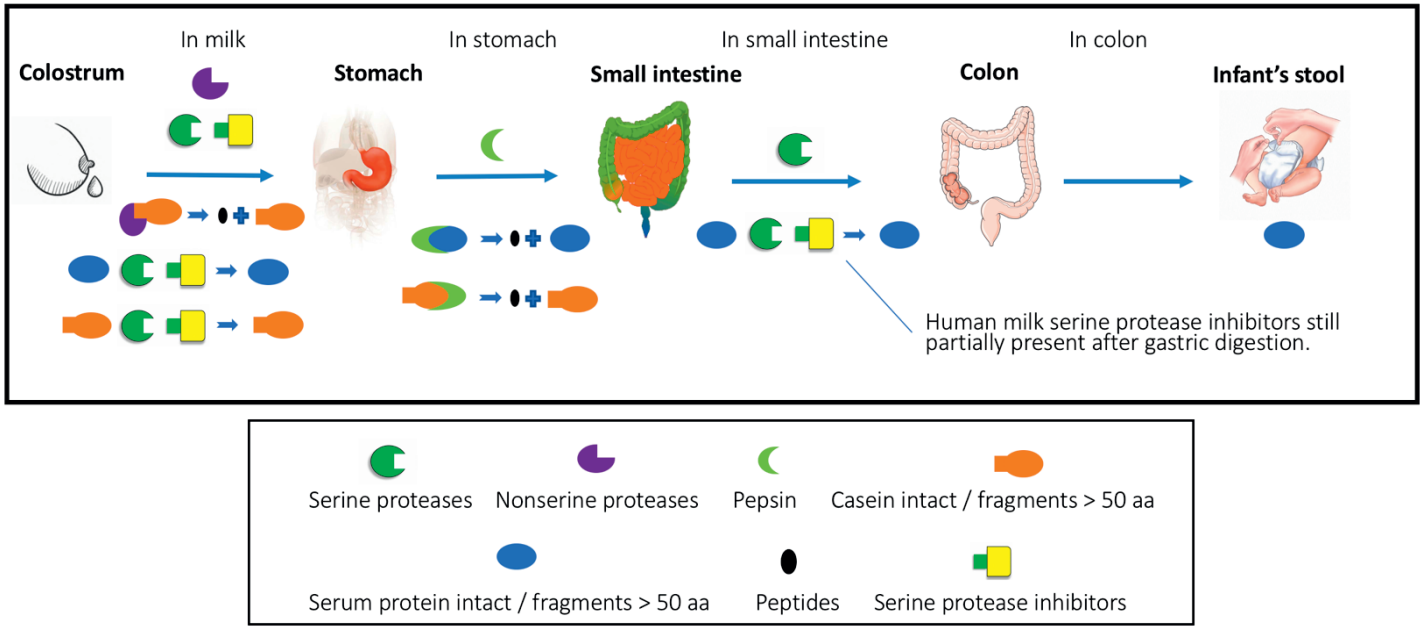


Figure 2. Possible mechanism for the survival of colostral serum proteins in human milk and the infant's digestive tract, as based on literature[12]. Amino acids = aa. The mechanism and products formed after digestion are highlighted for each digestion phase.

The concentrations of α_1-antichymotrypsin (0.4–0.7 g/L) and α_1-antitrypsin (0.1–0.4 g/L) are higher in colostrum than in mature milk[23]. These latter two protease inhibitors are not completely digested by proteases, and may escape from fermentation in the colon as they may even end up in the infant's faeces[17].

Protease inhibitors are involved in both the innate and adaptive immune system of the newborn[21]. In addition, a positive correlation in human milk between the level of protease inhibitors and serum proteins content involved in the innate and adaptive immune system has been reported[11]. This might imply that protease inhibitors have a protective role against the degradation of these immune-active serum proteins in human milk[11]. It has also been suggested that α_1-antichymotrypsin and α_1-antitrypsin, especially in colostrum, play an important role in the protection against enzymatic protein hydrolysis by serine proteases in human milk (e.g. trypsin) and might limit the activity of pancreatic enzymes in the infant's small intestine (e.g. trypsin and chymotrypsin)[23,24], possibly causing proteins to remain intact and therefore bioactive in the small intestine. Especially during the first weeks after birth when the concentration of protease inhibitors is higher in colostrum than in mature milk (Figure 2), there may be a reduced protein digestion.

Extending an adult *in vitro* digestion model to infant populations

Studies on the role of protease inhibitors in resistance of immune-active serum proteins against digestion can be performed using an infant *in vitro* digestion model. The infant's digestive tract is not fully developed after birth[24-28]. The concentrations and activities of amylase, pepsin, chymotrypsin, and trypsin are typically lower for 3-month-old infants than in older infants and adults[24,26,]. The gastrointestinal tract of young infants might result in a reduced protein digestion compared to older infants and adults.

Several child *in vitro* methods have been developed over the years[20,21,29]. The majority of these child *in vitro* digestion models are based on the INFOGEST adult *in vitro* digestion model[30]. This INFOGEST model makes use of 3 phases, namely the oral, gastric and intestinal phase. The simulated salivary, gastric, and intestinal fluid were based on human *in vivo* data retrieved from the mouth, stomach and small intestine. These fluids were investigated for their salt composition, pepsin activity, trypsin activity and the bile salt concentration.

In an infant (0–3 months) *in vitro* model, the oral phase can be excluded[31]. The next phase in the digestion model is the gastric phase. The gastric pH is higher in infants compared to adults[27]. Research showed that infants, in the first three months, had an average pH of 5 in the gastric phase[27]. The activity of pepsin for young infants is only 18% at four weeks after parturition, as compared to the full activity of the enzyme in older infants[24]. Another important parameter of the infant's stomach is the incubation time or gastric emptying time. It was found that 90% of the human milk was transferred from the stomach within one hour[31]. In previous studies it was found that the best incubation time to mimic the infant's

digestion was one hour[31]. Not only the transit time was investigated, but also the enzyme release in the small intestine and the bile salt concentrations[31]. Investigating the data obtained from infant *in vivo* experiments[26,31], the enzyme concentrations in the duodenal mixture is a twelvefold lower and the bile salt concentrations six times lower. Such a model has not yet been developed for a 0 to 3-month-old infants.

HMOs in human milk

The carbohydrate fraction in human milk consists of lactose and HMOs[32]. Lactose is a disaccharide made up of glucose and galactose joined by a β1,4 glycosidic linkage. HMOs are the most abundant solid component in human milk, after lactose and lipids[32]. HMO concentrations generally range between 20-23 g/L in colostrum, and 7-12 g/L in mature milk[32].

HMOs are complex lactose-based glycans synthesized in the mammary gland during lactation[33]. HMOs are composed of five monosaccharides; glucose (Glc), galactose (Gal), *N*-acetylglucosamine (GlcNAc), fucose (Fuc), and *N*-acetylneuraminic acid (NeuAc)[33]. During the synthesis of HMOs, lactose can be elongated with a GlcNAc residue through β1,3 or β1,6 linkages, and these core HMO structures can be further decorated with Gal, GlcNAc, Fuc or sialic acid NeuAc residues[33]. With these five monosaccharides a large number of structures can be formed, which are either branched or linear. A schematic representation of the main HMOs in human milk can be found in Figure 3.

HMOs have a remarkable structural and functional diversity, with >150 structures identified in human milk[32], while typically 15 of these structures account for >90% of the total HMO content[34]. HMOs can be classified as neutral or acidic HMOs, with acidic oligosaccharides generally being present at a 10-fold lower concentration than neutral oligosaccharides[33]. Fucose decorates up to 70% of all the oligosaccharides in human milk. The biosynthesis of the neutral HMOs includes the enzyme *N*-acetylglucosaminyltransferases, which is able to attach GlcNAc in a β1,3 or β1,6 linkage to a terminal Gal residue[33]. Galactosyltransferases attach a Gal residue via a β1,3 or β1,4 linkage to the GlcNAc residues[33]. The synthesis of sialylated HMOs requires sialyltransferases, which produce sialic acid containing HMOs with either α2,3 or α2,6 linkages[33].

HMO fucosylation depends on the presence of specific fucosyltransferases (FUT) enzymes, which are genetically determined by the mother's secretor (Se) and Lewis (Le) histo-blood group (Figure 4)[34,35]. The FUT2 Se gene determines the presence of α1,2-fucosylated oligosaccharides in human milk. On the basis of the Lewis (Le) blood group system, the FUT3 Le gene determines the presence of α1,4-fucosylated oligosaccharides in human milk.

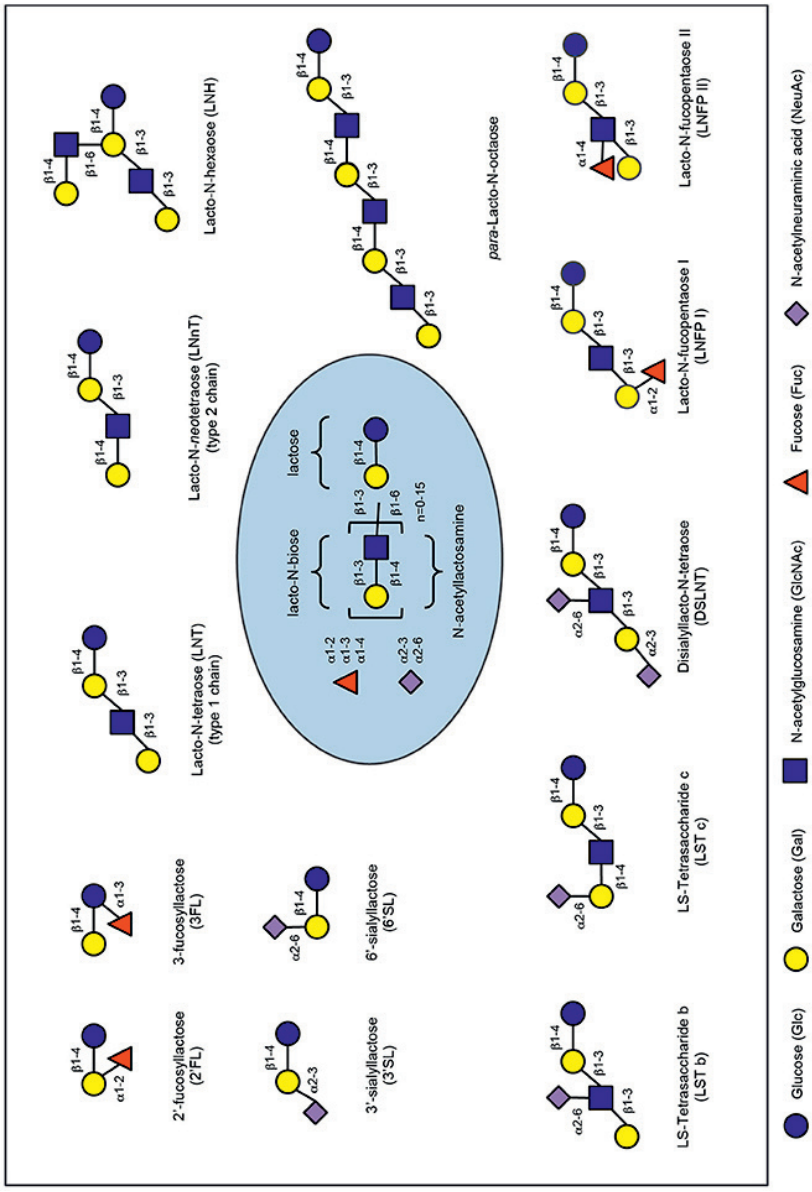


Figure 3. Schematic representation of HMOs in human milk[33].

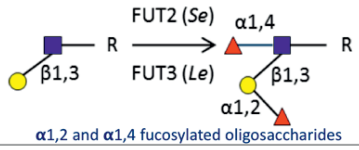
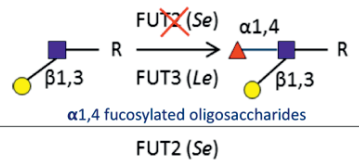
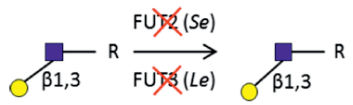
Milk-types	Type of linkage (Fucosylation)	% Worldwide
Group 1: (Se⁺Le⁺)		70%
Group 2: (Se⁻Le⁺)		20%
Group 3: (Se⁺Le⁻)		10%
Group 4: (Se⁻Le⁻)		<1%

Figure 4. The synthesis of HMOs in human milk depend on the mother's SeLe blood group status, as based on literature[34]. Secretor Lewis positive, Se^+Le^+; Nonsecretor Lewis positive, Se^-Le^+; Secretor Lewis negative, Se^+Le^-; Nonsecretor Lewis negative, Se^-Le^-. Human milk can be classified according to the presence and absence of α1,3/4- and α1,2-fucosylated oligosaccharides, which indicate the activity of the FUT3 and FUT2 gene of the mother.

Women with an active Se locus are classified as secretor (Se^+), whereas women with an active Le locus are classified as Lewis positive (Le^+). Women without FUT2 or FUT3 activity are classified as nonsecretor (Se^-) or Lewis negative (Le^-), lacking α1,2-fucosylated or α1,4-fucosylated oligosaccharides, respectively. Figure 4 shows that circa 70% of the women worldwide can be classified as Se^+Le^+ mother, containing in their milk α1,2-fucosylated, α1,3-fucosylated and α1,4-fucosylated HMOs. Se^-Le^+ mothers (20%) produce milk containing only α1,3-fucosylated and α1,4-fucosylated HMOs, due to the lack of the FUT2 gene. The Se^+Le^- mothers (10%) produce milk without α1,4-fucosylated HMOs. The Se^-Le^- milk-type is found in rare cases (<1%), with the α1,3-fucosylated HMOs being the only type of fucosylated HMOs. The variability in type and levels of HMOs between and within the 4 SeLe groups are not yet investigated in depth.

Having many biological functions, HMOs are able to prevent pathogens from binding to epithelial cell surfaces, and to alter host epithelial and immune cell responses[35-37]. HMOs

may bind directly to bacteria in the gut lumen, causing conformational change in binding sites, preventing binding to cell receptors (Figure 5). Additonally, HMOs may bind directly to gut epithehial cells causing altered availability of cell receptors, which may prevent pathogen binding to the gut epithelial cells (Figure 5). HMOs also offer natural protection against necrotizing enterocolitis, and specific HMOs containing sialic acid residues provide nutrients for brain development[35-37]. In addition, HMOs are not digested in the small intestine and these HMOs might serve as substrates for most beneficial microbes, contributing to the shaping of infant's gut flora[38-40]. It has been reported that, e.g. Bifidobacteria are equiped with enzymes capable of breaking down HMOs. These Bifidobacteria contribute up to 90% of the microbial community in the gut of breastfed infants[38]. Different microbial species and strains in the infant's digestive tract might have their own mechanisms for the degradation of HMOs leading to potentially diverse metabolisation products in the feces[41-42]. In addition, intact HMOs have been discovered in the feces of breastfed infants[40]. Overall, HMOs are complex indigestible carbohydrates, which are fermented in the large intestine, and important for the healthy development of the newborn's microbiota. HMOs are present in large amounts in human milk, however, these structures are not yet present in infant formula[10]. The most common prebiotic supplementation in bovine mik and infant formula is galacto- and fructo-oligosaccharides. Despite these latter oligosaccharides having simplified structures compared to HMOs, they have been recognized for their ''bifidogenic'' or prebiotic effects[38].

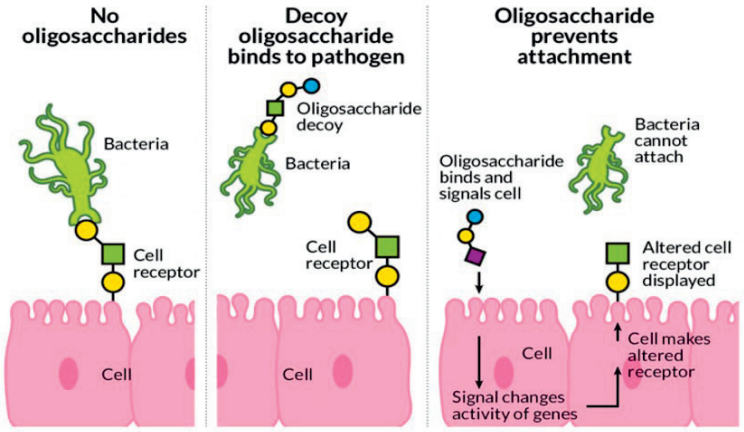


Figure 5. Mechanism of action of HMOs to prevent abberant pathogen colonization[37].

Glycoproteins containing *O*- and *N*-glycans

Glycans can be covalently attached to proteins and lipids to form glycoconjugates[43]. The majority of the glycolipids in human milk are present in the milk fat globule membrane, while most of the glycoproteins can be found in the milk serum[44]. The glycosylation of proteins occurs in the endoplasmic reticulum and Golgi apparatus of the mammary gland via various glycosyltransferases[44]. It has been shown that circa 70% of the human milk proteins are glycosylated to some extent[44].

It has been reported that α_{S1}- and β-casein in human milk do not have any glycosylated amino acid residues (Table 1), while κ-casein contains multiple *O*-glycosylation sites at various threonine (Thr) and serine (Ser) residues[45]. No specific sequence of amino acids is needed for *O*-glycosylation[45]. Caseinmacropeptide is released from κ-casein during digestion in the infant's stomach[14]. The sialic acid residues attached to the *O*-glycans in this caseinmacropeptide might be important for the infant's brain development. Serum proteins such as lactoferrin, immunoglobulins, serum albumin, and α-lactalbumin form the main portion of the glycoproteins present in human milk (Table 1), and mainly contain *N*-glycans. These glycans are attached to the amide nitrogen of an asparagine (Asn) residue of the protein[46,47]. The Asn belong to a specific sequence of amino acids, Asn-X-Ser-Thr. X can be any arbitrary amino acid, with the exception of proline[48]. The *N*-glycans attached to the immunoglobulins, serum albumin, and lactoferrin might protect these glycoproteins from digesting by proteolytic enzymes during digestion, and help these glycoproteins to arrive in the intestine[49].

A schematic representation of *O*- and *N*-glycotails attached to proteins can be found in Figure 6. More than 75% of the serum proteins in number are *N*-glycosylated, based on the most abundant human milk proteins (Table 1). *N*-glycans are composed of six monosaccharides; Fuc, Gal, mannose (Man), GlcNAc, *N*-acetylgalactosamine (GalNAc), and the sialic acid NeuAc structure[46,47]. The sialic acid residue, *N*-glycosylneuramimic acid can be only found in bovine milk[46]. The *N*-glycans in bovine and human milk have all an essential pentasaccharide core consisting of three Man residues and two GlcNAc residues[46]. The two GlcNAc residues are linked via a β1,4 linkage, whereas the GlcNAc dimer is β1,4-linked to a Man residue. The Man residue is connected with two other Man residues, through a α1,3- and α1,6-linkage.

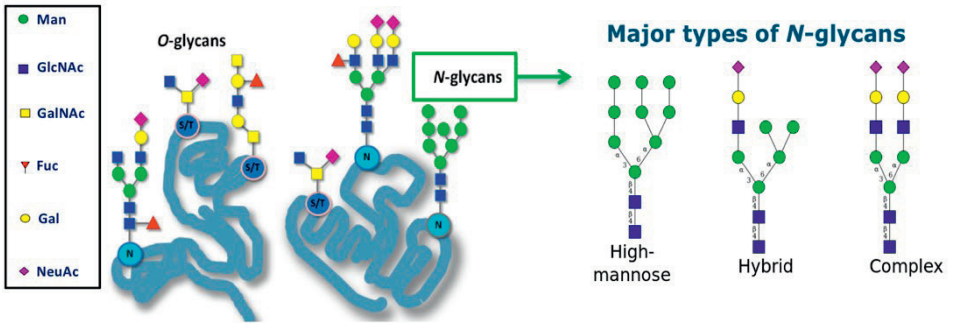


Figure 6. Schematic representation of *O*- and *N*-glycotails attached to proteins[44]. The essential core of *N*-glycans consists of three Man and two GlcNac residues.

The structure of *N*-glycans can be classified into three types, namely the high-mannose type, complex type, and hybrid type[46]. High mannose *N*-glycans merely consists of Man residues. Hybrid *N*-glycans contain a unsubstituted terminal Man residue (as are present in high mannose *N*-glycans) and substituted Man residue linked to a GlcNAc residue (as are present in complex *N*-glycans). These GlcNAc residue added to the *N*-glycan core in hybrid and complex *N*-glycans are part of the antennae (Figure 6). The biantennary *N*-glycan consists of two GlcNAc branches linked to the core, while triantennary *N*-glycan consists of three GlcNAc branches. Complex *N*-glycans differ from high mannose and hybrid *N*-glycans by having added GlcNAc residues at both α1,3- and α1,6 Man sites. Unlike the high mannose structures, complex *N*-glycans do not contain Man residues apart from the core structure. Both the hybrid and complex *N*-glycans are often decorated by Fuc and NeuAc residues[45-47]. Based on these building blocks, *N*-glycans can also be classified as acidic and neutral *N*-glycans, which can be further divided in nonfucosylated and fucosylated structures. The core fucosylation level (α1,6-linkages) of the complex and hybrid *N*-glycans is associated with the FUT8 enzyme[48]. Fucosylation of *N*-glycans might partially depend on the mother's SeLe histo-blood group, like with HMOs, which results in the presence and absence of α1,3/4- and α1,2-fucosylated oligosaccharides. So far, none of the studies have investigated the possible link of mother's SeLe status to the fucosylation of *N*-glycans in human milk.

The characterization of *N*-glycans from serum proteins in human milk might lead to a better understanding of the structural and functional properties of serum proteins. Several glycosylated serum proteins (e.g. lactoferrin, immunoglobulins, lysozyme, α_1-antichymotrypsin) can be partially found in the infant's stool, and they may preserve their bioactivity during digestion to protect the infant's small intestine[17]. These *N*-glycans might influence the stability of these proteins[49]. So far, no information is available on the composition of *N*-glycans in milk among mothers and over lactation.

Thesis outline

The work in this PhD thesis provides advanced knowledge on the variability in type and levels of HMOs and serum proteins in milk of Chinese and Dutch mothers during lactation. In addition, it gives insight in the variability of serum protein *N*-glycans in milk of Chinese mothers over time, and elaborates on the enzymatic digestion of proteins from colostrum and mature milk of Chinese mothers in an *in vitro* infant (0–3 months) digestion model.

In Figure 6 an overview of the PhD thesis is given including the used analytical techniques highlighted in orange. **Chapter 2** provides insights in the variability in type and levels of serum proteins in milk from Chinese and Dutch mothers over a 20-week lactation period. **Chapter 3** describes the disappearance of intact proteins in an *in vitro* infant (0–3 months) digestion model, and the potential role of protein content and protease inhibitors. Large peptides and undigested proteins remaining after digestion (**Chapter 3**) were analysed using sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) in combination with liquid chromatography mass spectrometry (LC-MS/MS). **Chapter 4** describes a developed method to release *N*-linked glycotails from serum proteins, and compared the serum protein *N*-glycans in colostrum (week 1) and mature milk (week 4) of Chinese mothers, using matrix assisted laser desorption ionization time of flight mass spectrometry (MALDI-TOF-MS). **Chapter 5** provides insights in the variability of the lactose and HMO concentrations in milk of Chinese and Dutch mothers over a 20-week lactation period, using capillary electrophoresis laser-induced fluorescence (CE-LIF). **Chapter 6** discusses all the findings and provides information on the peptide composition in human milk from 2 lactation periods, and in digesta from *in vitro* infant (0–3 months) digestion.

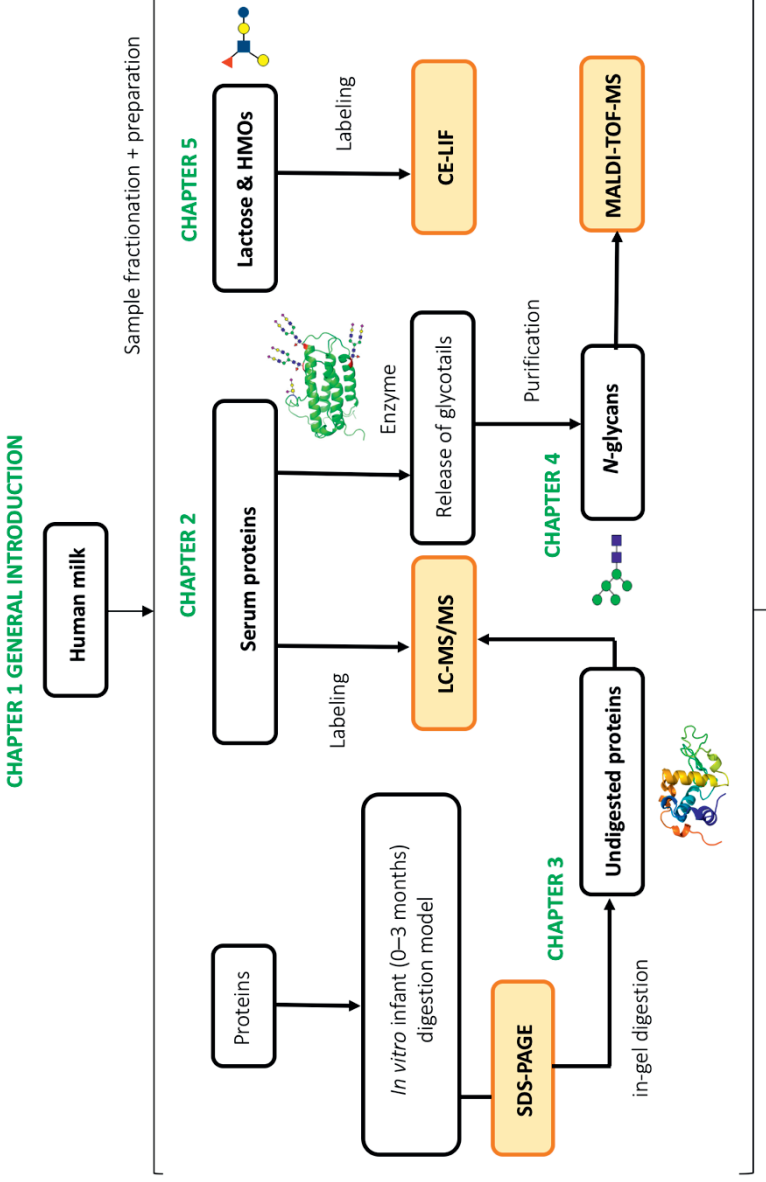


Figure 6. Schematic overview of the thesis outline: Proteomics and Glycomics of human milk, and the digestion of human milk proteins in an *in vitro* infant model.

References

1. Kramer, M.; Kakuma, R. The optimal duration of exclusive breastfeeding: A systematic review. Adv. Exp. Med. Bio. **2004**, 554, 63–77.
2. Gartner, L.; Morton, J.; Lawrence, R.; Naylor, A.; O'Hare, D.; Schanler, R.; Eidelman, A. Breastfeeding and the use of human milk. Pediatr. **2005**, 115, 496–506.
3. Field, C. The immunological components of human milk and their effect on immune development in infants. J. Nutr. **2005**, 135, 1–4.
4. Hanson, L.; Korotkova, M.; Lundin, S.; Haversen, L.; Silfverdal, S.; Mattsby-Baltzer, I.; Strandvik, B.; Telemo, E. The transfer of immunity from mother to child. Ann. N.Y. Acad. Sci. **2003**, 987, 199–206.
5. Duijts, L.; Jaddoe, V.; Hofman, A.; Moll, H. Prolonged and exclusive breastfeeding reduces the risk of infectious diseases in infancy. Pediatr. **2010**, 126, 18–25.
6. Global UNICEF Databases. **2018**, https://data.unicef.org/topic/nutrition/infant-and-young-child-feeding/
7. Theurich, M.; Davanzo, R.; Rasmussen, M.; Brennan, C.; McHugh, L.; Weikert, C.; Berthold, K, Breastfeeding rates and programs in Europe: A survey of 11 national breastfeeding committees and representatives. J. Pediatr. Gastroenterol. Nutr. **2019**, 68, 400–7.
8. Leung, A.; Sauve, R. Breast is best for babies. J. Natl. Med. Assoc. **2005,** 97, 1010–9.
9. Ballard, O.; Morrow, A. Human milk composition: Nutrients and bioactive factors. Pediatr. Clin. North Am. **2013**, 60, 49–74.
10. Labbok, M.; Clark, D.; Goldman, A. Breastfeeding: Maintaining an irreplaceable immunological resource. Nat. Rev. Immunol. **2004**, 4, 565–72.
11. Zhang, L.; de Waard, M.; Verheijen, H.; Boeren, S.; Hageman, J.A.; van Hooijdonk, T.; Vervoort, J.; van Goudoever, J.B.; Hettinga, K.A. Changes over lactation in breast milk serum proteins involved in the maturation of immune and digestive system of the infant. J. Proteomics, **2016**, 40–7.
12. Zhang, L.; van Dijk, A.; Hettinga, K.A. An interactomics overview of the human and bovine proteome over lactation. Proteome Sci. **2016**, 15, 1–14.
13. Hettinga, K.A.; Van Valenberg, H.; De Vries, S.; Boeren, S.; Van Hooijdonk, T.; Van Arendonk, J.; Vervoort, J. The host defense proteome of human and bovine milk. PLoS One, **2011**, 6, 1–8.
14. Dallas, D.; Guerrero, A.; Khaldi, N.; Borghese, R.; Bhandari, A.; Underwood, M.; Lebrilla, C.; German, J.; Barile, D. A peptidomic analysis of human milk digestion in the infant stomach reveals protein-specific degradation patterns. J. Nutr. **2014**, 144, 815–20.
15. Lönnerdal, B. Nutritional and physiologic significance of human milk proteins. Am. J. Clin. Nutr. **2003**, 6, 1537–43.
16. Kunz, C.; Rodriguez-Palmero, M.; Koletzko, B.; Jensen, R. Nutritional and biochemical properties of human milk, part I: General aspects, proteins, and carbohydrates. Clin. Perinatol. **1999**, 26, 307–33.
17. Davidson, L.; Lönnerdal B. Persistence of human milk proteins in the breastfed infant. Acta Pædiatr. **1987**, 76, 733–40.
18. Lönnerdal, B. Bioactive proteins in human milk: Mechanisms of action. J. Pediatr. **2010**, 156, 26–30.

19. Lönnerdal, B. Human milk proteins: Key components for the biological activity of human milk. Adv. Exp. Med. Biol. **2007**, 11–25.
20. Chatterton, D.; Rasmussen, J.; Heegaard, C.; Sorensen, E.; Petersen, T. *In vitro* digestion of novel milk protein ingredients for use in infant formulas: Research on biological functions. Trends Food Sci. **2004**, 15, 373–83.
21. Dall'Asta, C.; Florio, P.; Lammardo, A.; Prandi, B.; Mazzeo, T.; Budelli, A.; Pellegrini, N. Development of an *in vitro* digestive model for studying the peptide profile of breast milk. Int. J. Food Sci. Nutr. **2015**, 66, 409–15.
22. Hernandez-Ledesma, B.; Quiros, A.; Amigo, L.; Recio, I. Identification of bioactive peptides after digestion of human milk and infant formula with pepsin and pancreatin. Int. Dairy J. **2007**, 17, 42–9.
23. Chowanadisai, W.; Lönnerdal, B. α_1-Antitrypsin and antichymotrypsin in human milk: Origin, concentrations, and stability. Am. J. Clin. Nutr. **2002**, 76, 828–33.
24. Dallas, D.; Murray, N.; Gan, J. Proteolytic systems in milk: Perspectives on the evolutionary function within the mammary gland and the infant. J. Mammary Gland Biol. Neoplasia, **2015**, 20, 133–47.
25. Feher, J. Digestion and absorption of the macronutrients. In: Quantitative human Psysiology. Academic Press: Cambridge, U.S., **2012**, pp. 731–43. (Ed.)
26. Dupont, D.; Mandalari, G.; Molle, D.; Jardin, J.; Leonil, J.; Faulks, R.; Wickham, M.; Mills, E.; Mackie, R. Comparative resistance of food proteins to adult and infant *in vitro* digestion models. Mol. Nutr. Food Res. **2010**, 54, 767–80.
27. Dipalma, J.; Kirk, C.; Hamosh, M.; Colon, A.; Benjamin, S.; Hamosh, P. Lipase and pepsin activity in the gastric mucosa of infants, children, and adults. Gastroenterol. **1991**, 101, 116–21.
28. Hur, S.; Ou Lim, B.; Decker, E.; McClements, J. *In vitro* human digestion models for food applications. Food Chem. **2011**, 125, 1–12.
29. Liu, F.; Teodorowicz, M.; Wichers, H. J.; Van Boekel, M. A.; Hettinga, K. A.; Generation of Soluble Advanced Glycation End Products Receptor (sRAGE)-Binding Ligands during Extensive Heat Treatment of Whey Protein/Lactose Mixtures Is Dependent on Glycation and Aggregation. J. Agric. Food Chem. **2016**, 64, 6477–86.
30. Minekus, M; Alminger, M.; Alvito, P.; Balance, S.; Bohn, T.; Bourlieu, C.; Carriere, F.; Boutrou, R.; Corredig, M.; Dupont, D.; Dufour, C.; Egger, L.; Golding, M.; Karakaya, S.; Kirkhus, B.; Le Feunteun, S.; Lesmes, U.; Macierzanka, A.; Mackie, A.; Marze, S.; McClements, D.; Menard, O.; Recio, I.; Santos, C.; Singh, R.; Vegarud, G.; Wickham, M.; Weitschies, W.; Brodkorb, A. A standardized static *in vitro* digestion method suitable for food – An international consensus. Food Funct. **2014**, 5, 1113–24.
31. Bourlieu, C.; Menard, O.; Bouzerzour, K.; Mandalari, G.; Macierzanka, A.; Mackie, A.; Dupont, D. Specificity of infant digestive conditions: Some clues for developing relevant *in vitro* models. Crit. Rev. Food Sci. Nutr. **2014**, 54, 1427–57.
32. Kunz, C.; Rudloff, S.; Baier, W.; Klein, N.; Strobel S. Oligosaccharides in human milk: Structural, functional, and metabolic aspects. Annu. Rev. Nutr. **2000**, 20, 699–722.
33. Bode, L. Human milk oligosaccharides: Every baby needs a sugar mama. Glycobiology, **2012**, 22, 1147–62.

34. Thurl, S.; Henker, J.; Siegel, M.; Tovar, K.; Sawatzki, G. Detection of four human milk groups with respect to Lewis blood group dependent oligosaccharides. Glycoconj. J. **1997**, 14, 795–9.
35. Bode, L.; Jantscher-Krenn, E. Structure-function relationships of human milk oligosaccharides. Adv. Nutr. **2012**, 3, 383–91.
36. Kunz, C.; Rudloff, S. Health promoting aspects of milk oligosaccharides. Int. Dairy J. **2006**, 16, 1341–6.
37. Morrow, A.; Ruiz-Palacios, G.; Altaye, M.; Jiang, X.; Guerrero, M.; Meinzen-Derr, J.; Farkas, T.; Chaturvedi, P.; Pickering, L.; Newburg, D. Human milk oligosaccharides are associated with protection against diarrhea in breastfed infants. J. Pediatr. **2004**, 145, 297–303.
38. Davis, J.; Totten, S.; Smilowitz, J.; German, B.; Mills, D.; Lebrilla, C. Identification of oligosaccharides in faeces of breastfed infants and their correlation with the gut microbial community. MCP, **2016**, 15, 2987–3002.
39. LoCascio, R.; Ninonuevo, M.; Freeman, S.; Sela, D.; Grimm, R.; Lebrilla, C.; Mills, D.; German, B. Glycoprofiling of Bifidobacterial consumption of human milk oligosaccharides demonstrates strain specific, preferential consumption of small chain glycans secreted in early human lactation. J. Agric. Food Chem. **2007**, 55, 8914–9.
40. Albrecht, S.; Schols, H.A.; van den Heuvel, E.G.; Voragen, A.G.; Gruppen, H. Occurrence of oligosaccharides in faeces of breastfed babies in their first six months of life and the corresponding breast milk. Carbohydr. Res. **2011**, 346, 2540–50.
41. Sela, D.; Mills, D.; Nursing our microbiota: Molecular linkages between *bifidobacteria* and milk oligosaccharides. Trends Microbiol. **2010**, 18, 298–307.
42. Marcobal, A.; Barboza, M.; Froehlich, J.; Block, D.; German, B.; Lebrilla, C.; Mills, D. Consumption of human milk oligosaccharides by gut-related microbes. J. Agric. Food Chem. **2010**, 58, 5334–40.
43. Magnelli, P.; Bielik, A.; Gutrie, E. Identification and characterization of protein glycosylation using specific endo- and exoglycosidases. J. Vis. Exp. **2011**, 3749–54.
44. Froehlich, J.; Dodds, E.; Barboza, M.; McJimpsey, E.; Seipert, R.; Francis, J.; An, H.; Freeman, S.; German, J; Lebrilla, C. Glycoprotein expression in human milk during lactation. J. Agric. Food Chem. **2015**, 58, 6440–8.
45. Newburg, D.; Grave, G. Recent advances in human milk glycobiology. Pediatr. Res. **2014**, 75, 675–9.
46. Nwosu, C.; Aldredge, D.; Lerno, L.; Zivkovic, A.; German, B.; Lebrilla, C. Comparison of the human and bovine milk *N*-glycome via high-performance microfluidic chip liquid chromatography and tandem mass spectrometry. J. Proteome Res. **2012**, 11, 2912–24.
47. Dallas, D.; Martin, W.; Strum, J.; Zivkovic, A.; Smilowitz, J.; Underwood, M.; Affolter, M.; Lebrilla, C.; German, B. *N*-linked glycan profiling of mature human milk by high performance microfluidic chip liquid chromatography time of flight tandem mass spectrometry. J. Agric. Food Chem. **2011**, 59, 4255–63.
48. Vasconcelos-Dos-Santos, A.; Oliveira, I.; Lucena, M. Biosynthetic machinery involved in aberrant glycosylation: Promising targets for developing of drugs against cancer. Front Oncol. **2015**, 5, 1–23.
49. Georgi, G.; Bartke, N.; Wiens, F.; Stahl, B. Functional glycans and glycoconjugates in human milk. Am. J. Clin. Nutr. **2013**, 98, 578–85.

Chapter 2

Variability of serum proteins in Chinese and Dutch human milk during lactation

Published as: Elwakiel, M.; Boeren, S.; Hageman, J.A.; Szeto, I.M.; Schols, H.A; Hettinga, K.A. Nutrients, **2019**, 11, 499–513.

Abstract

To better understand the variability of type and level of serum proteins in human milk, the milk serum proteome of Chinese mothers over lactation was investigated using proteomic techniques and compared to the milk serum proteome of Dutch mothers. This study showed that total milk serum protein concentrations in Chinese human milk decreased over a 20-week lactation period, although with variation between mothers in rate of decrease. Variation was also found in the composition of serum proteins in both colostrum and mature milk, although the group of immune-active proteins, enzymes, and transport proteins were the most abundant for all mothers. These three protein groups encompass many of the 15 most abundant proteins, covering >95% of the total protein concentrations, in both the Chinese and Dutch milk serum proteome. The Dutch and Chinese milk serum proteome were also compared based on 166 common milk serum proteins, which showed that 22% of the 166 serum proteins differed in level. These differences were observed mainly in colostrum and concern several high abundant proteins. This study also showed that protease inhibitors, which are highly correlated to immune-active proteins, are present in variable amounts in human milk and could be relevant during digestion.

Keywords

Mammary gland, immune-active proteins, proteases, protease inhibitors, digestive tract

Introduction

Human milk is the best source of nutrition for babies, which enhances children's immune system and influences the microbiota[1-3]. Health benefits have been related to the presence and concentration of human milk components like oligosaccharides and proteins[4-5]. There are two distinct groups of proteins in human milk; caseins and milk serum proteins[6]. Human milk in early lactation consists of approximately 30% caseins and 70% serum proteins, with a 50:50 ratio typically found after a 6 month lactation period[6].

Serum proteins in human milk have been categorized according to their main and highly diverse biological functions[7-8]. It was found that immune-related proteins, transport proteins, and enzymes were present in the largest quantities, and their concentrations generally decrease over lactation[7-8]. Immune-active proteins not only protect infants against pathogenic microorganisms, but also confer passive immunity to the neonate until its own immune system has been fully developed[9-11]. Serum proteins in human milk also include an array of blood coagulation proteins, membrane proteins, signaling proteins, and protease inhibitors[9-11]. Protease inhibitors play a key role in the blood coagulation cascade and complement pathway[12-14], and might protect proteins against degradation by proteases in the mammary gland and even in the infant's gastrointestinal tract[12-18].

There is a wide range of proteins (e.g., α_{S1}-, β-, and κ-casein, lactoferrin, immunoglobulins, serum albumin, and α-lactalbumin) in relatively high concentrations in human milk[19]. Most milk proteins are synthesized in the mammary gland, except for immunoglobulins and serum albumin[19]. Serum albumin can enter milk via the paracellular pathway and immunoglobulins are transported from blood through mammary epithelial cells by a receptor-mediated mechanism[19]. Caseins are transport proteins that form micelles, and these micelles are capable of binding, and thereby transporting, minerals. Caseins can easily be digested in the infant's gastrointestinal tract[15-18], being a valuable source of amino acids and minerals, which can easily be absorbed. Milk serum proteins such as lactoferrin, immunoglobulins, serum albumin, and α-lactalbumin cover 90% of the milk serum proteome in abundance[20]. The milk serum protein α-lactalbumin is required for the synthesis of lactose, supplies infants with large amounts of tryptophan, and facilitates the absorption of essential minerals[21]. Several other milk serum proteins, like lactoferrin and immunoglobulins, protect infants against pathogens and decrease the risk of having acute or chronic diseases[21-22]. Lactoferrin, a globular glycoprotein of the transferrin family, partially ends up intact in the infant's faeces, and was shown to influence the microbiota composition of neonates[22]. Human milk is also a rich source of antibodies or immunoglobulins, which are able to recognize and bind to unique epitopes of pathogens, preventing their colonisation[23-25]. Serum albumin is a protein mainly involved in the transportation of hormones, fatty acids, and other milk components[21].

Individual differences in milk serum proteins between mothers have been reported, where it was found that there was a large overlap in identified proteins in human milk among

mothers, whereas there were also major quantitative changes, both between mothers and over time[7]. Given the various potential benefits of milk serum proteins, it would be of interest to obtain insights in the variability of serum proteins in human milk from mothers from other geographical and ethnic origin.

Therefore, the main objective of this study was to investigate the milk serum proteome of 7 Chinese mothers and to investigate the variability in type and level of serum proteins in Chinese human milk over a 20-week lactation period using liquid chromatography–tandem mass spectrometry (LC-MS/MS). Additionally, the type and level of serum proteins in Chinese human milk were compared to colostrum and mature milk from Dutch mothers.

Materials and Methods

Set-up study and sample collection

Chinese participants were recruited in the Hohhot region, China, between August 2014 and November 2015 by the Yili Innovation Center (Hohhot, CN). Yili organized the collection of the human milk, including sampling using a human milk pump. For every time point, a volume of 10 mL was collected in a polypropylene bottles. Milk bottles were shaken gently, aliquoted directly into 2 mL Eppendorf tubes, and stored at −20 °C. Milk samples of 7 healthy mothers who delivered term (38–42 weeks) infants were assessed in week 1, 2, 4, 8, 12, and 20 postpartum. Human milk collection was approved by the Chinese Ethics Committee of Registering Clinical Trials (ChiECRCT-20150017). Written informed consent was obtained from all mothers. Milk collection and analysis of milk of 4 Dutch mothers over a 24-week lactation period is described before and was a collaboration with the Dutch Human Milk Bank (Amsterdam, NL)[7]. Healthy women who delivered singleton term infants (38–42 weeks) were eligible for that study. The data from these analyses were re-used and made compatible with the Chinese data within this research facilitating direct comparison, as explained further in this section under data analysis.

Milk serum preparation and concentrations

Human milk samples (5 mL) were fractionated, as described previously[10]. Briefly, the milk fat was removed by centrifugation (10 min, 1,500 g, 4 °C) and the obtained skim milk was transferred to ultracentrifuge tubes. After ultracentrifugation (90 min, 100,000 g, 4 °C), the top layer represented the remaining milk fat still present, the middle layer was milk serum (with some free soluble caseins), and the bottom layer consisted of micellar casein. The free soluble caseins are part of the milk serum proteome. A comparative study previously showed that ultracentrifugation is the most effective method to separate caseins from serum proteins[26], although it is not possible to rule out low amounts of serum proteins in the casein pellet[6]. Milk serum concentrations were measured in duplicate using the bicinchoninic acid

(BCA) protein assay kit (Thermo Scientific Pierce, Massachusetts, U.S.), to ensure that the same amount of protein (10 µg) was used for further sample preparation. Bovine serum albumin was used as standard for making a BCA calibration curve.

Sample preparation, dimethyl labeling, protein digestion, and peptide analysis

Milk serum samples were prepared for protein analysis using filter-aided sample preparation and dimethyl labeling, as described previously[27]. Milk serum (20 µL) was mixed with a buffer containing sodium dodecyl sulfate (SDS) for protein denaturation and dithiothreitol (DTT) to reduce the disulfide bridges in proteins, after which the samples were loaded on a Pall 3 K omega filter (10-20 kDa cutoff, OD003C34, Pall, Washington, U.S.) for protein digestion. The lysis buffer contained 0.1 M Tris/HCl pH 8.0 + 4% SDS + 0.1 M DTT to get a 1 µg/µL protein solution. Next, 180 µL of 0.05 M iodoacetamide/urea (0.1 M Tris/HCl pH 8 + 8 M urea) was used for protein alkylation. Samples were washed three times with 100 µL of 8 M urea, using centrifugation, followed by 110 µL of 50 mM ammonium bicarbonate (ABC). Then 0.5 µg trypsin in 100 µL ABC was added, followed by overnight incubation at room temperature while mildly shaking, and centrifuged to separate peptides from undigested material. The trypsin digested samples were then labeled, using distinct combinations of isotopic isomers of formaldehyde and cyanoborohydride, leading to a unique stable isotope composition of labeled peptide doublets with different masses[27]. After dimethyl labeling, the prepared samples were analysed using LC-MS/MS, as described before[7]. For LC-MS/MS, a 0.10x30 mm ProntoSil 300-5-C18H (Bischoff, Leonberg, DE) pre-concentration column (prepared in house at a maximum of 270 bar) was used, and the full scan FTMS spectra were measured in positive mode between m/z 380 and 1400 on a Thermo LTQ-Orbitrap XL. MS/MS scans of the four most abundant doubly- and triply-charged CID fragmented peaks in the FTMS scan were obtained in data-dependent mode in the linear trap (MS/MS threshold = 5.000).

Data analysis

The MS/MS spectra obtained were processed by the software package Maxquant 1.3.0.5 with the Andromeda search engine, as described previously[28]. Protein identification and quantification was done according to literature[7]. Maxquant created a decoy database consisting of reversed sequences to calculate the false discovery rate (FDR). The FDR was set to 0.01 on peptide and protein level. Minimum required peptide length was 7 amino acids, and proteins were identified based on minimally 2 distinct peptides. The intensity based absolute quantification (iBAQ) values were selected, representing the total peak intensity as determined by Maxquant for each protein, and after correction for the number of measurable peptides[7]. The iBAQ values have been reported to have a good correlation with known absolute protein amounts over at least four orders of magnitude[29]. For data normalization, iBAQ values for each protein were transformed into BCA equivalent milk

serum protein concentrations, by dividing the iBAQ values of each protein in a sample by the summed iBAQ values of all protein within a sample, and multiplied with the corresponding milk serum protein concentration based on the BCA assay. To facilitate direct comparison between Chinese and Dutch data within this research, BCA equivalent values at time points week 12 and 20 postpartum were compared to week 16 and 24, respectively. The biological function was assigned to all the serum proteins using the online UniprotKB database, as done previously[7]. To assign a specific function to multifunctional proteins, DAVID Bioinformatics Resource 6.7 was used additionally for further protein biological function classification and clarification[30].

Statistical analysis

Statistics was based upon previously described methods[7], with modifications. For the BCA equivalent values of each protein in Chinese and Dutch human milk over lactation, a regression line was fitted using R (Lucent Technologies, New York, U.S.), summarizing the profile over time for each protein into an intercept and slope. The calculated intercepts are the protein BCA equivalent values at week 1, the calculated slopes indicate the decrease or increase in BCA equivalent values per week. To determine the significant different milk serum proteins over lactation per country, a comparison was done based on the calculated slope. Only BCA equivalent values of the common serum proteins found in both Chinese and Dutch human milk were used for comparison. The common serum proteins in Chinese and Dutch human milk were then evaluated based on the calculated intercept and slope using a two-tailed *t*-test, with a significance level set on $\alpha = 0.05$. Next, these common milk serum proteins were compared in Chinese and Dutch human milk using a two-tailed *t*-test in Perseus[31], separately for each lactation week, with correction for multiple testing based on permutation-based FDR. The BCA equivalent values of serum proteins in Chinese and Dutch human milk were also summed per function and were then compared using a two-tailed *t*-test. To quantify the relation between biological function groups, Pearson correlation coefficients were calculated for summed BCA equivalent values and visualized in correlation matrix plots. Pearson correlation coefficients > 0.5 were considered good. All the serum proteins in Chinese and Dutch human milk were plotted in a graph, to visualize the differences in serum proteins over lactation.

Results

The objective of this study was to investigate the variability in type and level of serum proteins in Chinese human milk over a 20-week lactation period. For this, the milk serum proteome of 7 mothers over lactation was investigated using LC-MS/MS.

Level and type of milk serum proteins in Chinese human milk

Total milk serum protein concentrations in Chinese human milk of the 7 mothers over lactation are presented in Figure 1. Concentrations ranging from 12 to 25 g/L decreased significantly ($\alpha < 0.05$) over a 20-week lactation period, although with large individual variations (Figure 1).

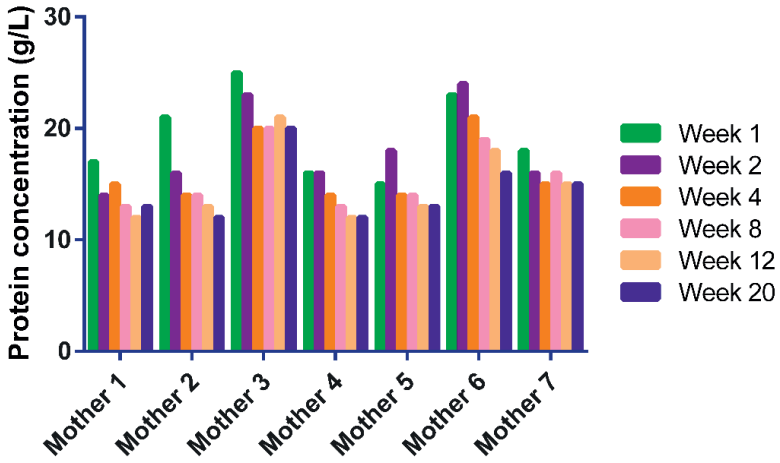


Figure 1. Total BCA serum protein concentrations (g/L) in Chinese human milk per mother over a 20-week lactation period.

Serum proteins in human milk were grouped based on their main biological functions (data not shown). Not only the total protein concentrations, but also the protein composition differed among mothers and over lactation as measured after protein digestion and subsequent LC-MS/MS analysis (Figure 2).

The figure shows that immune-active proteins, transport proteins, and enzymes were the most abundant for all mothers (Figure 2). The percentage of total protein attributable to these main biological functions, however, varied widely among mothers (Figure 2). Although the BCA equivalent values were always higher in colostrum than in mature milk, the rate of decline for the three main groups varied among mothers (Figure 2).

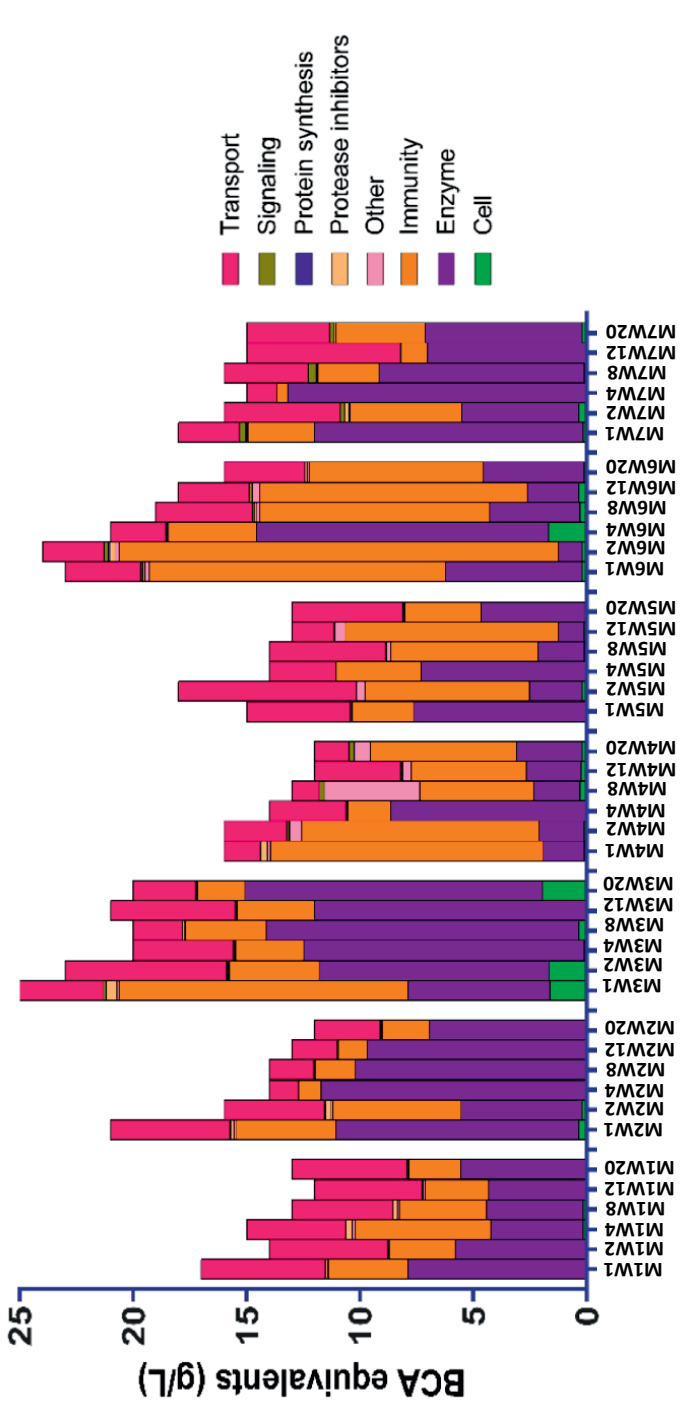


Figure 2. Serum protein composition in human milk of 7 Chinese mothers over a 20-week lactation period, based on BCA equivalent values (g/L). The number behind the M indicates the mother, and the numbers behind the W (1 to 20) indicates the number of weeks postpartum.

To facilitate the comparison between Chinese and Dutch human milk, data were averaged among mothers, as shown in Figure 3. Average total BCA equivalent values in Chinese human milk for enzymes, immune-active proteins, and transport proteins ranged between 4.5–10.0 g/L, 2.9–7.8 g/L, and 2.9–5.0, respectively (Figure 3).

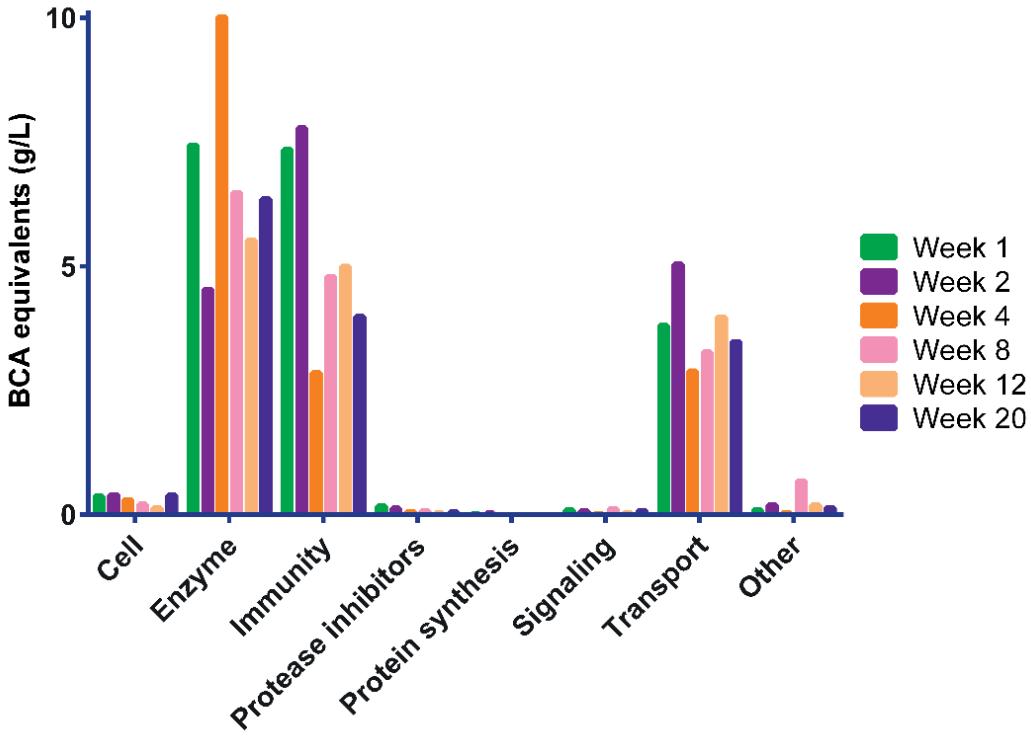


Figure 3. Averaged BCA equivalent values (g/L) of serum proteins for human milk from 7 Chinese mothers categorized per biological function over a 20-week lactation period.

Comparison of the Chinese and Dutch milk serum proteome

The type and level of serum proteins in Chinese human milk were also compared to Dutch human milk. The raw data on Dutch human milk was reprocessed to be compatible with the Chinese data. The total BCA milk serum protein concentrations in Dutch human milk per mother and over lactation are available as supplementary information (Figure S1). Total BCA equivalent values in Dutch human milk decreased over a 24-week lactation period from 21.6 to 13.6 g/L (Figure S2). Enzymes, immune-active proteins, and transport proteins were also the most abundant in Dutch human milk over lactation (Figure S2). The BCA equivalent values

for the groups enzymes, immune-active proteins, and transport proteins in Dutch human milk ranged from 4.5–9.0 g/L, 3.8–5.6 g/L, and 4.8–6.8 g/L, respectively. Although different patterns in Chinese and Dutch human milk can be observed, the difference was not significant between the same group of biological functions (data not shown), except for cell and signaling, where levels were higher in Chinese human milk.
The relations between the levels of different biological function groups of serum proteins within the Chinese and within the Dutch human milk population were visualized in a correlation matrix plot (Figure 4).

Individual milk serum proteins
A total of 469 and 200 serum proteins were measured in Chinese and Dutch human milk, respectively. The milk serum proteome of different Chinese and Dutch mothers were compared based on 166 common milk serum proteins. The overall 15 most abundant milk serum proteins can be found in Table 1.

Table 1. The 15 most abundant serum proteins categorized per function in both Chinese and Dutch human milk over lactation, with their corresponding BCA equivalent values (g/L) values at week 1.

Function	**Protein Name**	**BCA equivalent values (g/L)**	
		Chinese	**Dutch**
Enzyme	α-lactalbumin	6.98	8.73
	Bile salt-activated lipase	0.29	0.19
Immunity	Lactoferrin	3.74	2.10
	Ig α_1-chain c-region	0.91	0.71
	Ig λ_2-chain c-region	0.47	0.54
	Ig κ-chain c-region	0.39	0.90
	Polymeric immunoglobulin receptor	0.41	0.39
	Clusterin	0.23	0.17
	Osteopontin	0.17	0.19
	β_2-microglobulin	0.16	0.16
Protease inhibitors	α_1-antichymotrypsin	0.11	0.08
Transport	β-casein †	1.17	3.91
	α_{S1}-casein †	1.33	1.34
	Serum albumin	0.93	1.06
	κ-casein †	0.23	0.29
	Fatty acid-binding protein	0.07	0.13

† Micellar caseins were completely removed, while this was not the case for the free soluble part of the caseins.

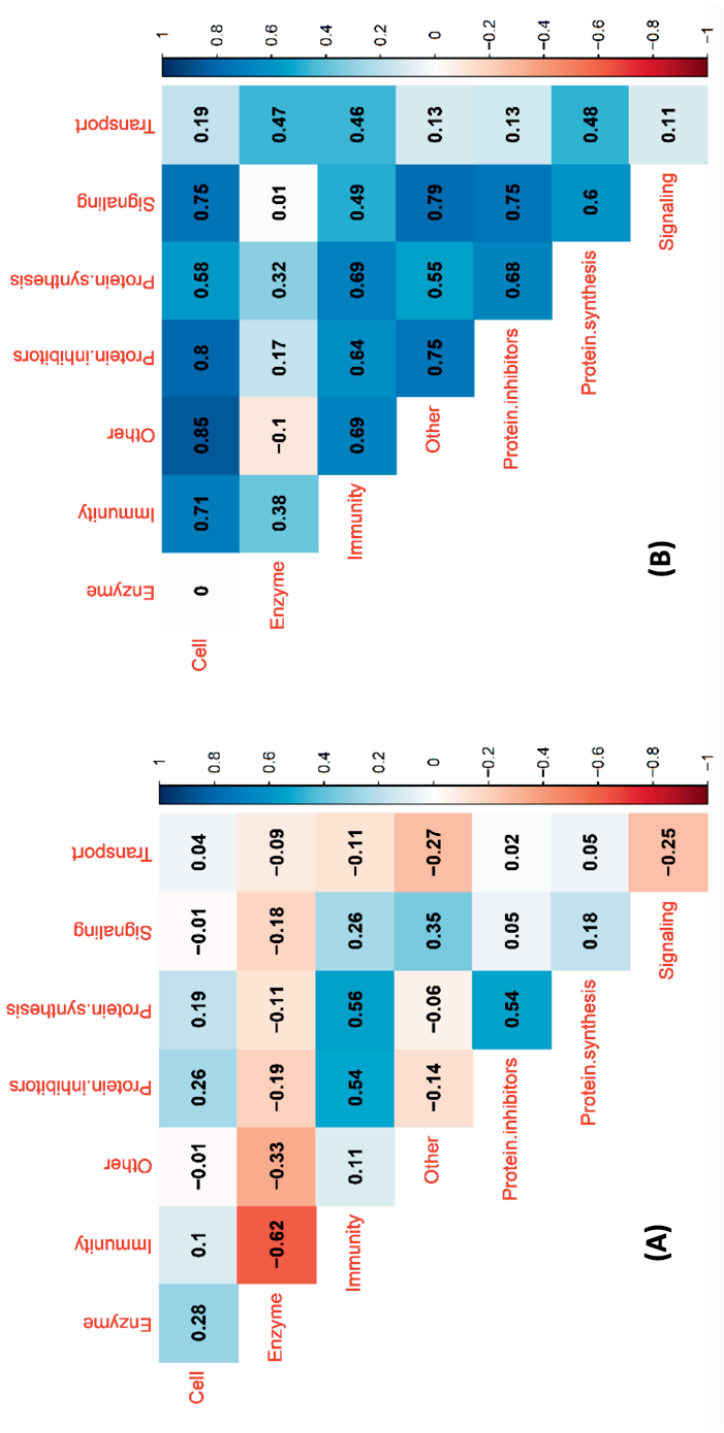


Figure 4. Calculated Pearson correlation coefficients between the different functional groups of serum proteins in Chinese and Dutch human milk, using the summed BCA equivalent values (g/L) over lactation. **(A)** Chinese human milk and **(B)** Dutch human milk.

In Chinese human milk, α_1-antichymotrypsin belongs to the top 15 serum proteins and not the transport protein fatty acid-binding protein (Table 1). Within the group enzymes, the highly abundant α-lactalbumin and bile salt-activated lipase are mainly responsible for the changes in this group in human milk over the course of lactation (Table 1). Many immune-active proteins, like lactoferrin, osteopontin, different types of immunoglobulins, polymeric immunoglobulin receptor, and clusterin, belong to the most abundant serum proteins in human milk (Table 1). The changes within the group of transport proteins over lactation can mainly be explained by the caseins (Table 1). The caseins in Table 1 probably refer to the free, non-micellar casein, as the micellar casein should have been removed by the sample preparation. With the majority of the caseins in milk being part of the micellar fraction, the caseins in Table 1 therefore do not reflect levels of total casein.

Differences in protein patterns between Chinese and Dutch human milk were examined by comparison of both the intercept (representing colostrum) and slope (representing the decline over lactation) of curves, fitted for the 166 common milk serum proteins. The p-values for these differences are shown in Figure 5, after using a two-tailed *t*-test.

The levels of two serum proteins (elongation factor 2 and myristoylated alanine-rich c-kinase substrate) varied in Chinese and Dutch human milk over lactation, as shown by the significantly different slope (Figure 5, area A). Next to that, levels of 35 serum proteins varied in intercept (Figure 5, area B), including several proteins from the top 15 (Table 1), as seen in green. The complete list of significantly different serum proteins in Chinese and Dutch human milk are shown in Table 2, grouped according to their biological function.

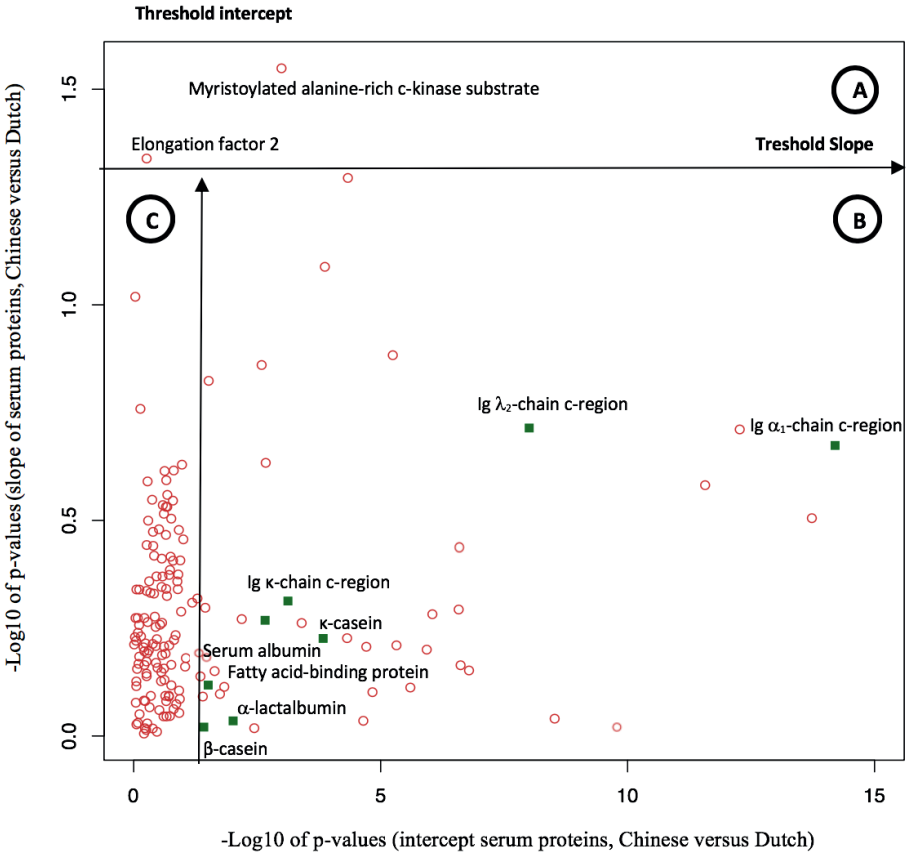


Figure 5. Comparison of the common serum proteins in Chinese and Dutch human milk over lactation. Green squares indicate the proteins displayed in Table 1. For each serum protein in Chinese and Dutch human milk over lactation, a regression line was fitted, summarizing the profile for each protein into an intercept (representing week 1) and slope (representing rate of change over lactation). These profiles were used for comparison between Chinese and Dutch human milk and the *p*-values for differences between them plotted. **(A)** Significant different proteins in Chinese and Dutch human milk over lactation, based on difference in slope, **(B)** significant different proteins in Chinese and Dutch human milk at week 1, based on intercept, and **(C)** no significant difference.

Table 2. Significantly different serum proteins in Chinese and Dutch human milk, with p-values for week 1 (intercept) and over lactation (slope).

Function	Protein Name	p-values of serum proteins (Chinese versus Dutch)	
		Intercept	**Slope**
Cell	Actin	0.002 *	0.540
	Calreticulin	0.000 *	0.620
	Follistatin-related protein 1	0.003 *	0.140
	MARCKS-like protein 1	0.004 *	0.959
	Protein deglycase DJ-1	0.000 *	0.051
	Peroxiredoxin 2	0.002 *	0.233
Enzyme	4-trimethylaminobutyraldehyde dehydrogenase	0.000 *	0.590
	α-lactalbumin	**0.010** *	**0.922**
	Fructose-bisphosphate aldolase A	0.000 *	0.710
	Isocitrate dehydrogenase 1	0.000 *	0.310
	L-lactate dehydrogenase A	0.000 *	0.772
	Nucleoside diphosphate kinase A	0.000 *	0.082
	Protein disulfide-isomerase	0.000 *	0.685
	Transketolase	0.023 *	0.707
	Triosephosphate isomerase	0.000 *	0.912
	Tryptophan-tRNA ligase	0.000 *	0.131
	UTP-glucose 1-phosphate uridylyltransferase	0.000 *	0.630
Immunity	Complement C4B	0.000 *	0.200
	Ig α_1-chain c-region	**0.000** *	**0.210**
	Ig γ_3-chain c-region	0.000 *	0.190
	Ig κ-chain c-region	**0.001** *	**0.490**
	Ig λ_2-chain c-region	**0.045** *	**0.640**
	Granulins	0.018 *	0.800
	Lysozyme C	0.000 *	0.937
	Monocyte differentiation antigen CD14	0.015 *	0.770

Table 2 (continued)

Function	Protein Name	p-values of serum proteins (Chinese versus Dutch)	
		Intercept	**Slope**
Protease inhibitors	Inter-α-trypsin inhibitor heavy chain H2	0.000 *	0.522
Protein synthesis	Elongation factor 2	0.547	0.050 *
Signaling	14-3-3 protein β/α	0.000 *	0.372
Transport	Apolipoprotein E	0.036 *	0.500
	β-casein †	**0.000** *	**0.590**
	Fatty acid-binding protein	**0.000** *	**0.790**
	Heat shock protein HSP 90-beta	0.040 *	0.810
	κ-casein †	**0.038** *	**0.950**
	Selenium-binding protein 1	0.006 *	0.536
	Serum albumin	**0.031** *	**0.760**
	Transcobalamin 1	0.000 *	0.509
Other	Myristoylated alanine-rich c-kinase substrate	0.001	0.028 *

Bold: indicates the proteins displayed in Table 1. † Micellar caseins were completely removed, while this was not the case for the free soluble part of the caseins. * Corresponding p-values (two-tailed *t*-test, $\alpha < 0.05$).

The levels of the 166 common milk serum proteins in the Chinese and Dutch population that increased or decreased over lactation, can be found as supporting information (Table S1). The levels of 17 (10%) and 21 (12%) of the 166 common milk serum proteins changed over lactation in Chinese and Dutch human milk, respectively. In addition, the 166 common serum proteins were compared between Chinese and Dutch human milk for each week separately (Table S2). This showed that 16 of 17 proteins that significantly differed in week 1 were also significantly differing in one or more of the other weeks.

Discussion

The level and type of serum proteins in Chinese human milk

The total protein concentrations decrease significantly over a 20-week lactation period per mother, although with individual variations (Figure 1). These milk serum protein concentrations match with those observed in earlier studies, ranging from 12 to 25 g/L[7,32–34], although other studies report lower values from 7 to 16 g/L over lactation[3,24,35–36]. These differences may be explained by the BCA method[37–38], which generally overestimates the total protein in human milk by about 25–40%[37–38]. The serum protein levels in this study should thus be regarded as semi-quantitative, although this did not influence the comparisons reported here, as they are all based on the BCA method. Although the protein

content seems high for milk serum, it should be taken into account that the samples with the highest protein content are actually those in early lactation. These samples are known to have higher protein and relatively lower casein contents[6], leading to higher milk serum protein contents. In addition, part of the casein remained in the sample after sample preparation and therefore also counted towards the BCA protein content.

As described previously[5], human milk becomes fully mature between 4 to 6 weeks postpartum, with amounts of bioactive components decreasing relatively to the nutrients. In early life, infants have an immature intestinal immune system, making them more vulnerable to infection by opportunistic pathogens[5]. The high levels of immune-related milk serum proteins in colostrum (Figure 3) may provide protection to the infant in this sensitive stage of its development.

It was also observed that a large variability exists in the milk serum protein composition in colostrum among Chinese mothers (Figure 2). The results in this study comprising milk of 7 mothers shows that immune-active proteins, enzymes, and transport proteins are highly abundant in Chinese human milk (Figure 3), which can also be observed from the individual data of mothers (Figure 2). Earlier studies had already shown that immune-active proteins, enzymes, and transport proteins were present in the largest quantities over lactation[7,9,11].

The top 15 most abundant milk serum proteins

The large quantities of immune-active proteins are especially driven by the abundance of lactoferrin, immunoglobulins, polymeric immunoglobulin receptor, clusterin, osteopontin, and β_2-microglobulin (Table 1), which may protect infants against pathogenic microorganisms, and confer passive immunity to the neonate until its own immune system has been developed[9–11]. As shown in Table 1, transport proteins, like free soluble caseins, serum albumin, and fatty acid binding protein were present in large quantities over lactation. Free soluble caseins could not be removed from the milk, unlike the micellar casein that can be pelleted by ultracentrifugation, a phenomena that is also reported by others[7,19,24]. Free soluble and micellar caseins belong to the most abundant proteins in human milk and these proteins mainly supply infants with amino acids and minerals needed for their growth[23–25]. It can also be observed from Table 1 that enzymes are the largest group of proteins across lactation. The large quantities of enzymes in human milk can be explained by the presence of α-lactalbumin, which is known to be the most abundant milk serum protein (Table 1). This enzyme is required for the synthesis of lactose, the main macronutrient in milk[5,21]. It should be noticed that α-lactalbumin does not have enzymatic activity on its own. Besides α-lactalbumin, bile salt-activated lipase belongs to the 15 most important enzymes in Chinese and Dutch human milk over lactation (Table 1). Bile salt-activated lipase supports the digestion of fats in the immature infant digestive tract, and facilitates the absorption of cholesterol, vitamin A, and triacylglycerols[7]. The protease inhibitor α_1-antichymotrypsin is

also among the 15 most abundant human milk serum proteins, and, like other protease inhibitors and proteases, might play a key role in digestion of human milk[12–14]. Overall, the 15 most abundant proteins identified in this study were in levels dominating the entire milk composition, covering >95% of both the Chinese and Dutch milk serum proteome.

Proteases and protease inhibitors

Proteases may play a key role in digestion of human milk. Although trypsin was the most abundant protease in Chinese and Dutch human milk, many other proteases (e.g., cytosol aminopeptidase, elastase, kallikrein, plasmin, cathepsins) were found, albeit to a lesser extent (data not shown). As described by others, proteases might be present in human milk to hydrolyze proteins in the mammary gland to regulate casein micelle size[14–15]. Protein digestion in human milk by proteases target specific proteins (e.g., caseins, polymeric immunoglobulin receptor, osteopontin) that do not have an extensive tertiary structure and are thus more accessible to proteolytic cleavage[16,18]. These proteins were, in this study, part of the overall 15 most abundant proteins in Chinese and Dutch human milk over lactation (Table 1). Especially the caseins are much digested[16–18], which indicates that proteases and bile salt-activated lipase in human milk aids overall in the digestion of two of its main macronutrients, fats and proteins[19].

Besides proteases, human milk also contains protease inhibitors. The ratio between protease inhibitors and proteases in colostrum is circa 10:1. The most abundant protease inhibitors were α_1-antichymotrypsin, α_1-antitrypsin, cystatin C, and phosphatidyletanolamine-binding protein. As described by others, α_1-antichymotrypsin binds to chymotrypsin and other chymotrypsin-like serine proteases in human milk, while α_1-antitrypsin inhibits proteases, such as trypsin, elastase, plasmin, and thrombin, and irreversibly deactivates trypsin *in vitro*[12–15]. A correlation was found between protease inhibitors and immune-active proteins in Chinese and Dutch human milk (Figure 4). Previous literature focused specifically on the relation between serine protease inhibitors and immunoglobulins[7], which also in our data showed stronger correlations than for all protease inhibitors and all immune proteins (Figure S3). A correlation higher than 0.7 was also found in both Chinese and Dutch milk between proteases and protease inhibitors specifically (data not shown). A previous study presented an overview of the proteolytic system network in human milk[15], which consists of several proteases, protease inhibitors, and blood coagulation proteins, indicating that these protein groups share a common biochemical pathway, which may explain their correlations.

Where some of the major proteins are partially digested by milk proteases in human milk, most immune-active proteins are less sensitive to digestion by these proteases, due to their compact folded globular structure, that can't be as easily digested[16]. For these immune-active proteins to have an immune-activating role in the small intestine, they must be protected against intestinal digestion, because they are sensitive to chymotrypsin and

trypsin[17–18]. That might be the reason that protease inhibitors present in human milk seem to target intestinal enzymes, specifically blocking trypsin, chymotrypsin, and other proteases[17–18], especially through the relative abundant α_1-antichymotrypsin and α_1-antitrypsin. Overall, protease inhibitors may thus ensure that specific proteins stay intact in the infant's digestive tract. This may also explain previous findings that several immune-active proteins (e.g., lactoferrin, lysozyme, immunoglobulins) and protease inhibitors (e.g. α_1-antichymotrypsin, α_1-antitrypsin) can be found intact in the stool of breastfed infants[17–18]. The intact proteins in the infant's stool may also be related to the simultaneous decrease in the content of immune-active proteins and protease inhibitors over lactation. Protection is less necessary later in lactation due to the development of the infant's immune system and digestive tract over time, while digestion becomes important for the release of nutrients later in lactation.

Comparison of high and low abundant serum proteins in Chinese and Dutch human milk

It appears that the milk serum proteome of Chinese and Dutch mothers is similar (Figure 3 and S2). The main purpose of this study was to evaluate the common serum proteins in Chinese and Dutch human milk over lactation. A total of 469 and 200 serum proteins were found in Chinese and Dutch human milk, respectively. Although a lower number of serum proteins was identified in Dutch human milk, still 166 serum proteins in Chinese human milk overlapped, which represents more than 95% of the milk serum proteome in concentrations. The reason for the higher number of serum proteins annotated in Chinese human milk might be due the larger sample size (48 versus 24 human milk samples), which generally leads to more identified proteins[28].

In total, 22% (37 out of 166) of the common serum proteins in human milk differed between Chinese and Dutch mothers either at week 1 or over lactation. The levels of 35 of the 166 (circa 21%) common serum proteins varied between Chinese and Dutch mothers in week 1 (Figure 5, area B). This, together with the results presented in table 2 and table S2, indicates that the differences between Chinese and Dutch human milk serum proteins were mainly in their level throughout lactation, and not in their changes over lactation, as the levels of only 2 of the 166 (circa 1%) common serum proteins identified in this study (myristoylated alanine-rich c-kinase substrate and elongation factor 2) differed over the course of lactation (Figure 5, area A, showing difference in slope). Overall, the main differences in the milk serum proteome between Chinese and Dutch human milk were observed in the level of individual proteins, and not in rate of change over lactation.

Conclusions

The milk serum proteome of Chinese and Dutch mothers was similar in relative abundance of different functional groups as well as the most abundant proteins. Some quantitative

differences were found, especially in absolute levels and not in rate of change over lactation. Human milk contains enzymes that can assist the digestion of milk proteins and lipids in the immature infant's digestive tract. Protease inhibitors, which are highly correlated to the immune-active proteins, are present in variable amounts in human milk and could be relevant during digestion and might be involved in controlling protein breakdown in the infant's intestinal tract.

References

1. Kramer, M.; Kakuma, R. The optimal duration of exclusive breastfeeding: A systematic review. Adv. Exp. Med. Biol. **2004**, 554, 63–77.
2. Liao, Y.; Alvarado, R.; Phinney, B.; Lönnerdal, B. Proteomic characterization of human milk whey proteins during a twelve-month lactation period. J. Proteome Res. **2011**, 10, 1746–56.
3. Martin, C.; Ling, P.; Blackburn, G. Review of infant feeding: Key features of breastmilk and infant formula. Nutrients, **2016**, 8, 279–90.
4. Gartner, L.; Morton, J.; Lawrence, R.; Naylor, A.; O'Hare, D.; Schanler, R.; Eidelman, A. Breastfeeding and the use of human milk. Pediatrics, **2005**, 115, 496–506.
5. Elwakiel, M.; Hageman, J.A.; Wang, W.; Szeto, I.M.; van Goudoever, J.B.; Hettinga, K.A.; Schols, H.A. Human milk oligosaccharides in colostrum and mature milk of Chinese mothers: Lewis positive secretor subgroups. J. Agric. Food Chem. **2018**, 66, 7036–43.
6. Kunz, C.; Lönnerdal, B. Re-evaluation of the whey protein/casein ratio of human milk. Acta Paediatr. **1992**, 81, 107–12.
7. Zhang, L.; de Waard, M.; Verheijen, H.; Boeren, S.; Hageman, J.A.; van Hooijdonk, T.; Vervoort, J.; van Goudoever, J.B.; Hettinga, K.A. Changes over lactation in breast milk serum proteins involved in the maturation of immune and digestive system of the infant. J. Proteomics, **2016**, 147, 40–7.
8. Zhang, L.; van Dijk, A.; Hettinga, K.A. An interactomics overview of the human and bovine proteome over lactation. Proteome Sci. **2016**, 15, 1–14.
9. Hettinga, K.A.; Van Valenberg, H.; De Vries, S.; Boeren, S.; Van Hooijdonk, T.; Van Arendonk, J.; Vervoort, J. The host defense proteome of human and bovine milk. PLoS One, **2011**, 6, 1–8.
10. Hettinga, K.A.; Reina, F.; Boeren, S.; Zhang, L.; Koppelman, G. Difference in the breast milk proteome between allergic and non-allergic mothers. PLoS One, **2015**, 10, 1–11.
11. Beck, K.; Weber, D.; Phinney, B.; Smilowitz, J.; Hinde, K.; Lönnerdal, B.; Korf, I.; Lemay, D. Comparative proteomics of human and macaque milk reveals species-specific nutrition during postnatal development. J. Proteome Res. **2015**, 14, 2143–57.
12. Chowanadisai, W.; Lönnerdal, B. α_1-Antitrypsin and antichymotrypsin in human milk: Origin, concentrations, and stability. Am. J. Clin. Nutr. **2002**, 76, 828–33.
13. Lindberg, T. Protease inhibitors in human milk. Pediatr. Res. **1982**, 16, 479–83.
14. Dallas, D.; Murray, N.; Gan, J. Proteolytic systems in milk: Perspectives on the evolutionary function within the mammary gland and the infant. J. Mammary Gland Biol. Neoplasia. **2015**, 20, 133–47.
15. Kelly, A.; O'Flaherty, P.; Fox, P. Indigenous proteolytic enzymes in milk: A brief overview of the present state of knowledge. Int. Dairy J. **2006**, 16, 563–72.

16. Dingess, K.A.; de Waard, M.; Boeren, S.; Vervoort, J.; Lambers, T.T.; van Goudoever, J.B.; Hettinga, K.A. Human milk peptides differentiate between the preterm and term infant and across varying lactational stages. Food Funct. **2017**, 8, 3769–82.
17. Dallas, D.; Guerrero, A.; Khaldi, N.; Borghese, R.; Bhandari, A.; Underwood, M. A peptidomic analysis of human milk digestion in the infant stomach reveals protein-specific degradation patterns. J. Nutr. **2014**, 144, 815–20.
18. Su, M.; Broadhurst, M.; Liu, C.; Gathercole, J.; Cheng, W.; Qi, X. Comparative analysis of human milk and infant formula derived peptides following *in vitro* digestion. Food Chem. **2017**, 221, 1895–903.
19. Lönnerdal, B. Human milk proteins: Key components for the biological activity of human milk. Adv. Exp. Med. Biol. **2007**, 11–25.
20. Lönnerdal, B. Bioactive proteins in human milk: Mechanisms of action. J. Pediatr. **2010**, 156, 26–30.
21. Newburg, D. Bioactive components of human milk: Evolution, efficiency and protection. Adv. Exp. Med. Biol. **2001**, 501, 3–10.
22. Spik, G.; Brunet, B.; Mazurier-Dehaine, C.; Fontaine, G.; Montreuil, J. Characterization and properties of the human and bovine lactotransferrins extracted from the faeces of newborn infants. Acta Pædiatr. Scand. **1982**, 71, 979–85.
23. Davidson, L.; Lönnerdal, B. Persistence of human milk proteins in the breastfed infant. Acta Pædiatr. **1987**, 76, 733–40.
24. Lönnerdal, B. Nutritional and physiologic significance of human milk proteins. Am. J. Clin. Nutr. **2003**, 77, 1537–43.
25. Jakaitis, B.; Denning, P. Human breast milk and the gastrointestinal innate immune system. Clin. Perinatol. **2014**, 41, 423–35.
26. Jensen, H.; Poulsen, N.; Moller, H.; Stensballe, A.; Larsen, L. Comparative proteomic analysis of casein and whey as prepared by chymosin-induced separation, isoelectric precipitation or ultracentrifugation. J. Dairy Res. **2012**, 79, 451–58.
27. Lu, J.; Boeren, S.; de Vries, S.; van Valenberg, H.; Vervoort, J.; Kasper, K.A. Filter-aided sample preparation with dimethyl labeling to identify and quantify milk fat globule membrane proteins. J. Proteomics, **2011**, 75, 34–43.
28. Cox, J.; Mann, M. MaxQuant enables high peptide identification rates, individualized ppb-range mass accuracies and proteome-wide protein quantification. Nature, **2008**, 26, 1367–72.
29. Schwanhausser, B.; Busse, D.; Li, N.; Dittmar, G.; Schuchhardt, J.; Wolf, J.; Chen, W.; Selbach, M. Global quantification of mammalian gene expression control. Nature, **2011**, 337–42.
30. Huang, D.; Sherman, B.; Lempicki, R. Systematic and integrative analysis of large gene lists using DAVID Bioinformatics Resources. Nat. Protoc. **2009**, 4, 44–57.
31. Tyanova, S.; Temu, T.; Sinitcyn, P.; Carlson, A.; Hein, M.; Geiger, T.; Mann, M.; Cox, J. The Perseus computational platform for comprehensive analysis of proteomics data. Nat. methods, **2016**, 13, 731–40.
32. Breckwoldt, J.; Neulen, J.; Keck, C. Lactation. From cellular mechanisms to integration. In Comprehensive Human Physiology. Greger, R., Windhorst, U., Eds.; Springer: Berlin, Germany, **1996**, pp. 2365–73.

33. Monaco, M.; Donavan, S. Human milk: Nutritional properties. In Nutrition in pediatrics: Basic science and clinical applications.; Walker, W., Watkins, J., Duggan, C., Eds,; BC Decker Inc.: Hamilton, Canada, **2008**, pp. 341-53.
34. Trend, S.; Strunk, T.; Hibbert, J.; Kok, C.; Zhang, G.; Doherty, D.; Richmond, P.; Burgner, D.; Simmer, K.; Davidson, D. Antimicrobial protein and peptide concentrations and activity in human breast milk consumed by preterm infants at risk of late-onset neonatal sepsis. PLoS One, **2015**, 10, 1–20.
35. Liao, Y.; Weber, D.; Xu, W.; Durbin-Johnson, B.; Phinney, B.; Lönnerdal, B. Absolute quantification of human milk caseins and the whey/casein ratio during the first year of lactation. J. Proteome Res. **2017**, 16, 4113–21.
36. Lönnerdal, B.; Erdmann, P.; Thakkar, S.; Sauser, J.; Destaillats, F. Longitudinal evolution of true protein, amino acids and bioactive proteins in breast milk: A developmental perspective. J. Nutr. Biochem. **2017**, 41, 1–11.
37. Wu, X.; Jackson, R.; Khan, S.; Ahuja, J.; Pehrsson, P. Human milk nutrient composition in the United States: Current knowledge, challenges, and research needs. Curr. Dev. Nutr. **2018**, 2, 1–18.
38. Perrin, M.; Fogleman, A.; Newburg, D.; Allen, J. A longitudinal study of human milk composition in the second year postpartum: Implications for human milk. Matern. Child Nutr. **2017**, 13, 1–12.

Supporting information

Table S1. Significantly different serum proteins in Chinese and Dutch human milk over lactation, based on the BCA equivalent values (g/L) over lactation (slope).

	Function	Protein Name	p-value slope
Chinese	Cell	Actin (↓)	0.040
		Platelet glycoprotein 4 (↓)	0.040
	Enzyme	Legumain (↓)	0.047
		Lipoprotein lipase (↓)	0.049
		Prosaposin (↓)	0.042
		Triosephosphate isomerase (↓)	0.045
	Immunity	Chitinase 3-like protein 1 (↓)	0.042
		Ig heavy chain V-III-region TRO (↓)	0.049
		Ig γ_3-chain c-region (↓)	0.050
		Ig κ-chain c-region (↓)	**0.046**
		Lactadherin (↓)	0.035
		Lactoferrin (↓)	0.034
		Mucin 1 (↓)	0.031
		Xanthine dehydrogenase (↓)	0.044
	Protease inhibitor	Plasma protease C_1-inhibitor (↓)	0.043
	Signaling	G-protein coupled receptor family C-5B (↓)	0.048
	Other	Fibrinogen gamma chain (↓)	0.045

Table S1 (continued)

	Function	Protein Name	p-value slope
Dutch *	Cell	Galectin 3-binding protein (↓)	0.013
		Leucine-rich α_2-glycoprotein (↓)	0.039
		Nucleobindin 2 (↓)	0.003
		Tenascin (↓)	0.018
	Enzyme	**Bile salt-activated lipase** (↓)	**0.010**
		L-lactate dehydrogenase (↓)	0.000
		Sulfhyldryl oxidase 1 (↓)	0.031
		UTP-glucose 1 phosphate uridylyltransferase (↓)	0.012
	Immunity	Haptoglobin (↑)	0.031
		Lysozyme C (↓)	0.034
		Zinc α_2-glycoprotein (↑)	0.002
	Protease inhibitors	α_1-antitrypsin (↓)	0.046
		Phosphatidylethanolamine binding protein 1 (↓)	0.015
	Transport	**α_{S1}-casein** (↓)†	**0.006**
		Fatty acid-binding protein (↑)	**0.012**
		κ-casein (↑)†	**0.010**
		Selenium binding protein 1 (↓)	0.025
		Serotransferrin (↓)	0.018
		Serum albumin (↑)	**0.001**
	Other	Chordin-like protein 2 (↓)	0.034
		Gelsolin (↓)	0.011

Bold: Indicate the proteins displayed in Table 1. † Micellar caseins were completely removed, while this was not the case for the free soluble part of the caseins. Arrows are indicative for an increase or decrease over lactation for each protein. * Significantly different serum proteins in Dutch human milk over lactation were already previously reported[7].

Table S2. Serum proteins that were significantly different in at least one of the lactation weeks. Numbers are the p-value for the difference between the Chinese and Dutch human milk serum proteins. To facilitate direct comparison between Chinese and Dutch data within this research, the time points week 12 and 20 postpartum were compared to week 16 and 24, respectively.

Protein Name	Week 1	Week 2	Week 4	Week 8	Week 12/16	Week 20/24
Ig α_2-chain c-region	**<0.0001**	**<0.0001**	**<0.0001**	**<0.0001**	**<0.0001**	**0.0000**
Ig μ-chain c-region	**<0.0001**	**0.0113**	**0.0254**	**0.0032**	**0.0007**	**0.0457**
Complement C4A	**0.0001**	**0.0029**	**0.0298**	**0.0064**	**<0.0001**	**0.0031**
Galectin-3-binding protein	**0.0001**	0.7715	0.4929	0.3741	0.4760	0.0787
Ig α_1-chain c-region	**0.0004**	**0.0002**	**<0.0001**	**<0.0001**	**0.0001**	**0.0001**
Ig γ_3-chain c-region	**0.0010**	**0.0025**	**0.0011**	**0.0003**	**0.0009**	**0.0021**
Ig κ-chain c-region	**0.0016**	0.0764	0.5174	**0.0236**	0.2549	0.1576
Mucin1	**0.0017**	**0.0328**	0.3881	0.6801	0.2955	**0.0190**
Protein S100-A9	**0.0026**	0.1452	0.5719	0.0526	**0.0137**	0.1594
Chordin-like protein 2	**0.0031**	**0.0035**	0.6800	0.8939	0.1924	0.5632
Complement C4B	**0.0041**	**0.0004**	**0.0100**	**0.0010**	**<0.0001**	**0.0005**
Sclerostin domain-containing protein 1	**0.0048**	**0.0430**	**0.0183**	0.1233	**0.0204**	0.1964
Apolipoprotein E	**0.0050**	**0.0012**	**0.0038**	**0.0119**	**0.0091**	0.0550
Transcobalamin-1	**0.0084**	0.1100	0.5721	0.1655	0.4203	**0.0004**
Ezrin	**0.0169**	**0.0100**	**0.0230**	0.0774	**0.0032**	**0.0229**
Myristoylated alanine-rich C-kinase substrate	**0.0196**	0.7578	0.1413	0.5531	**0.0030**	**0.0306**
Apolipoprotein B-100	**0.0398**	**0.0147**	**0.0000**	**0.0489**	0.1484	0.1247
β-casein	0.0533	0.8070	0.7004	**0.0015**	0.8484	**0.0491**
Protein disulfide-isomerase	0.1661	**0.0074**	0.1549	0.0599	**0.0121**	**0.0055**
Selenium-binding protein 1	0.2163	0.1200	0.0577	**0.0235**	**0.0024**	**0.0011**
45 kDa calcium-binding protein	0.3497	0.0942	0.2351	**0.0084**	**0.0047**	0.2351
Ribonuclease T2	0.3545	**0.0076**	**0.0094**	0.9075	**0.0028**	NA
Beta-2-glycoprotein 1	0.5483	**0.0002**	0.3685	**0.0230**	0.0545	0.6026
Legumain	0.6446	0.9789	0.6999	0.4638	0.8119	**0.0025**
Complement C3 β-chain	0.6921	0.1056	0.0542	**0.0422**	**0.0000**	**0.0197**
Protein S100-A11	0.7731	0.8635	**0.0007**	0.8139	0.4929	0.0940
Apolipoprotein D	0.7739	0.0991	0.0657	**0.0006**	0.5928	0.3749
Gamma-glutamyltranspeptidase 1	0.8008	0.2640	**0.0301**	0.2631	0.1578	**0.0024**
Heat shock cognate 71 kDa protein	0.9118	0.1782	**0.0241**	0.0977	**0.0002**	**0.0380**
Triosephosphate isomerase	0.9866	**0.0144**	0.0912	0.1336	**0.0028**	**0.0172**

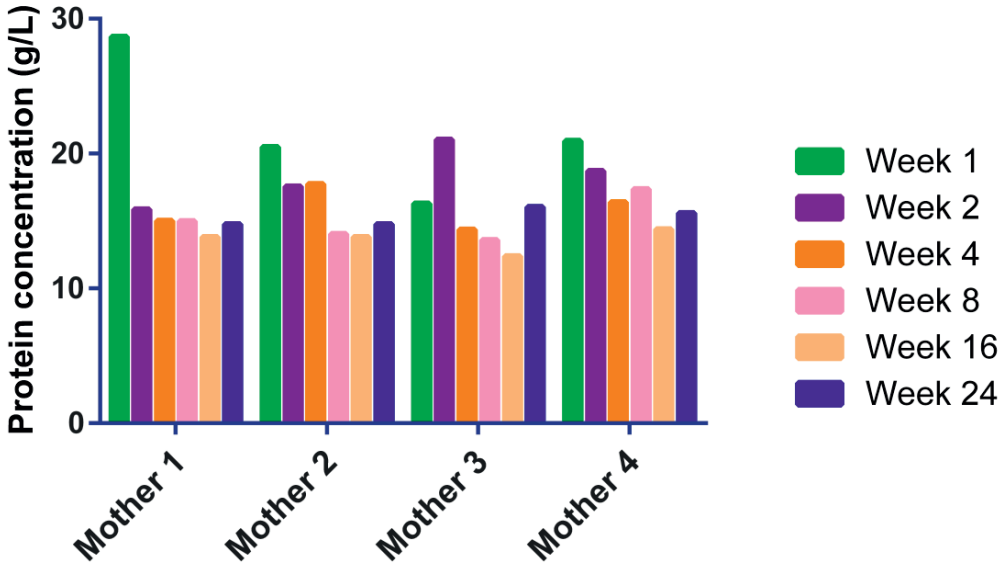


Figure S1. Total BCA serum protein concentrations (g/L) in Dutch human milk per mother over a 24-week lactation period. Raw data from Dutch human milk was re-used[7].

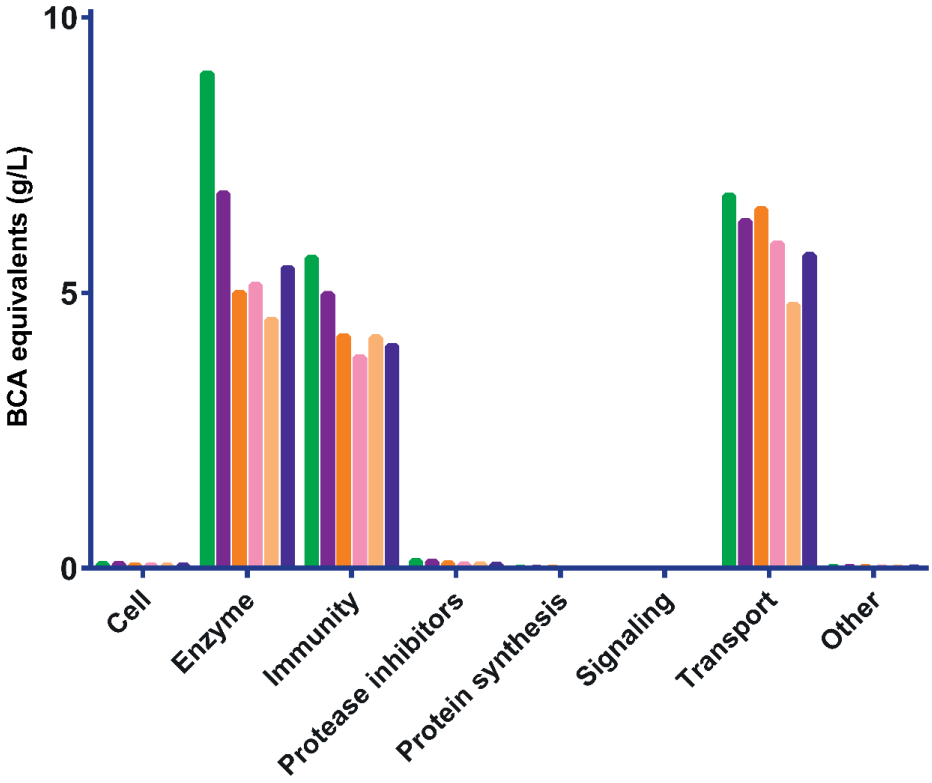


Figure S2. BCA equivalent values (g/L) of serum proteins in human milk of 4 Dutch mothers categorized per biological function over a 24-week lactation period. Raw data from Dutch human milk was re-used[7].

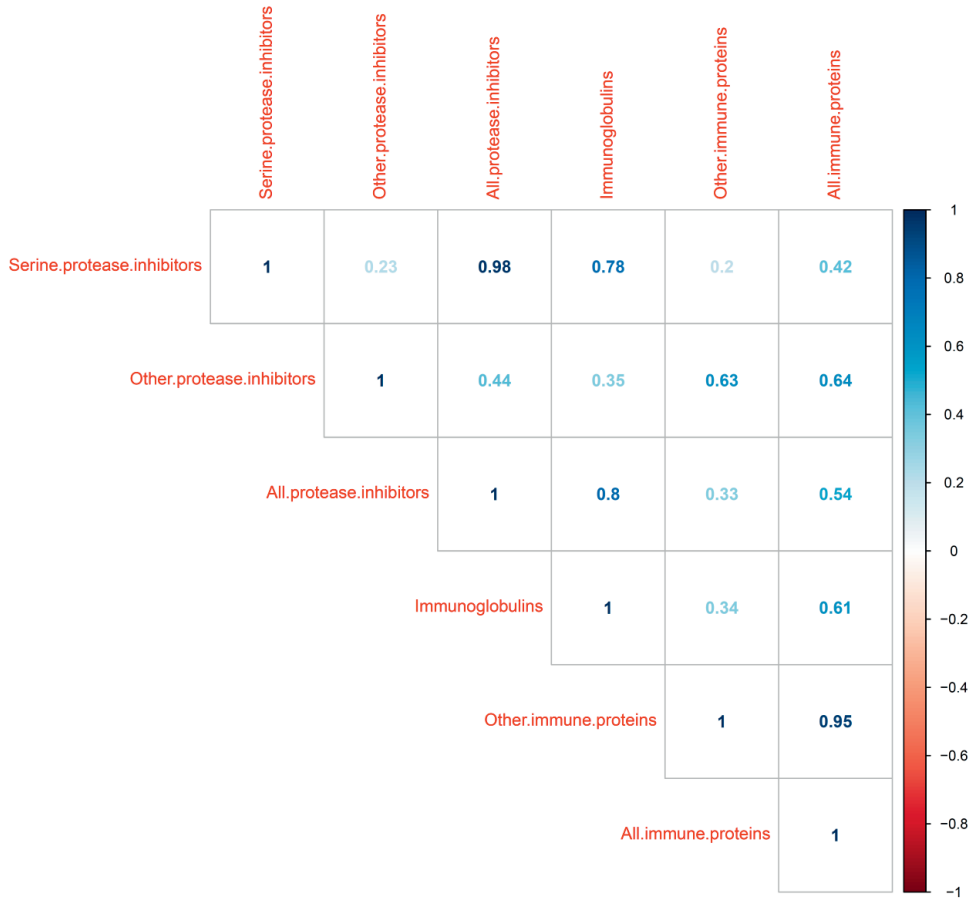


Figure S3. Correlations between the functional groups consisting of protease inhibitors (including serine and non-serine protease inhibitors) and immune-active proteins (including immunoglobulins and non-immunoglobulins) in Chinese human milk, using BCA equivalent values (g/L) over a 20-week lactation period.

Chapter 3

Infant *in vitro* digestion of proteins from colostrum and mature milk of Chinese mothers

Manuscript submitted for publication:
Elwakiel, M.; Boeren, S.; Wang, W.; Schols, H.A; Hettinga, K.A.

Abstract

This study gives insight in the degradation of human milk proteins in an *in vitro* infant digestion model, comparing colostrum (week 1) and mature milk (week 4) of 7 Chinese mothers individually. In this study, we used a new *in vitro* model, representing 0 to 3-month-old infants. The level of undigested proteins was analysed after gel-electrophoretic separation and in-gel tryptic digestion by LC-MS/MS. Using the BCA protein assay, it was shown that the total undigested milk protein content decreased from the start to the end of digestion with large variation between mothers, especially in the gastric phase (25–80%). It was also observed that undigested proteins could still be found after intestinal digestion, ranging from 0.5% to 4.2% of initial protein content. Based on LC-MS/MS analysis, protein digestion varies for the milk from the mothers individually, especially during gastric digestion. No differences could be observed between protein digestion from colostrum and mature milk after intestinal phase. The highest levels of proteins remaining after intestinal digestion can be linked to the group immune-active proteins, for all mothers. The level of protease inhibitors and total protein content in the milk did not correlate with the overall proteolysis during digestion. The results also showed that serum proteins were not completely digested during gastric digestion: with 33% remaining for colostrum and 37% remaining for mature milk, respectively. More than 40 serum proteins could be detected after intestinal digestion. Some of the highly abundant serum proteins (lactoferrin, serum albumin, bile salt-activated lipase, immunoglobulins, α_1-antichymotrypsin) were still partially present intact after intestinal digestion, for all mothers. Caseins were also not completely digested during gastric digestion: with 35% remaining for colostrum and 13% remaining for mature milk, respectively, whereas caseins were almost completely digested after intestinal digestion in most mothers. The complete degradation of these caseins into peptides might be related to their structural features. Overall, this study showed that digestion differed for the various human milk proteins, as well as between milk of the different mothers, using a newly developed infant *in vitro* digestion model, and the main differences between digestion could be observed after gastric digestion among mothers over lactation.

Keywords

Human milk proteins, infant digestion, protease inhibitors, casein, serum proteins

Introduction

Human milk is a complex mixture of nutrients and bioactive constituents, contributing to the infant's growth, development and health[1]. Human milk proteins among others play a pivotal role in protecting the infant's gut mucosa against pathogens[1]. There are two distinct types of proteins in human milk, caseins and serum proteins, with changing quantities and ratios over lactation[1]. Caseins (α_{S1}-, β-, and κ-casein) are generally described as transport proteins due to their calcium-binding properties, and become bioavailable after digestion[2]. The hydrophobic regions of caseins consist of a high number of proline residues, which prevents the formation of close-packed secondary structures[3]. The serum proteins in human milk have many different functions[4]. The most abundant serum protein groups in human milk are enzymes (e.g. α-lactalbumin, bile salt-actived lipase), transport proteins (e.g. serum albumin, fatty acid-binding protein), and immune-active proteins (e.g. lactoferrin, immunoglobulins)[5]. This latter study also showed that the serum protein composition varied among mothers and between different populations[5].

It has been assumed that caseins in human milk are fully digested in infant's digestive tract, facilitating the uptake of relatively small peptides, essential amino acids and minerals associated with the micelles[6-8]. Human milk serum proteins that do not have an extensive tertiary structure (e.g. polymeric immunoglobulin receptor, osteopontin) also may be broken down completely during infant digestion, in contrast to more tightly folded serum proteins like lactoferrin and immunoglobulins[9-13]. Some of the major proteins (caseins, osteopontin, clusterin, PIGR) are predigested by proteases, resulting in the visibility of peptides in human milk, independent of lactation period[12]. A variety of studies have reported that specific serum proteins like lactoferrin, immunoglobulins, and α_1-antitrypsin in human milk can be found intact in the stool of breastfed infants, showing that those proteins are able to partially survive digestion in the infant's digestive tract[14-18].

It has been suggested that the extent of protein digestion might be reduced by the presence of protease inhibitors (e.g. α_1-antichymotrypsin and α_1-antitrypsin) or by the high total protein levels in human milk. Protease inhibitors might inhibit the function of trypsin and other serine proteases during small intestinal digestion[19-22]. Colostrum contains a relative higher quantity of protease inhibitors than mature milk[5], which might lead to more undigested proteins from colostrum at the end of digestion. In addition, protein digestion might also be influenced by the total protein content in colostrum. Colostrum contains a higher total amounts of proteins (14–16 g/L) compared to mature milk (7–10 g/L)[1]. Therefore, it would be of interest to investigate both the variation in level of protease inhibitors as well as total protein content in relation to the degree of protein hydrolysis.

Different static *in vitro* digestion models have been developed over the years, mimicking the gastrointestinal tract of adults and 3-year-old infants[23-31]. Colostrum (between 0–2 weeks) and mature milk (> 4 weeks postpartum) are quite different in protein content and

composition[4]. Based on the gene expression of the infant's gastrointestinal tract, infant's at 4 weeks of age have among others 40% of their chymotrypsin capacity, and only 10% of their pepsin capacity available compared to adults[32]. To elaborate and better comprehend the digestion of proteins from colostrum and mature milk, a new *in vitro* digestion model was developed representing 0 to 3-month-old infant's digestion. This model is different from an existing adult *in vitro* digestion model[23], by having a higher gastric pH of 5, lower enzyme activities (pepsin 200 U/mL in gastric phase; trypsin 8.33 U/mL in intestinal phase), and shorter transition time (1 h each for gastric and intestinal)[32]. Based on these modifications, this would more mimic the situation in infants, who have generally decreased protein digestion than adults.

The aim of this study was to better understand the variation in protein digestion. The enzymatic hydrolysis of proteins in milk of 7 Chinese mothers from 2 different lactation periods (colostrum, week 1; mature milk, week 4) was investigated in a newly developed *in vitro* digestion model representing 0 to 3-month-old infants. The level of undigested proteins was analyzed by a combination of bicinchoninic acid (BCA) protein assay, sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) and liquid chromatography–tandem mass spectrometry (LC–MS/MS).

Material and Methods

Sample collection

Human milk was collected as described previously[5] and samples of 7 different Chinese mothers from 2 different lactation stages (colostrum, week 1; mature milk, week 4) were used. The samples used in this study were aliquots, but not the same samples, as described previously[5]. Healthy Chinese women who delivered singleton term infants (38–42 weeks) were eligible for this study. Human milk collection was approved by the Chinese Ethics Committee of Registering Clinical Trials (ChiECRCT-20150017). Written informed consent was obtained from these 7 mothers.

The infant in vitro protein digestion model

The INFOGEST static adult *in vitro* digestion model, as described previously[23], was modified to an *in vitro* infant (0–3 months) protein digestion model[32]. In comparison to adults, the pH and porcine pepsin concentrations in the gastric phase were adjusted, as well as the porcine pancreatin and porcine bile salt concentrations in the intestinal phase. The time to mimic each digestion phase was changed to 1 h. Briefly, skim milk (8 mL) was mixed with 6 mL of simulated gastric fluid, after which 5 µL of 0.3 M calcium chloride ($CaCl_2$) and 695 µL of water was added. Porcine pepsin diluted in simulated gastric fluid was added to reach an enzyme activity of 200 U/mL instead of 2000 U/mL for the adult model in the final gastric mixture.

The pH of the chyme was adjusted to 5 instead of 3 with 1 M HCl. The mixture was then incubated at 37 °C for 1 h while mildly shaking at 200 rpm. After incubation, the pH was adjusted to 7 with 1 M sodium hydroxide solution. For duodenal digestion, 7.5 mL of gastric chyme was mixed with 4 mL of simulated intestinal fluid electrolyte stock solution[23]. Porcine pancreatin was added to reach a trypsin enzyme activity of 8.33 U/mL instead of 100 U/mL for the adult model. After that, 2.5 mL of bile salts (40 mM), 40 µL of 0.3 M $CaCl_2$, and 1.31 mL water was added. The pH of the chyme was then again adjusted to 7 with 1 M hydrogen chloride solution. The duodenal chyme was then incubated at 37 °C for 1 h in a water bath while mildly shaking at 200 rpm. After incubation, the enzyme was inactivated with 50 µL of the irreversible serine protease inhibitor 4-(2-aminoethyl) benzenesulfonylfluoride (100 mM) in the duodenal chyme. The skim milk was diluted 4 times, while the samples after the gastric phase were diluted twice, facilitating direct comparison with the samples of the duodenal phase[32].

Total concentrations of proteins before and after infant in vitro digestion

The total protein concentration of the blank and digesta were measured in duplicate using the BCA protein assay kit 23225 (Thermo Scientific Pierce, Massachusetts, U.S.). Standards and reagents were prepared according to the manufacturer's instructions. Before analysis, 1 mL of the sample was mixed 1:1 with absolute trichloroacetic acid (TCA). After centrifugation (1,500 g for 30 min, 4 °C), the supernatant containing peptides was removed. TCA-precipitated proteins were washed twice with cold acetone to completely remove TCA, and the pellet dried at 70 °C in a heating block (Labtherm Graphit, Liebisch, DE) for 60 min. The dried proteins were re-dissolved in 2 mL of the BCA working reagent, and incubated at 37 °C in a water bath for 30 min. After cooling down to room temperature, the samples were ready for spectrophotometric measurements. Based on the BCA calibration curve using bovine serum albumin, concentrations were expressed as g/L for the diluted samples.

SDS-PAGE, in-gel digestion, and purification by solid phase extraction

For every sample, 2 µL was taken and diluted in 5 µL lithium dodecyl sulfate (LDS) sample buffer (pH 8.4, Life Technologies, Carlsbad, U.S.) and 15 µL water. This mixture was centrifuged at 1,500 g for 1 min and the supernatant heated at 70 °C in a heating block for 10 min. Samples and prestained marker (5–165 kDa, Jena Bioscience, De) were then loaded onto NuPAGE 12% Bis-Tris gels (Life Technologies, Carlsbad, U.S.). Gels were run with a LDS running buffer (containing 50 mM MES, 50 mM Tris base, 0.1% SDS, 1 mM EDTA, pH 7.3) under non-reducing conditions at 120V in a vertical electrophoresis cell (Bio-Rad, Hercules, U.S.). The gels were stained with colloidal Coomassie G-250 (Bio-Rad, Hercules, CA), followed by destaining with water and washing buffer (10% ethanol, 7.5% acetic acid in water). The

gels were scanned after staining with Image Lab version 4.1 (Bio-Rad) to visualize the protein patterns, and determine the location of α-lactalbumin on the gel.

In-gel digestion was used to digest proteins into peptides. As described previously[33], SDS-PAGE gels were incubated in 25 mL of ammonium bicarbonate (ABC) containing 0.039 g dithiothreitol (= 10 mM DTT, pH 8) for 1 h at 60 °C. Subsequently, the gels were incubated in 25 mL of Tris buffer pH 8 containing 0.092 g iodoacetamide (= 20 mM IAA, pH 8) for 1 h at room temperature in the dark. To separate proteins (> 10 kDa) from small peptides, single lanes of the SDS-PAGE gels were cut into <1 mm^3 small pieces and pieces above the band of α-lactalbumin were transferred to 1.5 mL Eppendorf low protein binding tubes. The gel pieces were frozen and thawed 3 times to increase the accessibility for trypsin. Then, 100 µL of 50 mM ABC (pH 8) containing 0.5 µg trypsin was added to the gel pieces, followed by 100 µL ABC to cover the gel pieces completely. According to the manufactures, the activity of bovine sequencing grade trypsin was ≥7500 benzoyl-L-arginine-ethyl-ester U/mL protein (Roche, Basel, CH). After trypsin digestion overnight, 15 µL of 10% trifluoroacetic acid in water was added to adjust the pH between 2 and 4 (pH-indicator strips).

Solid phase extraction was done to purify peptides, as described previously[12]. Stage tips containing 2 mg Lichroprep C18 (25 um particles) column material (C18+ Stage tip) were made in-house. The peptides were transferred to a methanol washed and 0.1% formic acid equilibrated C18 stage tip column. The peptides were eluted with 50 µL of 50% acetonitrile in water containing 0.1% formic acid. The samples were dried in a vacuum concentrator (Eppendorf, Nijmegen, NL) at 45 °C for 1 h until the volume of each sample decreased to 15 µL or less. The content of the tubes was transferred to empty low protein binding tubes, and samples reconstituted to 50 µL by adding water containing 0.1% formic acid.

LC-MS/MS and data analysis

LC-MS/MS was used to measure the amounts of distinct peptides. As described previously[33], 18 µL of each sample was injected on a 0.10x30 mm ProntoSil 300-3-C18H (Bischoff, Leonberg, DE) pre-concentration column (prepared in house at a maximum of 270 bar), peptides were eluted from the pre-concentration column onto a 0.10x200 mm ProntoSil 300-5-C18H analytical column, and the full scan FTMS spectra were measured in positive mode between m/z 380 and 1400 on a LTQ-Orbitrap XL (Thermo Fisher Scientific, Waltham, U.S.). MS/MS scans of the four most abundant doubly- and triply-charged CID fragmented peaks in the FTMS scan were obtained in data-dependent mode in the linear trap (MS/MS threshold = 5.000)[33].

MS/MS spectra for each run were obtained and analysed using Maxquant and the built-in Andromedia search engine with the Uniprot human protein database[34]. Protein identification and quantification was done as described previously[4]. Maxquant created a decoy database consisting of reversed sequences to calculate the false discovery rate (FDR).

The FDR was set to 0.01 on peptide and protein level. The minimum required peptide length was seven amino acids, and proteins were identified based on a minimum of two distinct peptides. The intensity based absolute quantification (iBAQ) values were used, representing the total peak intensity as determined by Maxquant for each protein, after correction for the number of measurable peptides[5]. The iBAQ values have been reported to have a good correlation with known absolute protein amounts over at least four orders of magnitude[35].
The Uniprot database was also used to assign functions to all individual identified proteins, as described previously[5]. The iBAQ values for each protein were used individually and summed per function, and per digestion phase (predigestion, gastric phase, intestinal phase). The iBAQ values of the proteins individually and grouped per function per phase were also converted in percentages of the total iBAQ intensity. The total iBAQ intensities of the skim milk from colostrum and mature milk were set to 100%.
To compare colostrum and mature milk of the 7 Chinese mothers on total protein (based on the BCA protein assay) after both gastric and intestinal digestion, a *t*-test for independent samples was used (R, Lucent Technologies, New York, U.S.). For this comparison, the total BCA protein concentrations were preferred over the summed iBAQ values. For statistical analysis, a FDR adjusted p-value of < 0.05 was considered significant. Scatterplots were generated with both the total protein (based on the BCA protein assay) and the levels of protease inhibitors (iBAQ values) in human milk against the relative change of total protein from milk to gastric and to intestinal digestion.

Results and discussion

Determination of total protein before and after digestion by the BCA protein assay
An *in vitro* digestion model was developed representing 0 to 3-month-old infant's digestion, which differed from an existing adult *in vitro* digestion model[23], by having a higher gastric pH of 5, lower enzyme activities in gastric phase and intestinal phase, and shorter digestion times. The parameters of the newly developed *in vitro* model, which was based on literature[32] including references and citations in that paper, represents the *in vivo* infant digestive conditions better than the adult model. Bovine milk serum was used for method development and validation of the model, as more was known about *in vitro* digestion of bovine milk proteins in older infants[2,6,11]. As the findings for bovine milk were showing similar trends with previous studies[2,6,11], although using different age-specific models, the newly developed 0 to 3-month-old infant *in vitro* digestion model was then used for human milk samples.
The disappearance of human milk proteins was quantified using the BCA protein assay. Information was obtained on the total protein content before and after digestion for milk of 7 mothers from 2 lactation periods (colostrum, week 1 and mature milk, week 4). The total

BCA protein concentrations in skim milk were ranging from 7.1 to 16.6 g/L for colostrum, and from 6.6 to 13.8 g/L for mature milk (Figure 1A). The lowest protein concentrations in colostrum (7.1 g/L) and mature milk (6.6 g/L) were for mother 7, whereas the highest protein concentrations in colostrum (16.6 g/L) and mature milk (15.3 g/L) were found for mother 4 and 1, respectively (Figure 1A).

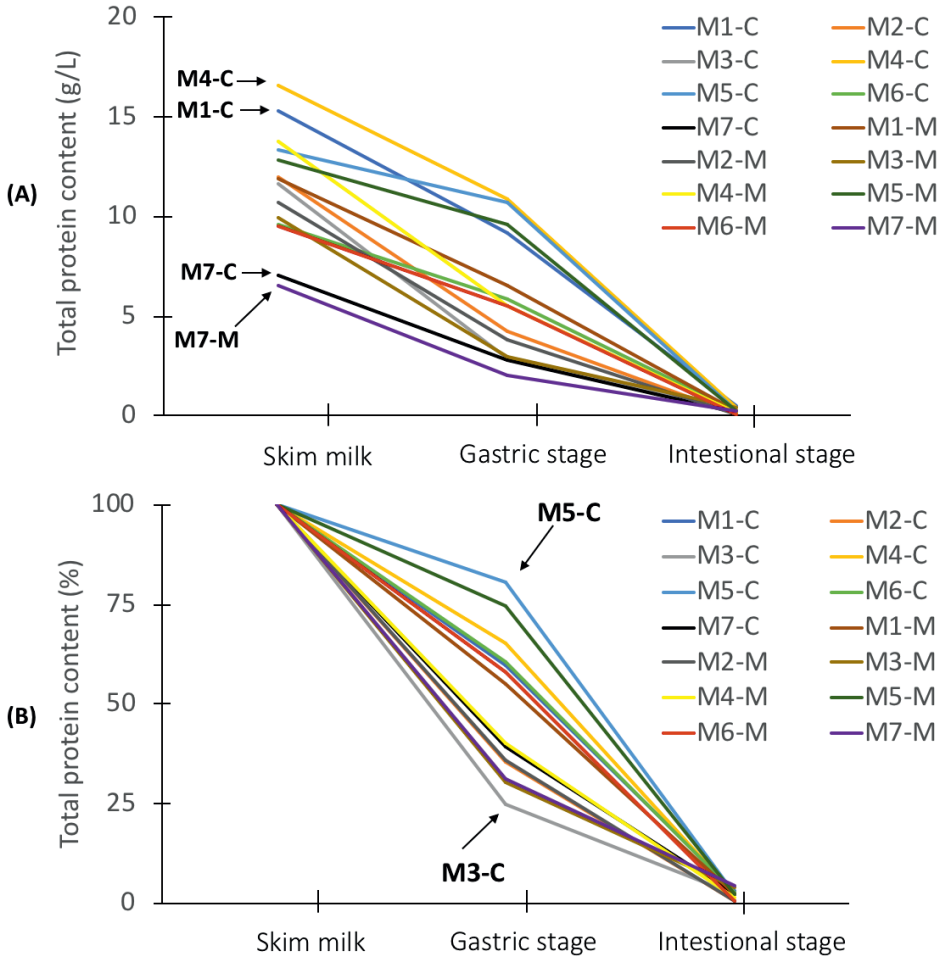


Figure 1. The total BCA protein content (g/L) **(A)** and total BCA protein content (%) **(B)** in skim milk for colostrum (week 1) and mature milk (week 4) of 7 Chinese mothers after dilution of the samples, and after gastric and intestinal stage. The total BCA protein concentrations (g/L) in diluted skim milk for colostrum and mature milk of the individual mothers was set to 100%. MX; X = mother. The letter behind the hyphen indicates colostrum (C) and mature milk (M).

Colostrum contained higher total protein concentrations compared to mature milk, although the rate of decline varied among mothers during digestion (Figure 1A).
It can be observed that the total BCA milk protein content decreased from the start to the end of digestion with large variation in decline between mothers in the gastric phase (25–80%) (Figure 1B). The total BCA protein concentrations in colostrum for mother 5 and 3 both started at 13.3 g/L and 11.7 g/L (Figure 1A), although showing the lowest (19.6%) and highest (75.3%) decline during gastric digestion (Figure 1B), respectively. It was also observed that still some undigested proteins can be found after intestinal digestion, ranging from 0.5% to 4.2% of total protein content (Figure 1B). The higher starting total protein concentrations in colostrum did not seem to be associated with a higher degree of total undigested proteins, as will be discussed in more detail later.

The most abundant human milk proteins identified by SDS-PAGE
The undigested human milk proteins were initially monitored using SDS-PAGE before performing in-gel digestion and LC-MS/MS analysis. An example of a SDS-PAGE gel is given in Figure 2, showing the outcomes of *in vitro* digestion of the most abundant proteins in colostrum and mature milk for a single mother. It can be seen from Figure 2 that caseins and α-lactalbumin from colostrum and mature milk were not readily digested in the gastric phase, but are completely digested after intestinal digestion. Lactoferrin and serum albumin, on the other hand were still partially present after the intestinal digestion (Figure 2). Bands of the individual human milk proteins on the SDS-PAGE gels were not quantified, however, it can be observed that proteins in mature milk were digested to a rather similar extent by the *in vitro* infant model as colostrum (Figure 2).
The survival of human caseins in the gastric phase, as shown in Figure 2, might be explained by the casein micellar size in human milk. It has been reported that the mean casein micellar size varies between mammal species (e.g. human, bovine, equine species) in milk and that casein micelles in human milk are much smaller in size than bovine milk[24]. The milk from bovine and human showed different degradation patterns when digested with adult gastrointestinal enzymes for 30 min at their respective pH values (pH 2.5 in gastric phase, pH 8.0 in the intestinal phase)[24]. During adult gastric digestion the caseins in human and bovine milk were poorly degraded with 69% and 39% of remaining protein, respectively[24]. Further digestion of caseins with intestinal enzymes resulted in a very fast digestion of the caseins from all species[24], with 20% of the caseins remaining intact after 5 min, while after 30 min almost no caseins were left[24].

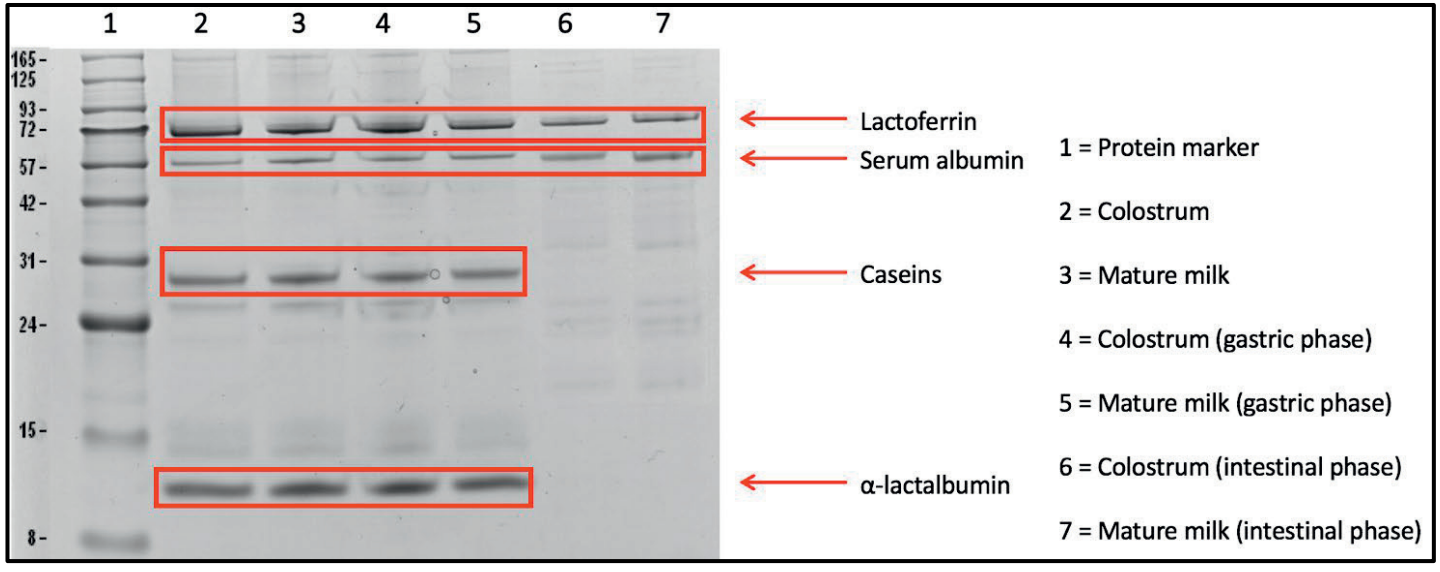


Figure 2. SDS-PAGE gel showing the *in vitro* digestion of colostrum (week 1) and mature milk (week 4) of one mother, highlighting lactoferrin, serum albumin, caseins, and α-lactalbumin.

A randomized controlled trial with 12 hospitalized tube-fed preterm infants, showed that α-lactalbumin, lactoferrin, β-casein, and serum albumin were *in vivo* only partially digested in the infant's stomach[17], which matched with our findings (Figure 2). Overall, specific human milk serum proteins partially survived *in vitro* digestion, which will be further discussed while evaluating the findings obtained by LC-MS/MS.

The protein composition before digestion by LC-MS/MS

Although a more quantitative analysis of all proteins together was done by BCA protein assay, BCA protein assay can only determine the total protein content (Figure 1). A general decrease of individual proteins was observed by SDS-PAGE, whereas SDS-PAGE is only able to identify some of the most abundant human milk proteins (Figure 2). LC-MS/MS can provide a more complete overview of the protein composition during *in vitro* infant digestion. Lanes of the SDS-PAGE gels were cut above the band of α-lactalbumin, which was thus used as threshold for in-gel digestion, meaning that proteinaceous material >10 KDa was assumped to be identified and quantified as intact protein.

The average relative protein composition in human milk of 7 Chinese mothers from 2 different lactation periods can be found in Figure 3. Proteins as present in both colostrum and mature milk were dominated by immune-active proteins, enzymes, and transport proteins (Figure 3). Colostrum contained relatively higher quantities of protease inhibitors, cell proteins, and ''other'', as compared to mature milk (Figure 3, insert). The relative levels of the latter protein groups were much lower than the transport, enzymes, and immune active proteins.

The higher relative levels of transport proteins in colostrum and mature milk in Figure 3, as compared to a previous study of the Chinese human serum proteome[5], might be explained by the presence of both free soluble and micellar caseins. In this previous study during isolation of serum proteins, micellar caseins were completely removed, but not the free soluble part of the caseins[5]. The levels of the groups enzymes and immune-active proteins in human milk are showing a similar trend as described previously[5]. The relative levels of enzymes were significantly higher in mature milk than colostrum, whereas the relative levels of immune-active proteins were significantly higher in colostrum compared to mature milk[5]. It was also reported before that large variation in protein composition existed among mothers, between populations and over time[5]. In this study, it could also be observed that the protein composition differed among mothers and over lactation (Figure S1). For Chinese mother 5, both the relative levels of immune-active proteins and enzymes increased over lactation, whereas the relative levels of transport proteins decreased from colostrum to mature milk (Figure S1). For Chinese mother 6, immune-active proteins became relatively less abundant over lactation (Figure S1).

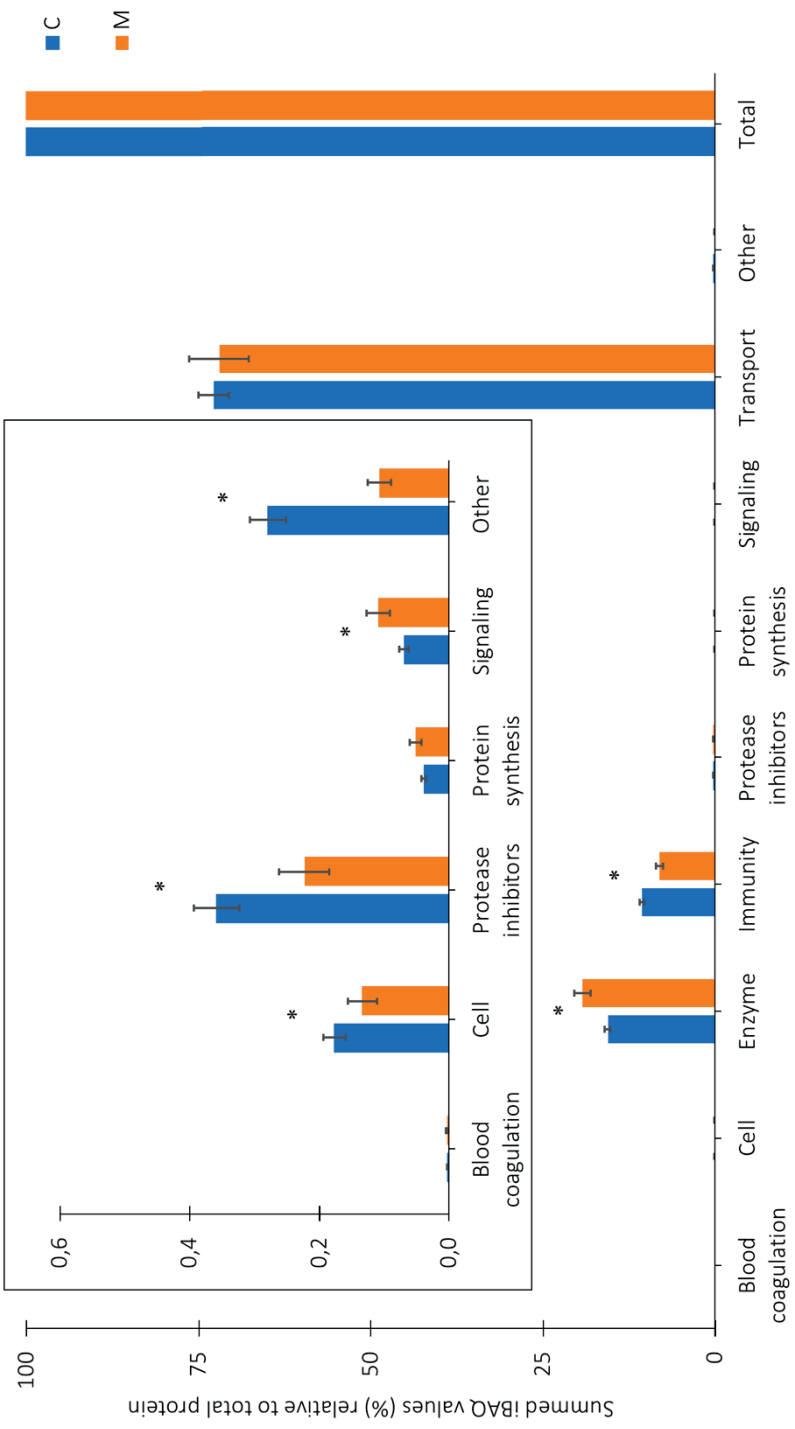


Figure 3. Proportion of the protein groups (%) relative to total protein, based on the summed iBAQ values obtained by LC-MS/MS, in colostrum (C, week 1) and mature milk (M, week 4) of 7 Chinese mothers averaged. * Corresponding p-values (two-tailed *t*-test, $\alpha < 0.05$). Insert illustrates zooming in on the low abundant protein groups.

The protein composition after gastric and intestinal digestion by LC-MS/MS

The relative levels of undigested protein for each protein group after gastric and intestinal digestion can be found in Table 1, for both colostrum and mature milk. Comparing colostrum and mature milk, it can be seen that higher percentages of undigested protein can be found in the gastric phase for most protein groups in colostrum (Table 1). In addition, the relative levels of undigested protein for the groups protease inhibitors and protein synthesis in colostrum was the highest during gastric digestion (Table 1). For mature milk protease inhibitors and immune-active proteins were digested to a lesser extent during gastric digestion (Table 1).

Table 1. Undigested human milk proteins (%) in an *in vitro* infant (0–3 months) model, based on the summed iBAQ values obtained by LC-MS/MS, categorized per biological function, for both colostrum and mature milk of 7 Chinese mothers averaged, and for gastric and intestinal digestion. The data has been normalized to 100% per protein group, based on predigestion data.

	Colostrum (week 1)		**Mature milk (week 4)**	
Protein function groups	Gastric (%)	Intestinal (%)	Gastric (%)	Intestinal (%)
Blood coagulation	25	0	7	0
Cell	31	0	18	0
Enzyme	22	0	20	0
Immunity	29	2	32	3
Protease inhibitors	44	1	35	3
Protein synthesis	41	1	14	1
Signaling	27	1	12	4
Transport	30	0	17	0
Other	22	0	22	0
Total	**29**	**0.3**	**18**	**0.4**

It should be noticed that a lot of variation in the human milk protein composition exists among mothers post-intestinal digestion (Figure S1), however, with all samples having in common that the survival level was the highest for the immune-active proteins after intestinal digestion (Figure S1). This can also be observed in Table 1. Transport proteins were quite stable against gastric digestion, but not against intestinal digestion (Table 1). This might be explained by the caseins, which were the main proteins within the transport protein group. Caseins were almost completely digested after intestinal digestion, as can be observed from Table 2. The variation in protein digestion might be affected by the initial milk

composition and protein profiles of the individual mothers. The static *in vitro* digestion model did not attribute to the variability in milk digestion for all the 7 mothers, as the conditions for both lactation points were identical.

The effect of the higher levels of protease inhibitors and total protein content in colostrum on protein digestion

The variation in levels of protease inhibitors and total protein were investigated in relation to the level of undigested protein, as both were hypothesized to reduce overall proteolysis. It can be observed that more intact proteins from colostrum (29%) can be found after gastric digestion compared to mature milk (18%) (Table 1). This effect is not expected to be related to the level of protease inhibitors, as no pepsin inhibitor has been found in human milk[5]. Large fragments of proteins and undigested proteins leaving the stomach may be relevant *in vivo*, as they may still be biologically active before they are further degraded in smaller peptides, amino acids, and finally absorbed.

In addition, still some undigested proteins can be found after intestinal digestion (Table 1). The total protein content after intestinal digestion was similar: with 0.3% and 0.4% remaining for colostrum and mature milk, respectively (Table 1). Protease inhibitors were not completely digested during intestinal digestion, with 1% and 3% remaining for colostrum and mature milk, respectively (Table 1). In addition, it can be observed in Figure 1, that protein digestion varies between mothers, with no indications that higher total protein content in colostrum than mature milk influences protein digestion. The degradation of human milk proteins during gastric and intestinal digestion could not be explained by the higher levels of protease inhibitors and total protein content in colostrum than mature milk, as the r-squared values were ranging between 0–0.3 (Figure S2), and as seen above (Figure 1 for total protein).

The enzymatic hydrolysis of individual human milk proteins in an infant in vitro model after gastric digestion

It can also be observed in Figure 2 that some human milk proteins remained partially intact after *in vitro* infant intestinal digestion. Table 2 is based upon the 15 most abundant proteins in Chinese human milk. The other human milk proteins still present intact after intestinal digestion can be found in Table S1, but not in all cases for both lactation periods.

The results showed that both serum proteins and caseins were not completely digested during gastric digestion: with 33% and 35% remaining for colostrum and 37% and 13% remaining for mature milk, respectively (Table 2). Interestingly, the levels of undigested caseins after gastric digestion were higher for colostrum than for mature milk (Table 2). The ratio between serum proteins and caseins in our study was ranging from 33:67 to 38:62 over lactation, respectively (data not shown). The findings in Table 2 might be attributed to the

hypothesis that lower levels of caseins in mature milk become easier to digest in the gastric phase than in colostrum.

Table 2. The top 15 most abundant Chinese human milk proteins after *in vitro* infant (0–3 months) digestion, categorized per function (N = 7). The data has been normalized to 100% per protein.

		Colostrum (week 1)		**Mature milk (week 4)**	
Protein function	Name of proteins	Gastric (%)	Intestinal (%)	Gastric (%)	Intestinal (%)
Enzymes	α-lactalbumin	22	0	20	0
	Bile salt-activated lipase	55	3	21	1
Immunity	Lactoferrin	33	1	12	3
	Ig α_1-chain c-region	12	2	84	2
	Ig λ_2-chain c-region	16	6	46	6
	Ig κ-chain c-region	9	8	64	12
	PIGR	39	0	61	0
	Clusterin	46	0	43	0
	Osteopontin	36	0	14	0
	β2-microglobulin	24	0	4	0
Protease inhibitor	α_1-antichymotrypsin	53	1	40	1
Transport	Serum albumin	45	2	32	3
	β-casein	29	0	15	0
	α_{S1}-casein	39	0	2	0
	κ-casein	37	0	23	0
Based on top 15	**Total serum proteins**	33	2	37	2
	Total caseins	35	0	13	0

The reason for the high level of gastric casein digestion (Table 2), but still partial survival of caseins, may be due to the low pepsin activity. Pepsin exerts its maximum activity at a pH of 2[23]. Another reason for this phenomenon may be the curd forming properties of casein, which happened at pH 5[3,17,31]. These "curds" are difficult to digest, which might result in a lower gastric casein digestion[11]. Caseins can be degraded more easily by enzymes during intestinal digestion due to their flexible non-compact structure and no curd forming at the pH 7[6-8]. As mentioned, there is also survival of many serum proteins in the gastric phase (Table 2). For example, lactoferrin can appear in the iron rich form, hololactoferrin, and the iron free form, apolactoferrin. This iron rich form of lactoferrin might stabilize the protein structure above pH 4, making hololactoferrin more resistant against enzymatic degradation in an infant *in vitro* model during gastric digestion. Previous studies have confirmed that lactoferrin is resistant against digestion in the infant's gastrointestinal tract[17,24-25], although

using different *in vitro* digestion models. Another example, α-lactalbumin in colostrum and mature milk is not fully digested (remaining 22% and 20%) (Table 2), respectively, which might be due to the fact that the gastric pH is too high to form the molten globule structure[3]. On lowering the pH to 3 in another *in vitro* model, the acidic side chains are protonated and α-lactalbumin adopts the molten globule state, which is a less compact conformation that is easier to digest by pepsin[3]. With regard to immunoglobulins, the heavy and light chains of the different immunoglobulins are connected via disulphide bridges in human milk, making them more resistant against digestive enzymes[18], which might explain their low gastric digestion (Table 2). The reason for the large differences in rate of digestion between immunoglobulins from colostrum and mature milk remains unclear.

The survival of individual human milk proteins after infant in vitro intestinal digestion

After gastric digestion, the human milk proteins where further hydrolyzed by intestinal enzymes. It can be observed in Table 2 that some of the highly abundant serum proteins (lactoferrin, bile salt-activated lipase, immunoglobulins, α_1-antichymotrypsin, serum albumin) from both colostrum and mature milk were still partially present after intestinal digestion (range: 1–12%) (Table 2). These specific serum proteins were always present in all the individual intestinal digesta samples (data not shown). Other abundant serum proteins, α-lactalbumin, polymeric immunoglobulin receptor (PIGR), clusterin, osteopontin, β_2-microglobulin, and the 3 caseins (β-, α_{S1}- and κ-casein) from both colostrum and mature milk were almost completely digested after *in vitro* intestinal digestion (Table 2). These proteins were absent in the intestinal digesta samples of most (N = 6) mothers (data not shown). Thirty-seven other human milk proteins were found after intestinal digestion (Table S1), although, the survival levels of these proteins varied between colostrum and mature milk during digestion. Among these 37 serum proteins, several low abundant immunoglobulins can be found after intestinal digestion (Table S1), but also e.g. lysozyme, α_1-antitrypsin, and fatty acid-binding protein.

Lactoferrin, lysozyme, immunoglobulins, antichymotrypsin, α_1-antitrypsin, and serum albumin were previously found in the infant's faeces[16], even surviving fermentation. The proportion of intact human milk proteins found in the faeces varied with the age of the infants, and about 10% of the total protein intake of the breastfeed infants was undigested in the faeces during the early neonatal period up to 1 month of age, while only 3% was found at 4 months of age[16]. The findings in Figure 1 and Table 1 are in the same range[16]. From the current study, it became clear that overall more than 40 serum proteins, including several immune-active proteins (e.g. lactoferrin, immunoglobulins) and protease inhibitors (e.g. α_1-antichymotrypsin and α_1-antitrypsin), were still partially present intact after *in vitro* intestinal digestion, and these proteins might be involved in supporting the infant's digestive tract against pathogens[16,24].

Undigested immunoglobulins after gastric and intestinal digestion might be important for infants in the first months of life, and might provide additional protection when the infant's immune system and digestive tract is not fully developed[1]. This might also account for the other immune-active proteins, which survive gastric digestion, as they might still be biologically active before being further degraded during intestinal digestion. PIGR, clusterin, osteopontin, and β_2-microglobulin are highly abundant in human milk and exert important functions for the development of the infant's immune system[5]. Bile salt-activated lipase and serum albumin, which are highly abundant in human milk, were also found to be resistant (relative levels ranging from 3 to 1%) against gastrointestinal enzymes from both colostrum mature milk, and able to survive intestinal digestion in an *in vitro* infant model.

Conclusions

This study provided, for the first time, detailed information on the digestion of proteins in a newly developed infant (0–3 months) *in vitro* digestion model using both colostrum and mature milk of 7 individual Chinese mothers. LC-MS/MS was used to provide a more complete overview of the human milk protein composition before, during and after *in vitro* infant digestion. Protein digestion vary for the milk from the mothers individually. Large variation in total undigested protein was found between mothers after gastric digestion. Colostrum and mature milk were digested after intestinal digestion to a similar extent. In contrast to expections, the extent of protein degradation was not directly influenced by protease inhibitors and the total protein content. Caseins were more digested after digestion than most serum proteins. The relative levels of the immune-active serum proteins were overall the highest after intestinal digestion. The resulting intact immune-active proteins, like antibacterial proteins, might supports the infant's intestine against pathogens.

References

1. Lönnerdal, B. Nutritional and physiologic significance of human milk proteins. Am. J. Clin. Nutr. **2003**, 77, 1537–43.
2. Rudloff, S.; Kunz, C. Protein and non-protein nitrogen components in human milk, bovine milk, and infant formula: Quantitative and qualitative aspects in infant nutrition. J. Pediatr. Gastroenterol. Nutr. **1997**, 24, 328–44.
3. Lönnerdal B. Human milk proteins: Key components for the biological activity of human milk. Adv. Exp. Med. Biol. **2004**, 554, 11–25.
4. Zhang, L.; Van Dijk, D.; Hettinga, K.A. An interactomics overview of the human and bovine milk proteome over lactation. Proteome Sci. **2017**, 15, 1–14.
5. Elwakiel, M.; Boeren, S.; Hageman, J.A.; Szeto, I.M.; Schols, H.A.; Hettinga, K.A. Variability of serum proteins in Chinese and Dutch human milk during lactation. Nutrients, **2018**, 11, 499–513.

6. Chatterton, D.; Rasmussen, J.; Heegaard, C.; Sørensen, E.; Petersen, T. *In vitro* digestion of novel milk protein ingredients for use in infant formulas: Research on biological functions. Trends Food Sci. Technol. **2004**, 15, 373–83.
7. Ross, R.; Fitzgerald, G.; Stanton, C. Unraveling the digestion of milk proteins. Am. J. Clin. Nutr. **2013**, 97, 1161–2.
8. Holton, T.; Vijayakumar, V.; Dallas, D.; Guerrero, A.; Borghese, R.; Lebrilla, C.; German, B.; Barile, D.; Underwood, M.; Shields, D.; Khaldi, N. Following the digestion of milk proteins from mother to baby. J. Proteome. Res. **2014**, 13, 5777–83.
9. Picariello, G.; Ferranti, P.; Fierro, O.; Mamone, G.; Caira, S. Peptides surviving the simulated gastrointestinal digestion of milk proteins: Biological and toxicological implications. J. Chromatogr. B, **2010**, 878, 295–308.
10. Dallas, D.; Guerrero, A.; Khaldi, N.; Borghese, R.; Bhandari, A.; Underwood, M.; Lebrilla, C.; German, B.; Barlie, D. A peptidomic analysis of human milk digestion in the infant stomach reveals protein-specific degradation patterns. J. Nutr. **2014**, 144, 815–20.
11. Dupont, D.; Tome, D. Milk proteins: Digestion and absorption in the gastrointestinal tract. In milk proteins; Boland, M., Singh, H., Thompson, A., Eds.; Academic Press: London, **2014**, pp. 557–69.
12. Dingess, K.A.; de Waard, M.; Boeren, S.; Vervoort, J.; Lambers, T.T.; van Goudoever, J.B.; Hettinga, K.A. Human milk peptides differentiate between the preterm and term infant and across varying lactational stages. Food Funct. **2017**, 8, 3769–82.
13. Su, M.; Broadhurst, M.; Liu, C.; Gathercole, J.; Cheng, W.; Qi, X. Comparative analysis of human milk and infant formula derived peptides following *in vitro* digestion. Food Chem. **2017**, 221, 1895–903.
14. Lönnerdal, B. Bioactive proteins in human milk: Health, nutrition, and implications for infant formulas. J. Pediatr. **2016**, 173, 4–9.
15. Dallas, D.; Underwood, M.; Zivkovic, A.; German, J. Digestion of protein in premature and term infants. J. Nutr. Disord. Ther. **2012**, 2, 112–33.
16. Davidson, L.; Lönnerdal B. Persistence of human milk proteins in the breastfed infant. Acta Pædiatr. **1987**, 76, 733–40.
17. De Oliveira, S.; Bellanger, A.; Ménard, O.; Pladys, P.; Le Gouar, Y.; Dirson, E.; Kroell, F.; Dupont, D.; Deglaire, A.; Bourlieu, C. Impact of human milk pasteurization on gastric digestion in preterm infants: A randomized controlled trial. Am. J. Clin. Nutr. **2017**, 105, 379–90.
18. Gan, J.; Bornhorst, G.; Henrick, B.; German, B. Protein digestion of baby foods: Study approaches and implications for infant health. Mol. Nutr. Food Res. **2018**, 62, 1–25.
19. Lindberg, T.; Ohlsson, K.; Westrom, B. Protease inhibitors and their relation to proteases activity in human milk. Pediatr. Res. **1982**, 16, 479–83.
20. Chowanadisai, W.; Lönnerdal, B. α_1-Antitrypsin and antichymotrypsin in human milk: Origin, concentrations, and stability. Am. J. Clin. Nutr. **2002**, 76, 828–33.
21. Ehlers, M. Immune-modulating effects of α_1-antitrypsin. Biol. Chem. **2014**, 395, 1187–93.
22. Dallas, D.; Murray, N.; Gan, J. Proteolytic systems in milk: Perspectives on the evolutionary function within the mammary gland and the infant. J. Mammary Gland Biol. Neoplasia, **2015**, 20, 133–47.

23. Minekus, M; Alminger, M.; Alvito, P.; Balance, S.; Bohn, T.; Bourlieu, C.; Carriere, F.; Boutrou, R.; Corredig, M.; Dupont, D.; Dufour, C.; Egger, L.; Golding, M.; Karakaya, S.; Kirkhus, B.; Le Feunteun, S.; Lesmes, U.; Macierzanka, A.; Mackie, A.; Marze, S.; McClements, D.; Menard, O.; Recio, I.; Santos, C.; Singh, R.; Vegarud, G.; Wickham, M.; Weitschies, W.; Brodkorb, A. A standardized static *in vitro* digestion method suitable for food – An international consensus. Food Funct. **2014**, 5, 1113–24.
24. Inglingstad, A.; Devold, T.; Eriksen, E.; Holm, H.; Jacobsen, M.; Liland, K.; Rukke, E.; Vegarud, G. Comparison of the digestion of caseins and whey proteins in equine, bovine, caprine and human milks by human gastrointestinal enzymes. Dairy Sci. Technol. **2010**, 90, 549–63.
25. Lindberg, T.; Engberg, S.; Sjoberg, L.; Lönnerdal, B. *In vitro* digestion of proteins in human milk fortifiers and in preterm formula. J. Pediatr. Gastroenterol. Nutr. **1998**, 27, 30–6.
26. Dipalma, J.; Kirk, C.; Hamosh, M.; Colon, A.; Benjamin, S.; Hamosh, P. Lipase and pepsin activity in the gastric mucosa of infants, children, and adults. Gastroenterol. **1991**, 101, 116–21.
27. Dall'Asta, C.; Florio, P.; Lammardo, A.; Prandi, B.; Mazzeo, T.; Budelli, A.; Pellegrini, N. Development of an *In vitro* digestive model for studying the peptide profile of breast milk. Int. J. Food. Sci. Nutr. **2015**, 66, 409–15.
28. Hamosh, M. Digestion in the newborn. Clin. Perinatol. **1996**, 23, 191–209.
29. Sondheimer, J.; Clark, D.; Gervaise, E. Continuous gastric pH measurement in young and older healthy preterm infants receiving formula and clear liquid feedings. J. Pediatr. Gastroenterol. Nutr. **1985**, 4, 352–5.
30. Hur, S.; Ou Lim, B.; Decker, E.; McClements, J. *In vitro* human digestion models for food applications. Food Chem. **2011**, 125, 1–12.
31. Dupont, D.; Mandalari, G.; Molle, D.; Jardin, J.; Leonil, J.; Faulks, R.; Wickham, M.; Mills, E.; Mackie, R. Comparative resistance of food proteins to adult and infant *in vitro* digestion models. Mol. Nutr. Food Res. **2010**, 54, 767–80.
32. Bourlieu, C.; Menard, O.; Bouzerzour, K.; Mandalari, G.; Macierzanka, A.; Mackie, A.; Dupont, D. Specificity of infant digestive conditions: Some clues for developing relevant *in vitro* models. Crit. Rev. Food Sci. Nutr. **2014**, 54, 1427–57.
33. Hettinga, K.A.; Van Valenberg, H.; De Vries, S.; Boeren, S.; Van Hooijdonk, T.; Van Arendonk, J.; Vervoort, J. The host defense proteome of human and bovine milk. PLoS One, **2011**, 6, 1–8.
34. Cox, J.; Mann, M. MaxQuant enables high peptide identification rates, individualized ppb-range mass accuracies and proteome-wide protein quantification. Nature, **2008**, 26, 1367–72.
35. Schwanhausser, B.; Busse, D.; Li, N.; Dittmar, G.; Schuchhardt, J.; Wolf, J.; Chen, W.; Selbach, M. Global quantification of mammalian gene expression control. Nature, **2011**, 473, 337–42.

Supporting information

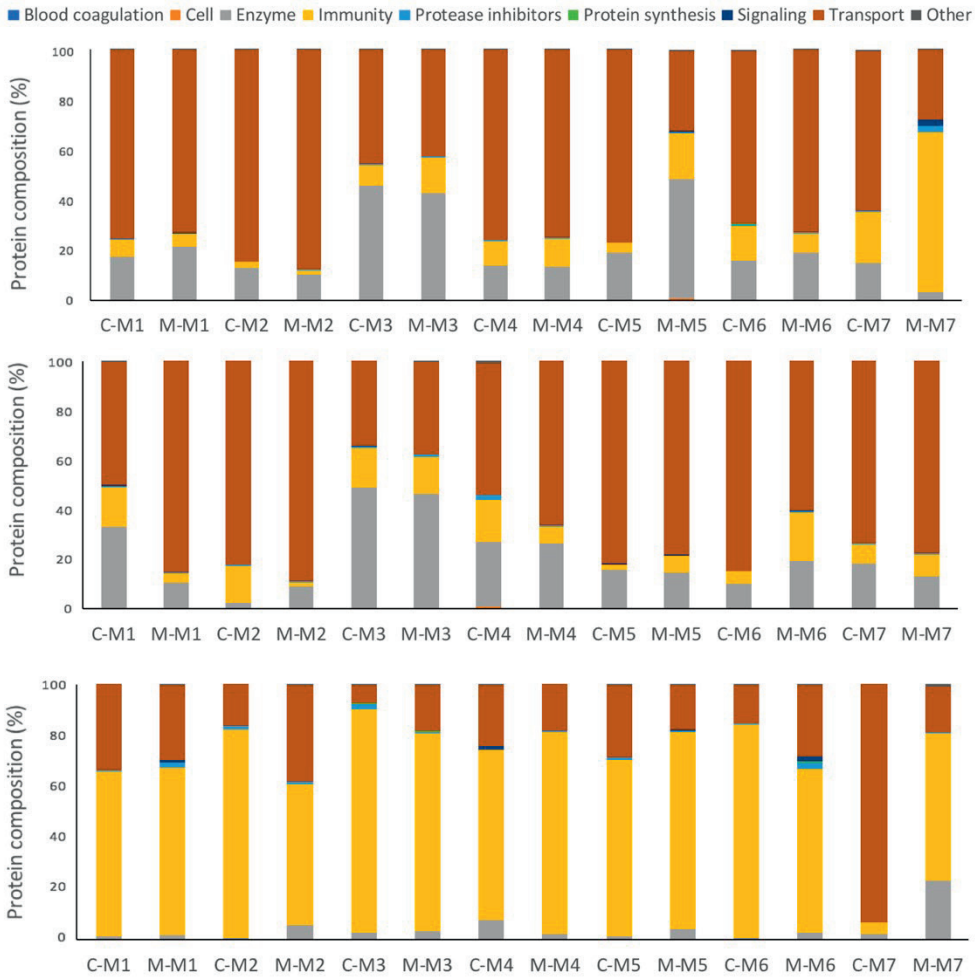


Figure S1. Classes of proteins in colostrum (week 1) and mature milk (week 4) of 7 Chinese mothers **(A)**, and in an *in vitro* infant (0–3 months) digestion model for the gastric phase **(B)** and intestinal phase **(C)**, with proteins grouped having similar functions. C and M stands for colostrum and mature milk and the number behind the hyphen indicates the mother. The composition in the gastric and intestinal phase were based on total remaining proteins set on 100%.

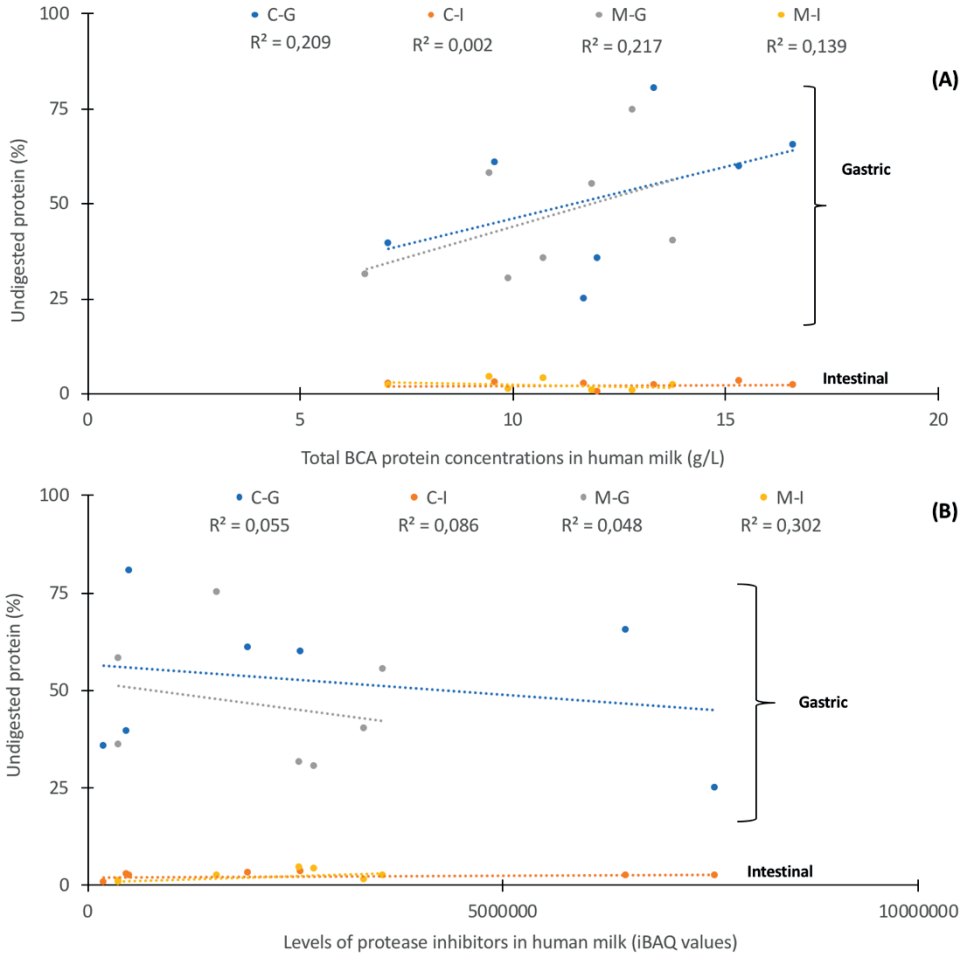


Figure S2. The initial total protein content (based on the BCA protein assay) **(A)** and levels of protease inhibitors (iBAQ values) **(B)** in milk of 7 Chinese mothers from 2 different lactational periods (week 1 and 4) plotted versus undigested protein to gastric and to intestinal digestion. A trendline with r close to zero indicates no relationship between the two variables. The different dots represent the different samples of the mothers per lactation stage.

Table S1. List of undigested proteins found after *in vitro* gastric (G) and intestinal digestion (I) using colostrum (C) and mature milk (M) of 7 Chinese mothers. The data has been normalized to a scale of 100% per protein. The most abundant human milk proteins from Table 3 are here excluded.

Protein names	Function	C-G	M-G	C-I	M-I
Actin, cytoplasmic 1	Cell	39	14	0	1
Erythrocyte band 7 integral membrane protein	Cell	20	28	1	2
Galectin-3-binding protein	Cell	98	3	0	2
α-enolase	Enzyme	46	39	1	0
Annexin	Immunity	19	32	4	3
Complement C3	Immunity	20	73	1	2
Complement C4A	Immunity	46	17	1	0
Ig α_2-chain c-region	Immunity	4	13	0	1
Ig γ_1-chain c-region	Immunity	27	10	3	12
Ig γ_2-chain c-region	Immunity	17	7	3	1
Ig γ_3-chain c-region	Immunity	15	4	1	2
Ig κ-chain V-II region TEW	Immunity	7	32	1	3
Ig λ_2-chain V-III region SH	Immunity	9	52	1	2
Immunoglobulin α-like polypeptide 5	Immunity	9	53	5	4
Lactadherin	Immunity	34	37	1	1
Lysozyme C	Immunity	41	16	1	1
Neutrophil defensin 3	Immunity	18	29	0	1
Xanthine dehydrogenase	Immunity	42	53	1	3
Zinc α_2-glycoprotein	Immunity	20	51	7	11
Heat shock protein HSP 90-alpha	Other	21	18	0	3
Heat shock protein HSP 90-alpha	Other	21	18	0	3
Hornerin	Other	70	33	24	5
Glyceraldehyde-3-phosphate dehydrogenase	Other	53	21	5	3
Macrophage mannose receptor 1	Other	69	17	1	0
Zinc finger protein 337	Other	53	6	0	1
Putative elongation factor 1 α-like 3	Protein synthesis	37	29	0	1
α1-antitrypsin	Protease inhibitor	35	44	1	2
Cystatin A	Protease inhibitor	72	50	30	9

Table S1 (continued)

Protein names	Function	C-G	M-G	C-I	M-I
14-3-3 protein ζ/ δ	Signaling	26	9	1	1
Guanine nucleotide-binding protein G subunit γ-12	Signaling	31	23	5	4
Ubiquitin-60S ribosomal protein L40	Signaling	56	13	0	16
Suprabasin	Signaling	40	36	15	5
ADP-ribosylation factor 1	Transport	41	30	12	0
Apolipoprotein B-100	Transport	56	32	24	18
Ceruloplasmin	Transport	25	36	6	0
Epididymal secretory protein E1	Transport	25	26	5	9
Fatty acid-binding protein	Transport	41	1	6	2
Serotransferrin	Transport	91	12	1	0

Chapter 4

Serum protein *N*-glycans in colostrum and mature milk of Chinese mothers

Manuscript submitted for publication:
Elwakiel, M.; Bakx, E.J.; Szeto, I.M.; Li, Y.; Hettinga, K.A.; Schols, H.A.

Abstract

To study the Chinese human milk serum glycoproteome over lactation, *N*-glycans were released and separated from serum proteins, purified by solid phase extraction, and analysed by MALDI-TOF-MS. In total, 66 different serum protein *N*-glycans were found in colostrum (week 1) and mature milk (week 4) of 7 Chinese mothers. A clear difference was observed between milk of 5 secretor and 2 non-secretor mothers, based on the type and relative amounts of the individual *N*-glycans. Principal component analysis enable a further distinction between colostrum and mature milk of the 5 secretor mothers. Further analysis showed that the relative levels of the total neutral nonfucosylated and the fucosylated *N*-glycans in milk of 5 secretor mothers increased and decreased over lactation, respectively. This pattern could not be observed for the milk from the 2 non-secretor mothers. The relative levels of the total acidic *N*-glycans remained constant over lactation for both genetic groups. The individual *N*-glycan profiles in milk of the 5 secretor and 2 non-secretor mothers mainly differed with regard to their levels of individual neutral fucosylated *N*-glycans. Despite these differences, the total neutral fucosylated glycan level was approximately the same for both genetic groups. The relative levels of the total neutral (sum of nonfucosylated and fucosylated) *N*-glycans covered >90% of the total *N*-glycan content, for all 7 mothers. Overall, this was the first study that provided detailed information on individual serum protein *N*-glycans in milk among mothers and over time, and that fucosylation of the serum protein *N*-glycans in milk was associated with the mother's secretor status.

Keywords

Glycoproteins, oligosaccharides, glycosylation, intestinal mucosal barrier

Introduction

Human milk is the best nutrition for infants during the first 6 months of life[1], and stimulates the maturation of the infant's intestinal immune system[2]. Human milk contains many biofunctional components, such as proteins and human milk oligosaccharides (HMOs)[3]. HMOs represent complex lactose-based glycans synthesized in the mammary gland during lactation, which reach the colon intact, and are able to stimulate the development of the bifidogenic flora[4,5]. Human milk proteins among others play a pivotal role in protecting the infant's gut mucosa against pathogens[6]. It has been reported that 70% of the human milk proteins in number are glycosylated[7]. These protein-bound glycans among others might affect the folding and stability of proteins and modulate neonatal immunity by altering host epithelial and immune cell responses in the infant's gut[7-10].

Caseins are divided into three main types (α_{S1}-, β-, and κ-casein), and are mainly described as transport proteins[11]. The proteins α_{S1}- and β-casein in human milk do not have any glycosylated amino acid residues, while κ-casein has multiple *O*-glycosylation sites at various threonine and serine residues[12]. Serum proteins such as lactoferrin, immunoglobulins, serum albumin, and α-lactalbumin form the main portion of the glycoproteins present in human milk, and mainly contain *N*-glycans[12]. These glycans are attached to the amide nitrogen of an asparagine residue of the protein[13-16]. There are also serum proteins that only contain *O*-glycans (e.g. osteopontin) or contain both *O*-glycans and *N*-glycans (e.g. bile salt-activated lipase). Osteopontin, lactoferrin and immunoglobulins have direct bactericidal properties[12]. In addition, many other glycosylated serum proteins can be found in human milk having a reported immune activity[12].

N-glycans are composed of five types of neutral monosaccharides; fucose (Fuc), galactose (Gal), mannose (Man), *N*-acetylglucosamine (GlcNAc), *N*-acetylgalactosamine (GalNAc), and the sialic acidic structure *N*-acetylneuraminic acid (NeuAc)[14-19]. All identified *N*-glycans in human milk have a pentasaccharide as common core, consisting of three Man and two GlcNAc residues[14-16]. *N*-glycans can be classified into three types, namely high mannose, complex, and hybrid *N*-glycans[20-23]. High mannose *N*-glycans typically contains between 2 and 6 Man residues attached to the pentasaccharide core, whereas complex *N*-glycans can be elongated at the mannose residues by GlcNAc residues, which are often further decorated by Gal and GalNAc residues[20-23]. Hybrid *N*-glycans are composed of a pentasaccharide core, with one branch of mannose residues and another branch with one or two GlcNAc residues[20-23].

Complex and hybrid *N*-glycans have a large variety of structures, and fucosylation might affect their conformation and functional properties[19]. Up to date, 13 different fucosyltransferases (FUT) have been detected in the human genome[24]. The FUT8 gene encodes for the α1,6-fucosyltransferase that transfers a Fuc reside to the innermost GlcNAc unit of *N*-glycan chains[24,25]. Fucosylation of *N*-glycans also might partially depend on the

FUT2 and FUT3 genes[24,25], which determine the mother's secretor (Se) and Lewis (Le) status. The FUT2 Se gene determines the presence of α1,2-linked fucosylated glycans[26]. The addition of Fuc residues by an α1,3/4-linkage on the antennae of GlcNAc might be regulated by FUT3 or other α1,3-genes (FUT4, 5, 6, 7 and 9)[26]. The fucosylated glycotypes on serum glycoproteins in milk from 43 healthy mothers were analysed semi-quantitatively by lectin-blotting, where three specific biotinylated lectins were able to recognize and differentiate between the α1,2-, α1,3- and α1,6-Fuc linkages[27].

The type and levels of *N*-glycans have been previously investigated in mature milk (week 12–16) of three mothers[14]. From the total 52 *N*-glycans identified, 34 (65%) *N*-glycans were fucosylated[14]. The relative levels of the fucosylated *N*-glycans covered >80% of the total *N*-glycan content[14]. However, the type and level of individual *N*-glycans in milk of the individual mothers have not yet been investigated[14,27]. More recently, the variation of *N*-glycans in human milk over lactation has been studied[16]. In this latter study, milk from 10 mothers was collected and combined per lactation stage (colostrum, 3 days; transition milk, 9 days; and mature milk, 40 days)[16]. It was reported that levels of fucosylated *N*-glycans dropped from circa 61% in colostrum to 37% in transition milk, and then remained constant in mature milk[16]. This current study set out to investigate the individual differences in type and levels of *N*-glycans in milk between mothers. The main objective was to profile and compare *N*-glycans in colostrum (week 1) and mature milk (week 4) of 7 Chinese mothers differing in secretor status using matrix assisted laser desorption ionization time of flight mass spectrometry (MALDI-TOF-MS).

Material and methods

Chemicals

Ammonium bicarbonate, sodium chloride, acetonitrile (MeCN), trifluoroacetic acid (TFA), and ethanol (EtOH) were purchased from Biosolve B.V. (Valkenswaard, NL). Water was filtered using the Milli-Q water purification system of Merck Millipore (Molsheim, FR). Sodium dodecyl sulfate (SDS), 2-mercaptoethanol (2-BME), 2.5 dihydroxybenzoic acid, and branched octylphenoxy poly(ethoxy)ethanol (IGEPAL CA-630) were provided by Sigma-Aldrich (St. Louis, U.S.). A mixture of maltodextrins was obtained from Avebe (Veendam, NL). The complex *N*-glycans NA_2 (No. 24) and NA_2F (No. 33), and the high mannose structures Man6 (No. 11) and Man8 (No. 27) (Table 1) were provided by Ludger (Oxfordshire, UK). Peptidyl *N*-glycosidase F (PNGase F) was bought from Asparia Glycomics (San Sebastian, ES).

Human milk collection and the mother's SeLe status

Milk of 7 Chinese mothers from 2 different lactation periods (colostrum, week 1; mature milk, week 4) was collected, as described previously[11]. Human milk collection was approved

by the Chinese Ethics Committee of Registering Clinical Trials (ChiECRCT-20150017). Written informed consent was obtained from all mothers. The milk was categorized based on the mother's SeLe group, as described previously (M1 = 26, M2 = 20, M3 = 22, M4 = 25, M5 = 23, M6 = 21, M7 = 24)[28]. Mother 1 and 4 belong to the the $Se^{-}Le^{+}$ group, whereas mother 2, 3, 5, 6 and 7 were assigned to the $Se^{+}Le^{+}$ group. Milk samples from $Se^{+}Le^{-}$ and $Se^{-}Le^{-}$ mothers are not represented in this study.

Isolation of human milk serum proteins

The lipid layer was removed from the human milk samples (7 mL) after centrifugation (10 min, 1,500 g, 4 °C), and the obtained skim milk was transferred to ultracentrifuge tubes[11]. After ultracentrifugation (90 min, 100,000 g, 4 °C), the top layer represented the remaining milk fat still present, the middle layer was milk serum (consisting of serum proteins and free oligosaccharides), and the bottom layer consisted of micellar casein[11]. Serum proteins were separated from the HMOs via EtOH precipication[15], with modifications. Milk serum (3 mL) was diluted twice, and then absolute EtOH was added till a relative concentration of 67% EtOH was reached. After 67% EtOH precipitation (60 min, 4 °C) and centrifugation (15 min, 1,500 g, 4 °C), the supernatant containing HMOs was discarded. The pellet containing serum proteins was re-dissolved in 0.5 mL of water and EtOH precipitation was repeated three times. Finally, serum proteins in the pellet were re-dissolved in 0.5 mL of 200 mM ammonium bicarbonate (pH 8) using alternately a vortex and an ultrasonic bath at room temperature. The final experiments (after method optimization and validation) were done in duplicate.

The release and purification of N-glycans from serum proteins

Methods were based upon previously described methods[14-15,29-32], with modifications. Briefly, 1 µL of 1M SDS in water and 10 µL of 2-BME were combined with 100 µL of the solution containing the human milk serum proteins, and kept 10 min at 95 °C. After cooling down to 37 °C, solutions with the denatured serum proteins were diluted with 50 µL 100 mM ammonium bicarbonate, and mixed with 50 µL of 4% (v/v) IGEPAL CA-630. A wide range of concentrations (range: 50–200 mM) have been used before starting incubation with PNGase F[15,29,32], and here the final concentration of the samples was 100 mM. In order to protect the PNGase F from denaturation by SDS, IGEPAL CA-630 as a non-ionic detergent was added, although the mechanism behind this protection effect is still unknown[29]. For the complete release of *N*-glycans from human milk serum proteins, the mixture was incubated with PNGase F (24 h, 37 °C), 6 µL of enzyme at t = 0 followed by 4 µL of enzyme after 16 h. After incubation, the mixtures containing *N*-glycans and deglycosylated proteins were mixed with absolute EtOH till a relative concentration of 67% EtOH was reached, and stored for 60 min at 4 °C. After centrifugation (15 min, 1,500 g, 4 °C), the supernatant was dried under a stream of air overnight, and the *N*-glycans thereafter reconstituted with 0.5 mL of water.

The *N*-glycans in solution were further purified by solid phase extraction using a graphitized carbon column cartridge (bed weight; 150 mg, tube size; 4 mL, Alltech, Deerfield, U.S.)[14]. The cartridge was prepared with 2 mL of water, followed by 2 mL of 80% MeCN containing 0.1% TFA. The cartridge was conditioned with 2 mL of water before loading 0.5 mL of the sample with *N*-glycans. The *N*-glycans on the cartridge were eluted with 0.5 mL of 10% MeCN, 0.5 mL of 20% MeCN, and 0.5 mL of 40% MeCN in water containing 0.05% TFA. The *N*-glycan mixtures were dried under a stream of air overnight. After reconstitution in 20 µL of water, the solution containing *N*-glycans was ready for MALDI-TOF-MS analysis.

Analysis of N-glycans by MALDI-TOF-MS

Analysis of *N*-glycans by MALDI-TOF-MS was done, as described previously[33]. MALDI-TOF mass spectra were recorded using an UltraFlextreme workstation controlled by FlexControl 3.3 software (Bruker Daltronics, Bremen, DE) equipped with a Smartbeam II laser of 355 nm and operated in positive mode. Spectra were collected from 1500 laser shots with an energy level of 30%. The spectrometer was calibrated using a mixture of maltodextrins in a mass range of 500 to 3000 Da. The complex *N*-glycans NA_2 (No. 24) and NA_2F (No. 33), and the high mannose structures Man6 (No. 11) and Man8 (No. 27) in solution were used as *N*-glycan standards. The numbers behind the four standards correspond with the numbers in Table 1. The matrix solution was prepared by mixing 25 mg of 2.5 dihydroxybenzoic acid in 1 mL 50% MeCN/ 50% water (containing 1 mM of sodium chloride), and subsequent centrifugation (5 min, 1,500 g, 4 °C). For each sample containing *N*-glycans, 1 µL was added directly on the ground steel MS target plate (Bruker Daltonics, DE), followed by 1 µL of the matrix solution, and dried under a stream of air.

Data analysis of *N*-glycans was performed with Flex Analysis 3.3 (Bruker Daltronics, DE). Peak intensities of the individual *N*-glycans were used if the peak height of the *N*-glycans was 3 times higher than background noise. For data normalization, the MALDI-TOF-MS peak intensities for each *N*-glycan were transformed into percentages, by relating the peak intensity of each *N*-glycan in a sample to the total signal intensity of all the identified *N*-glycans within a sample. The data of the individual *N*-glycans for 2 biological replicates were averaged. The structures of the *N*-glycans were assigned via the online database GlyTouCan using their molecular mass[34]. No distinction could be made between isomers. For each structure, just one possible isomer was selected for visualization. Interpretation of the *N*-glycan profiles in human milk was facilitated by principal component analysis (PCA) and heatmaps using R (Lucent Technologies, New York, U.S.). The relative levels of the individual *N*-glycans in milk per mother and per lactation stage were used. Information about the mother's SeLe status was omitted from the dataset during statistical analysis.

Results and discussion

Analysis of N-glycans by MALDI-TOF-MS

In order to analyse all *N*-glycans present in milk, colostrum (week 1) and mature milk (week 4) of 7 Chinese mothers were analysed. Milk serum was isolated by ultracentrifugation, after which the serum proteins were separated from the HMOs. Serum proteins were denatured, incubated with the enzyme PNGase F for 24 h, and the released *N*-glycans purified by solid phase extraction and analysed by MALDI-TOF-MS. Different methods for the release and isolation of *N*-glycans were tested, using and combining various methods from literature[14-15,29-32]. For example, the removal of HMOs with the right concentration of EtOH was crucial to get the optimal signal to noise ratio by Maldi-TOF-MS and to identify more individual *N*-glycans. The best method, after optimization and validation, can be found in the material and method section. An example of a MALDI-TOF mass spectrum for colostrum (week 1) from Chinese mother 4 can be found in Figure 1, highlighting the 15 most abundant *N*-glycans.The structures of the different *N*-glycans numbered in Figure 1 can be found in Figure 2, which were assigned via the online database GlyTouCan[34]. The top 15 *N*-glycans have a pentasaccharide as common core, consisting of three Man and two GlcNAc residues (Figure 1). More than half of the top 15 *N*-glycans contained a Fuc residue, and none of them contained a NeuAc residue (Figure 1). No distinction could be made between the different isomers by MALDI-TOF-MS. The moleculair mass of the different *N*-glycans numbered in Figure 1 can be found in Table 1. However, not all the 66 different *N*-glycans, as summarized in Table 1, were found in colostrum of each individual Chinese mother.

Overview of identified N-glycans in colostrum (week 1) and mature milk (week 4) of 7 Chinese mothers

An overview of all identified *N*-glycans in human milk can be found in Figure 2 and Table 1, combining the data obtained by MALDI-TOF-MS of the 7 mothers from 2 different lactation periods. In total, 66 different *N*-glycans were detected in human milk over time by MALDI-TOF-MS (Table 1), a higher number than previously reported in literature[14-16]. Of these the 66 *N*-glycans, 42 (64%) were found in all human milk samples: 48 (73%) and 43 (65%) unique structures were detected in colostrum and mature milk, respectively (Table S1). Among these 66 *N*-glycans, 42 (64%) structures can be classified as complex *N*-glycans, 5 (7%) as high mannose, 12 (18%) as hybrid (Table 1, classification system type I[15]), while 7 structures (11%) could not be classified in one of these 3 groups and are referred to in Table 1 as ''other''.

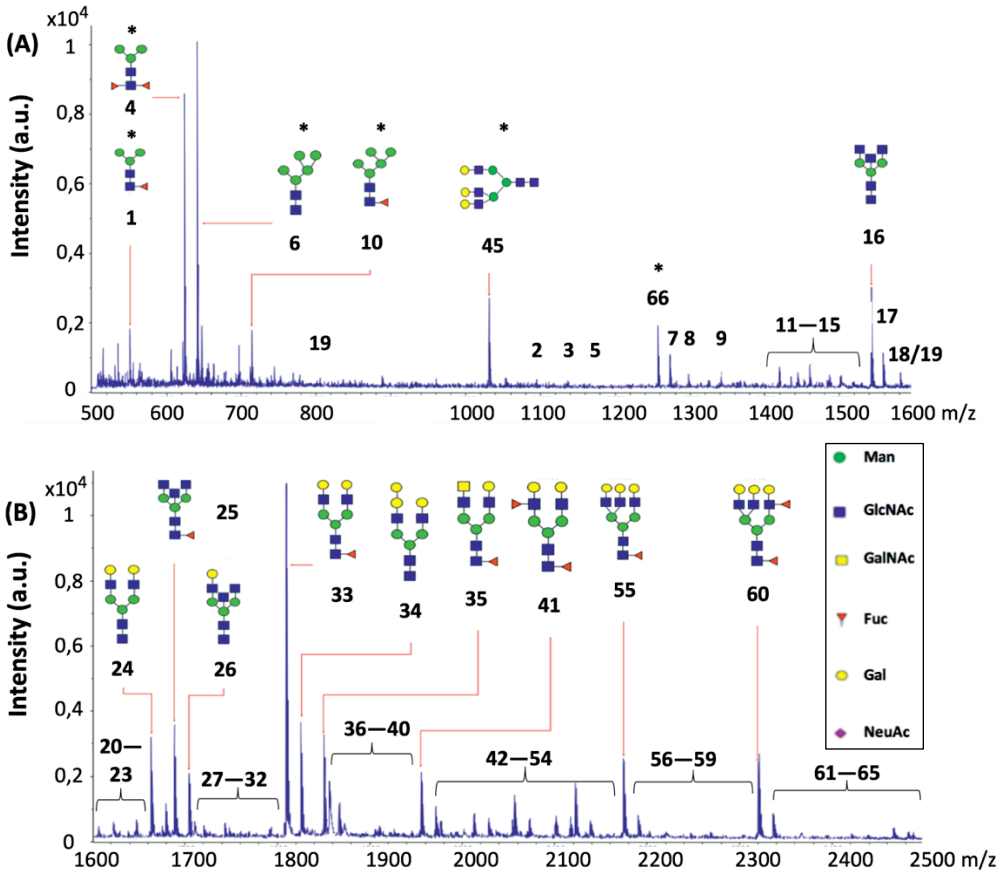


Figure 1. MALDI-TOF mass spectrum, highlighting the 15 most abundant *N*-glycans in colostrum (week 1) of Chinese mother 4. The numbers of the *N*-glycans correspond with the numbers in Table 1 and Figure 2. **A).** Spectrum with m/z ranging from 500 to 1600 and **B).** M/z ranging from 1600 to 2500. Just one possible isomer was selected for visualization. The *N*-glycans highlighted with an asterix were doubly charged.

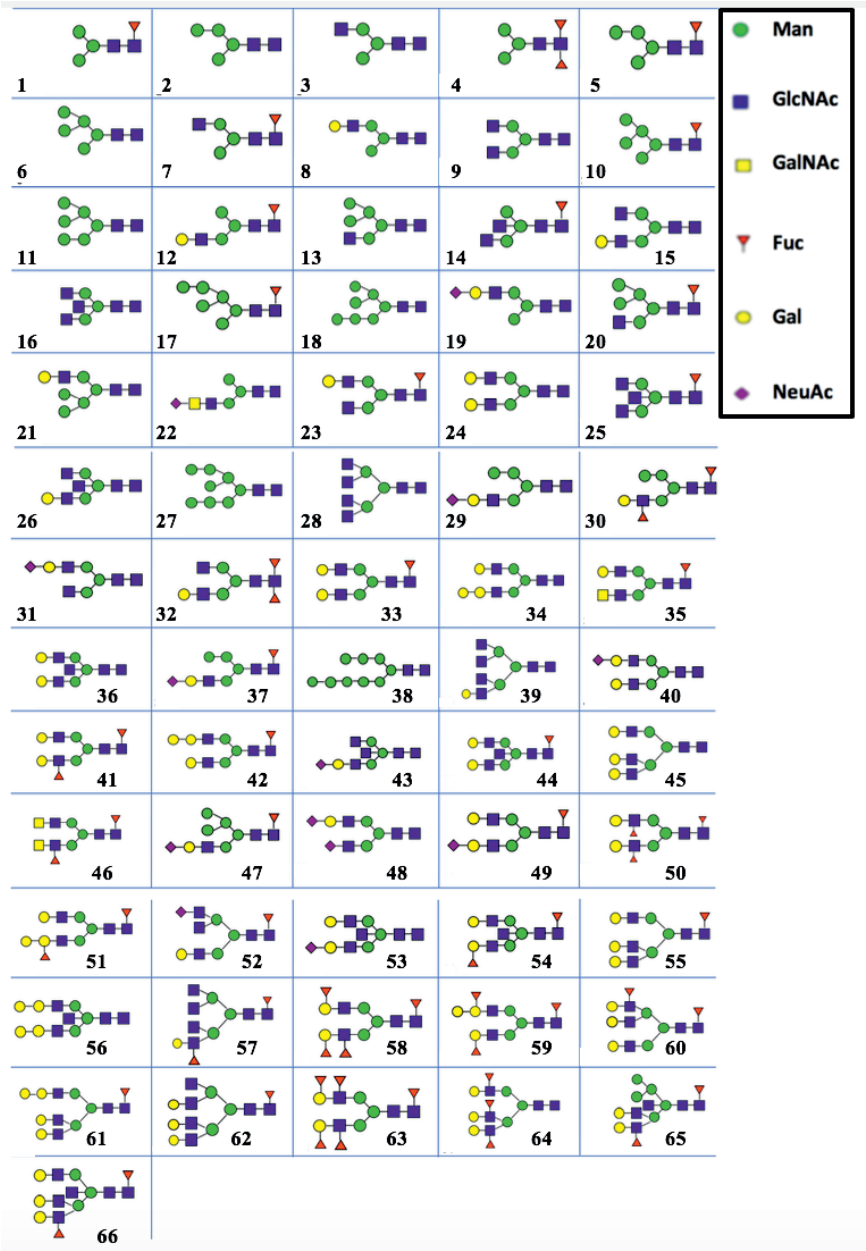


Figure 2. Overview of 66 *N*-glycans identified in colostrum (week 1) and mature milk (week 4) of 7 Chinese mothers using MALDI-TOF-MS. Numbers indicate the *N*-glycans displayed in Table 1. The structures of the identified *N*-glycans were assigned via the online GlyTouCan database based on their molecular mass[34]. Just one possible isomer is shown.

Table 1. *N*-glycans identified in colostrum (week 1) and mature milk (week 4) of 7 Chinese mothers, including sugar mass, building blocks, and using 2 classification systems.

No.	Mass	Composition				Classification: Type	
		Hex	HexNAc	Fuc	NeuAc	I *	II **
N1	1056,4 [15]	3	2	1	0	Other	NF
N2	1072,3 [15]	4	2	0	0	Other	N
N3	1113,3 [16]	3	3	0	0	Other	N
N4	1202,4	3	2	2	0	Other	NF
N5	1218,4 [15]	4	2	1	0	Other	NF
N6	1234,3 [14, 15, 16]	5	2	0	0	High mannose	N
N7	1259,3 [16]	3	3	1	0	Hybrid	NF
N8	1275,4 [14, 16]	4	3	0	0	Hybrid	N
N9	1316,3 [16]	3	4	0	0	Complex	N
N10	1380,6	5	2	1	0	Other	NF
N11	1396,3 [14, 15, 16]	6	2	0	0	High mannose	N
N12	1421,4 [14, 16]	4	3	1	0	Hybrid	NF
N13	1437,4 [15]	5	3	0	0	Hybrid	N
N14	1462,5 [14]	3	4	1	0	Complex	NF
N15	1478,4 [14, 15]	4	4	0	0	Complex	N
N16	1519,4 [14, 16]	3	5	0	0	Complex	N
N17	1542,4	6	2	1	0	Other	NF
N18	1558,4 [14, 15, 16]	7	2	0	0	High mannose	N
N19	1566,9 [14, 16]	4	3	0	1	Hybrid	A
N20	1583,4 [15]	5	3	1	0	Hybrid	NF
N21	1599,5 [14]	6	3	0	0	Hybrid	N
N22	1607,3	3	4	0	1	Hybrid	A
N23	1624,5 [14, 15]	4	4	1	0	Complex	NF
N24	1640,4 [14, 15]	5	4	0	0	Complex	N
N25	1665,5 [14, 16]	3	5	1	0	Complex	NF
N26	1681,5 [14]	4	5	0	0	Complex	N
N27	1720,4 [14, 15, 16]	8	2	0	0	High mannose	N
N28	1722,5	3	6	0	0	Complex	N
N29	1728,1 [15]	5	3	0	1	Hybrid	A
N30	1729,0 [15]	5	3	2	0	Hybrid	NF
N31	1769,5 [14]	4	4	0	1	Complex	A
N32	1770,8 [14]	4	4	2	0	Complex	NF
N33	1786,5 [14, 15, 16]	5	4	1	0	Complex	NF
N34	1802,5	6	4	0	0	Complex	N
N35	1827,5 [14, 16]	4	5	1	0	Complex	NF
N36	1843,5	5	5	0	0	Complex	N
N37	1874,4 [15]	5	3	1	1	Hybrid	AF
N38	1882,5 [14, 15, 16]	9	2	0	0	High mannose	N
N39	1884,5 [16]	4	6	0	0	Complex	N
N40	1931,9 [14, 15, 16]	5	4	0	1	Complex	A
N41	1932,9 [14, 15]	5	4	2	0	Complex	NF
N42	1948,5 [14,15]	6	4	1	0	Complex	NF

Table 1 (continued)

No.	Mass	Composition				Classification: Type	
		Hex	HexNAc	Fuc	NeuAc	I *	II **
N43	1972,7 [14]	4	5	0	1	Complex	A
N44	1989,5 [14, 16]	5	5	1	0	Complex	NF
N45	2005,6	6	5	0	0	Complex	N
N46	2014,5 [15]	3	6	2	0	Complex	NF
N47	2036,6	6	3	1	1	Complex	AF
N48	2060,9	4	4	0	2	Complex	A
N49	2077,5 [14, 15, 16]	5	4	1	1	Complex	AF
N50	2078,6 [14, 15]	5	4	3	0	Complex	NF
N51	2094,6 [15]	6	4	2	0	Complex	NF
N52	2118,7 [14, 15]	4	5	1	1	Complex	AF
N53	2134,7 [14]	5	5	0	1	Complex	A
N54	2135,5	5	5	2	0	Complex	NF
N55	2151,6 [14]	6	5	1	0	Complex	NF
N56	2167,6	7	5	0	0	Complex	N
N57	2176,7 [15]	4	6	2	0	Complex	NF
N58	2224,6 [14]	5	4	4	0	Complex	NF
N59	2240,7 [15]	6	4	3	0	Complex	NF
N60	2297,7 [14, 16]	6	5	2	0	Complex	NF
N61	2313,6	7	5	1	0	Complex	NF
N62	2354,8	6	6	1	0	Complex	NF
N63	2370,8	5	4	5	0	Complex	NF
N64	2443,8 [14, 16]	6	5	3	0	Complex	NF
N65	2459,8 [14]	7	5	2	0	Hybrid	NF
N66	2500,6 [15]	6	6	2	0	Complex	NF

* Type I classification: Complex, hybrid, high-mannose, and other *N*-glycans. ** Type II classification: N, neutral nonfucosylated *N*-glycans; NF, neutral fucosylated *N*-glycans; A, acidic nonfucosylated *N*-glycans; AF, acidic fucosylated *N*-glycans.

The ''other'' *N*-glycans 10 and 17 have both a Fuc residue (Figure 2), excluding them as high mannose *N*-glycans. The ''other'' *N*-glycans 1, 2, 4, and 5 did not have 5 to 9 Man residues (Figure 2). High mannose *N*-glycans merely consists of Man building blocks. The ''other'' *N*-glycan 3 lacks either a Fuc or Gal residue (Figure 2) to be classified as hybrid *N*-glycan, and does not have two or three GlcNAc residues like the complex *N*-glycans (Figure 2).

Another classification system has been introduced to group the different type of *N*-glycans[14]. Using this classification system, 11 (17%) and 55 (83%) structures can be grouped as acidic and neutral *N*-glycans, respectively, and 37 (56%) as fucosylated (Table 1, classification system type II[14]). The relative occurrence of fucosylated *N*-glycans in human milk was

mentioned in several studies[14-16]. Two earlier studies found that the numbers of fucosylated *N*-glycans were ranging between 65-75%[14,15]. A more recent paper, which used a larger sample size (10 mothers and three time points), found that 16 (55%) of the 29 *N*-glycans found were fucosylated[16]. Based on the structural features of the *N*-glycans, as mentioned in literature[26,35], the fucosylated *N*-glycans in Figure 2 with a single Fuc residue are probably α1,6-linked by FUT8 during biosynthesis to the GlcNAc residue at the reducing end. The *N*-glycans containing more than one Fuc residue (Figure 2), might be formed due to the presence of other fucosyltransferases. Multiple (>50) *N*-glycans from different blood and tissue glycoproteins have been structurally characterized, containing α1,2-, α1,3-, and α1,6-linked Fuc linkages[35]. The extra Fuc residues in the *N*-glycans of these blood and tissue glycoproteins were α1,2- and α1,3-linked to a Gal residue and a peripheral GlcNAc residue, respectively[35]. It has also been found that fucosylation of *N*-glycans is modified by FUT2[24], which decorate the Fuc residues by α1,2-linkages[24]. These fucose containing *N*-glycans might also be important for the infant's healthy development, as has been reported for fucosylated HMOs.

Untargeted statistics with the relative levels of individual N-glycans in milk of 7 Chinese mothers over lactation

The averaged levels of the individual *N*-glycans per mother and per lactation stage can be found in Table S1. Separation of the different clusters coincides with the type of secretor status and lactation time as indicated in Figure 3.

Mother 1 and 4 can be assigned to the $Se^{-}Le^{+}$ group, whereas the mother 2, 3, 5, 6 and 7 belong to the $Se^{+}Le^{+}$ group. Based on the PCA plot three different groups could be observed (I-III). For the first time, a clear difference can be observed with respect to milk of 2 $Se^{-}Le^{+}$ mothers (I) and 5 $Se^{+}Le^{+}$ mothers (II and III), based on the levels of the individual *N*-glycans. It can also be observed that milk of the $Se^{+}Le^{+}$ mothers was strongly grouped per lactation stage (II and III). A similar trend could be observed for the 2 $Se^{-}Le^{+}$ mothers, however, a larger sample size is needed for confirmation.

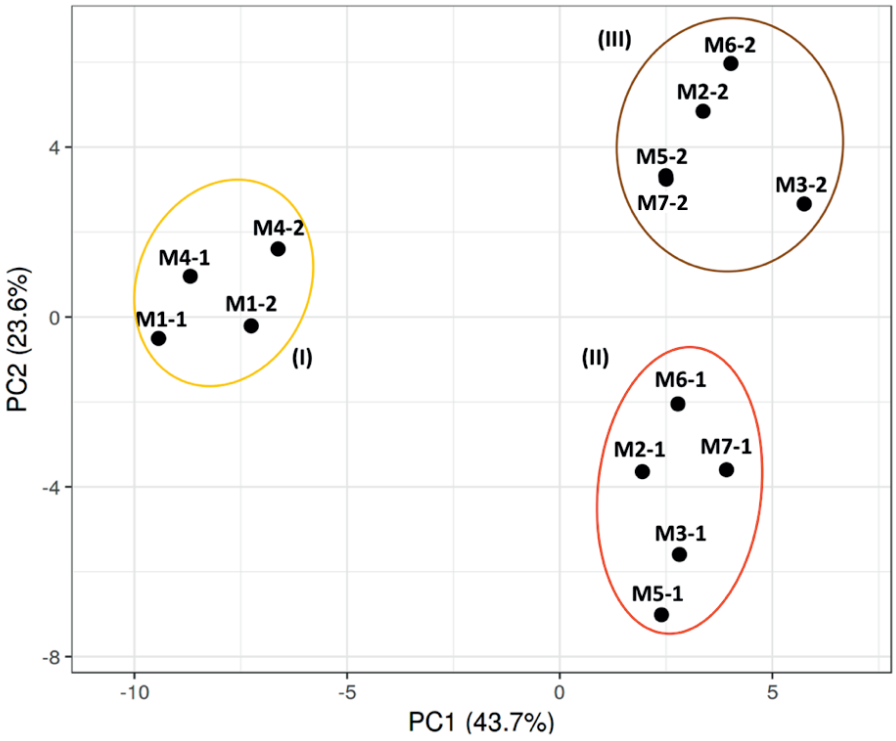


Figure 3. PCA plot of the Chinese human milk serum glycoproteome over lactation, using the relative level of each single *N*-glycan per mother and per lactation stage. Mother 1 and 4 were assigned to the $Se^{-}Le^{+}$ group, whereas mother 2, 3, 5, 6 and 7 were grouped in the $Se^{+}Le^{+}$ group. The numbers after the hyphen per mother indicates colostrum (1) and mature milk (2).

The individual N-glycans in milk of $Se^{+}Le^{+}$ and $Se^{-}Le^{+}$ mothers over lactation grouped on classification system I and II

The levels of the 66 different *N*-glycans were grouped, according to classification system I (Figure S1) and II (Figure 4) per mother and lactation stage. Based on classification system I (high mannose, complex, hybrid and other *N*-glycans), the relative levels of the total high mannose and total ''other'' *N*-glycans decreased over lactation for the 2 $Se^{-}Le^{+}$ mothers, whereas the relative levels of the total complex *N*-glycans increased over time (Figure S1). The relative levels of the total hybrid *N*-glycans remained constant over lactation (Figure S1). The group containing the complex *N*-glycans covered >65% of the total *N*-glycan content, for all 7 mothers (Figure S1). It was reported before that complex *N*-glycans individually are highly abundant in human milk, and the most dominant type of *N*-glycans present in mature milk among mothers[14].

As mentioned above, classification system II considered all the different structural features of the *N*-glycans. The levels of the total neutral fucosylated *N*-glycans in milk from 5 $Se^{+}Le^{+}$ mothers slightly decreased over lactation, while the total levels of acidic *N*-glycans remained constant, and the total levels of neutral nonfucosylated *N*-glycans increased (Figure 4). This pattern could not be observed for the milk from the 2 $Se^{-}Le^{+}$ mothers. The profiles of the three *N*-glycan groups stayed constant over lactation for the 2 $Se^{-}Le^{+}$ mothers (Figure 4). Despite the different patterns, the relative levels of the total neutral (sum of nonfucosylated and fucosylated) *N*-glycans ends up being the same for both genetic groups (Figure 4). The relative levels of total neutral *N*-glycans covered >90% of the total *N*-glycan content, respectively, for all 7 mothers (Figure 4).

The patterns for the total acidic and total neutral *N*-glycan content of human milk proteins over lactation (Figure 4) did not completely match with literature[16]. It has been reported by others that the levels of fucosylated *N*-glycans decreased from circa 61% in colostrum (3 days) to 37% in transition milk (9 days), and then remain constant in mature milk (40 days)[16]. This large drop could not be observed here for neutral fucosylated *N*-glycans over time (Figure 4). It was also reported that the levels of the nonfucosylated *N*-glycans increased over time, and the levels of acidic *N*-glycans in milk proteins over time were ranging from 5-12%, with a little increase over lactation[16]. In this latter study, it was also found that the overall *N*-glycan content was not affected by individual acidic *N*-glycans, which were highly abundant[16]. In this current study, the relative amounts of the total acidic *N*-glycans were ranging between 3–8% in milk of 7 mothers over time (Figure 4). However, none of the acidic *N*-glycans belong here to the most abundant *N*-glycans (Table S2).

Two other studies showed a complete different pattern for the acidic *N*-glycans in mature milk[14,15]. By abundance, 47% and 57% of the *N*-glycans were sialylated[14,15]. Twenty-seven *N*-glycans were not found in our study (Table S3), as compared to literature[14-16], including 13 acidic *N*-glycans (Table S3). It seems unlikely that these acidic *N*-glycans in Table 1 and Table S3 belong to the most abundant *N*-glycans. Some of the highly acidic *N*-glycans (e.g. 1915.7, 2881.1) were only found once in literature (Table S3), while other structures (e.g. 40, 49) were found in low quantities by us (Table S1) and others[15,16]. In addition, a recent study investigated the core fucosylation patterns of serum proteins in milk of 56 Chinese mothers[25]. In this latter study, acidic *N*-glycans were also not highly abundant in human milk[25]. The type and level of individual *N*-glycans in milk of the individual mothers were not investigated[25], as the study evaluated the role of FUT8 in respect to the formation of a healthy micriobiota[25].

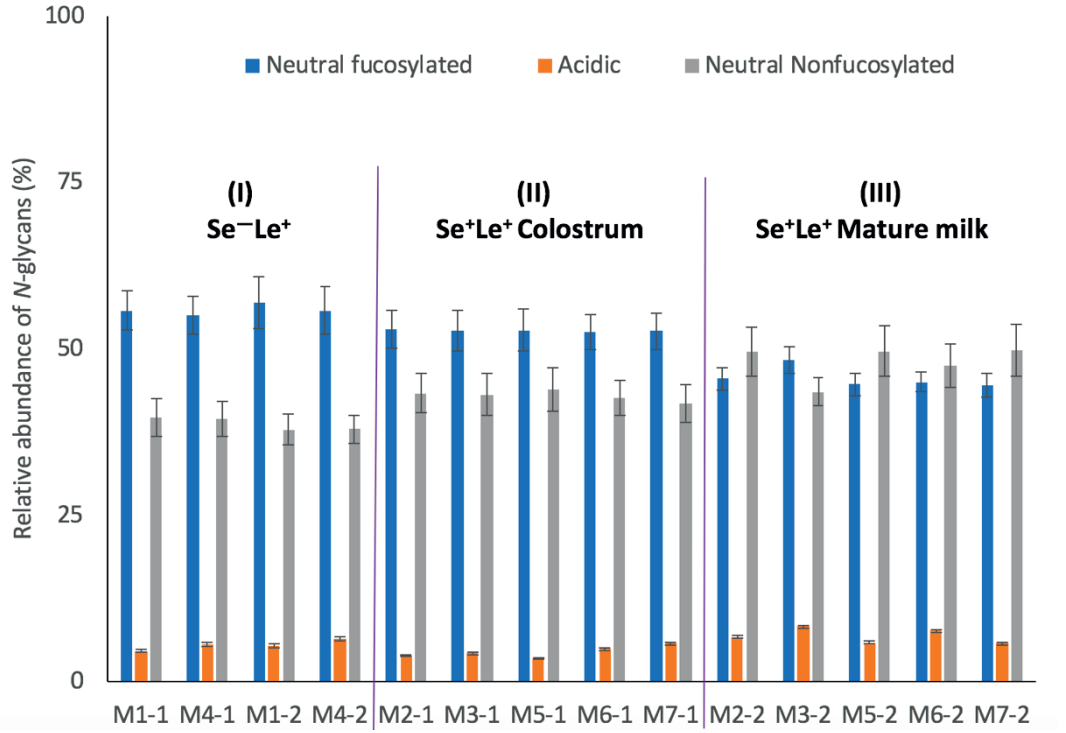


Figure 4. Total acidic and neutral (nonfucosylated and fucosylated) *N*-glycan content in milk of 7 mothers from 2 different lactation periods. The acidic (nonfucosylated and fucosylated) *N*-glycans were grouped together. Mother 1 and 4 were assigned to the $Se^{-}Le^{+}$ group, whereas mother 2, 3, 5, 6 and 7 belong to the $Se^{+}Le^{+}$ group. The numbers after the hyphen per mother indicates colostrum (1) and mature milk (2). Roman numerals (I-III) refer to the groups in the PCA plot in Figure 3.

Individual N-glycans in colostrum (week 1) and mature milk (week 4) of 7 Chinese mothers
Besides the PCA plot also a heatmap was generated. The differences in individual serum protein *N*-glycans in colostrum and mature milk of the Chinese mothers can be investigated using a heatmap, showing variation in both the type and levels of specific individual *N*-glycans among mothers and over time (Figure 5). For example, the levels of the neutral fucosylated *N*-glycan 25 for Chinese mother 2 was higher in colostrum than in mature milk (Figure 5). The level of the neutral fucosylated *N*-glycan 33 was higher in both colostrum and mature milk from Chinese mother 4 in comparison to Chinese mother 2 (Figure 5).
Although the heatmap (Figure 5) provided insights in the Chinese human milk serum glycoproteome over lactation, it is quite hard to observe accurately the differences in type and levels of individual *N*-glycans between the individual mothers and over lactation. Therefore, the profiles of the individual *N*-glycans over time can be found in Figure 6 and 7, for milk of Chinese mother 2 (Se^+Le^+ status) and Chinese mother 4 (Se^-Le^+ status), respectively.

The levels of the individual N-glycans in colostrum and mature milk of Chinese mother 2 with the Se^+Le^+ status
It can be seen in Figure 6 that neutral nonfucosylated *N*-glycan 16 and fucosylated *N*-glycan 25 are both highly abundant in colostrum (13.0%) of Chinese mother 2. Other highly abundant neutral *N*-glycans in colostrum of Chinese mother 2 are structures 35, 33, 26, 11, 41, 6, 24, 36, 23, 44, 34, 15, 50, ordered from most to least abundant (Figure 6).
The levels of the neutral nonfucosylated *N*-glycan 16 and 34 increased (14.4% → 17,3% and 1.6% → 2.3%) over time, respectively (Figure 6). The levels of the neutral fucosylated *N*-glycans 25, 35 and 41 decreased (12.3% → 7,9%, 9.3% → 5.0%, and 3.2% → 2.2%) over lactation, respectively (Figure 6). The levels of the neutral nonfucosylated *N*-glycan 36 also decreased (2.6% → 1.8%) from colostrum to mature milk, respectively (Figure 6). The neutral fucosylated *N*-glycans 4, 5, and 10 were completely absent in milk of Chinese mother 2 (Figure 6). The neutral fucosylated *N*-glycans 7 and 46, the nonfucosylated *N*-glycans 38 and 39, and acidic nonfucosylated *N*-glycan 40 were only present in colostrum, while the acidic nonfucosylated *N*-glycan 19 and neutral fucosylated *N*-glycan 29 were only present in mature milk (Figure 6, Figure 2).

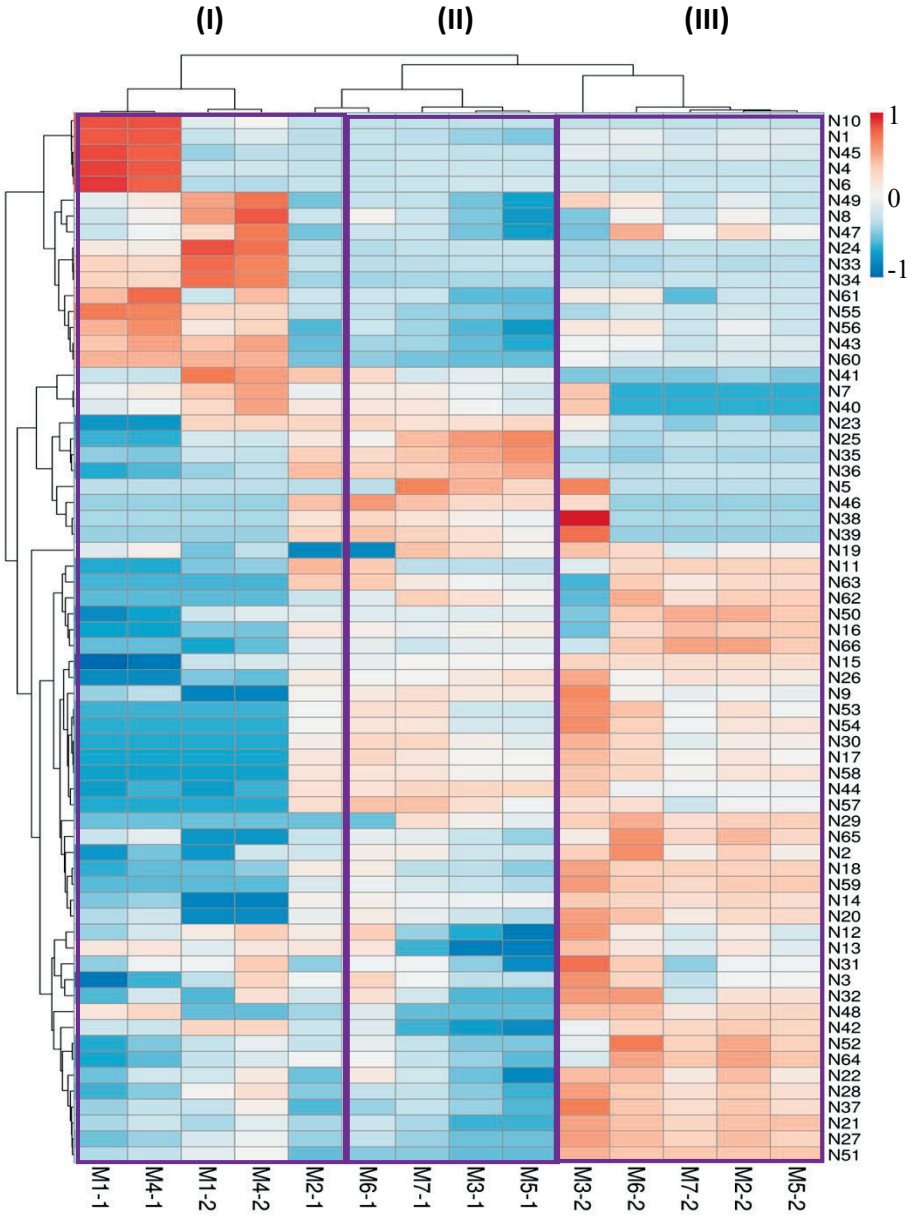


Figure 5. Heatmap of *N*-glycans in colostrum and mature milk per mother, using the relative abundancies of the each single *N*-glycan. The levels of individual *N*-glycans are represented as different colors. The colors blue and red represent the lowest and highest values of the *N*-glycans, respectively. Mother 1 and 4 were assigned to the $Se^{-}Le^{+}$ group, whereas mother 2, 3, 5, 6 and 7 were grouped in the $Se^{+}Le^{+}$ group. The numbers after the hyphen per mother indicates colostrum (1) and mature milk (2). Roman numerals (I-III) refer to the groups in the PCA plot in Figure 3.

The levels of the individual N-glycans in colostrum and mature milk of Chinese mother 4 with the $Se^{-}Le^{+}$ status

It can be seen in Figure 7 that the neutral fucosylated *N*-glycan 33 is highly abundant in colostrum (13.2%) of Chinese mother 4. The other highly abundant *N*-glycans in colostrum of Chinese mother 4 are 6, 4, 34, 25, 35, 24, 16, 45, 55, 60, 26, 41, 1, 10, ordered from most to least abundant (Figure 7). The majority of the top 15 structures can be categorized as neutral fucosylated *N*-glycans (Figure 7), despite the fact that this milk belongs to the $Se^{-}Le^{+}$ group.

The levels of the neutral fucosylated *N*-glycans 33, 25, 35 increased (13.2% → 18,7%, 4.5% → 8.8%, and 4.4% → 5.5%) over lactation, respectively, while the levels of the neutral nonfucosylated *N*-glycans 6 and 45 decreased (12.1% → 2,5% and 3.4% → 0,6%) over time, respectively (Figure 7). The levels of the neutral fucosylated *N*-glycans 4, 55, 1 and 10 decreased (9.1% → 0.5%, 3.2% → 2.4%, 2.1% → 0.7% and 1.9% → 0.5%) from colostrum to mature milk, respectively (Figure 7). The levels increased for the nonfucosylated *N*-glycans 34, 24, 16, 26 and 41 (4.6% → 6.4%, 3.9% → 6.2%, and 3.7 → 6.3%, 2.5% → 3.8% and 2.4% → 3.4%) over time, respectively (Figure 7).

The neutral fucosylated *N*-glycans 20 and 65, the neutral fucoyslated *N*-glycan 9, and the acidic nonfucosylated *N*-glycan 48 were only present in colostrum (Figure 7). The neutral fucosylated *N*-glycans 5, 17, 30, 46, 54, 57-59, 62, 63, the acidic nonfucosylated *N*-glycans 29 and 53, and the neutral nonfucosylated *N*-glycans 38 and 39 were completely absent in milk of Chinese mother 4 (Figure 7, Figure 2).

Comparison of the individual N-glycans in milk of Chinese mother 2 and 4 with the $Se^{+}Le^{+}$ and $Se^{-}Le^{+}$ status, respectively

The neutral fucosylated *N*-glycan 33 was present in milk of the $Se^{+}Le^{+}$ and $Se^{-}Le^{+}$ mother, with higher levels for the $Se^{-}Le^{+}$ mother in comparison to the $Se^{+}Le^{+}$ mother (Figure 6 and 7). The levels of the neutral *N*-glycan 6 was also higher in colostrum for the $Se^{-}Le^{+}$ mother compared to the $Se^{+}Le^{+}$ mother. The levels of the neutral fucosylated *N*-glycans 25 and 35 were higher in colostrum for the $Se^{+}Le^{+}$ mother in comparison to $Se^{-}Le^{+}$ mother. The levels of the neutral *N*-glycans 11 and 26 were both higher in colostrum and mature milk for the $Se^{+}Le^{+}$ mother in comparison to $Se^{-}Le^{+}$ mother. The neutral *N*-glycan 34 was more dominant in colostrum and mature milk for the $Se^{-}Le^{+}$ mother in comparison to $Se^{+}Le^{+}$ mother. From the total 66 *N*-glycans, the neutral fucosylated *N*-glycans 4, 5 and 10 were completely absent in the milk from the $Se^{+}Le^{+}$ mother (Figure 6), while more *N*-glycans 5, 17, 29-30, 38-39, 46, 53, 54, 57-59, 62-63 were missing in the milk of the $Se^{-}Le^{+}$ mother of which 10 structures were neutral fucosylated (Figure 7). These missing structures already discriminate between the two different milk-types.

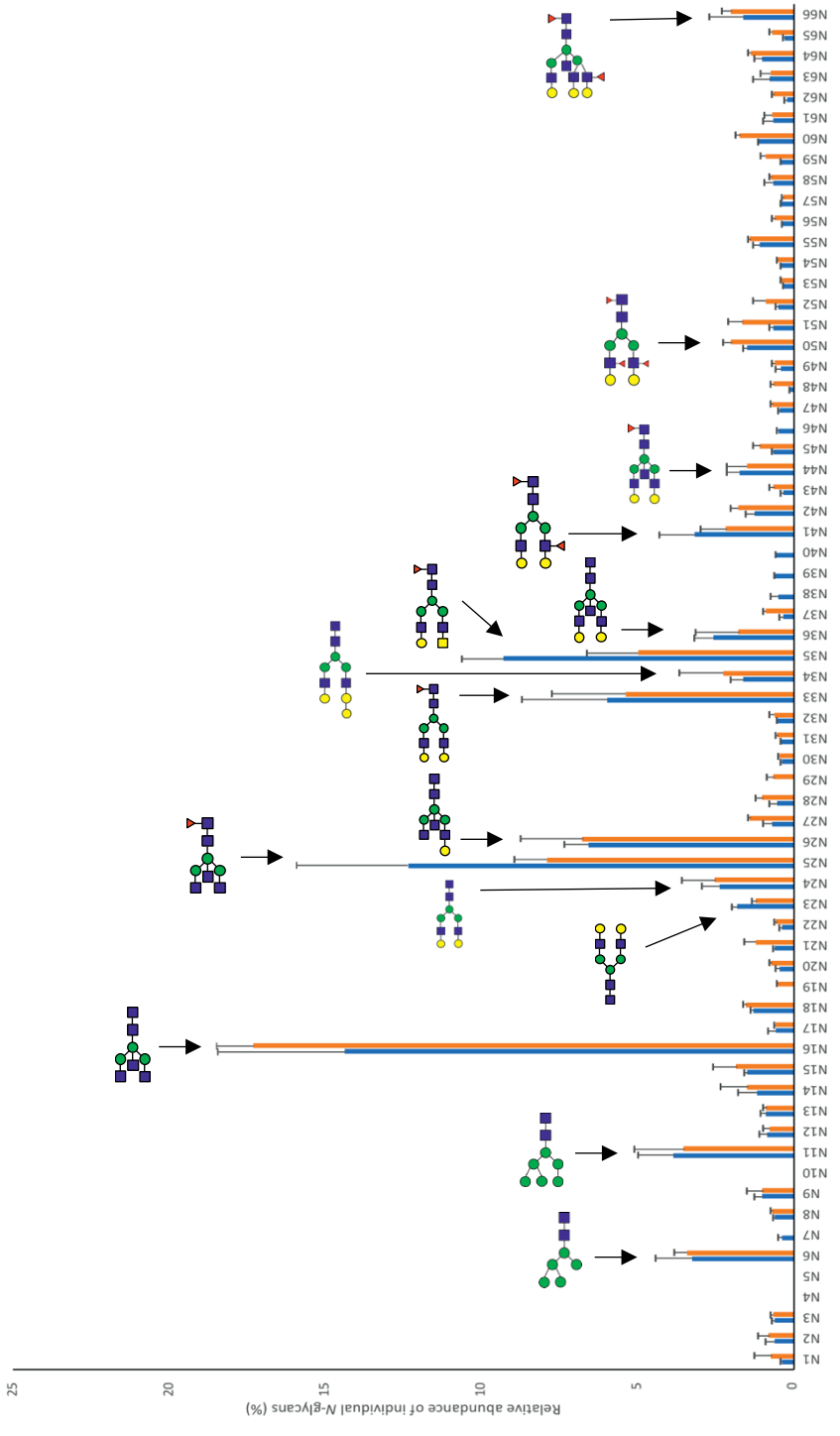


Figure 6. Individual *N*-glycan profiles in colostrum (blue bars) and mature milk (orange bars) of Chinese mother 2, as measured using MALDI-TOF-MS. Mother 2 belongs to the Se^+Le^+ group. Numbers on the x-axis indicate the *N*-glycans displayed in Table 1. Biological replicates (N = 2). The structures displayed highlight the 15 most abundant *N*-glycans.

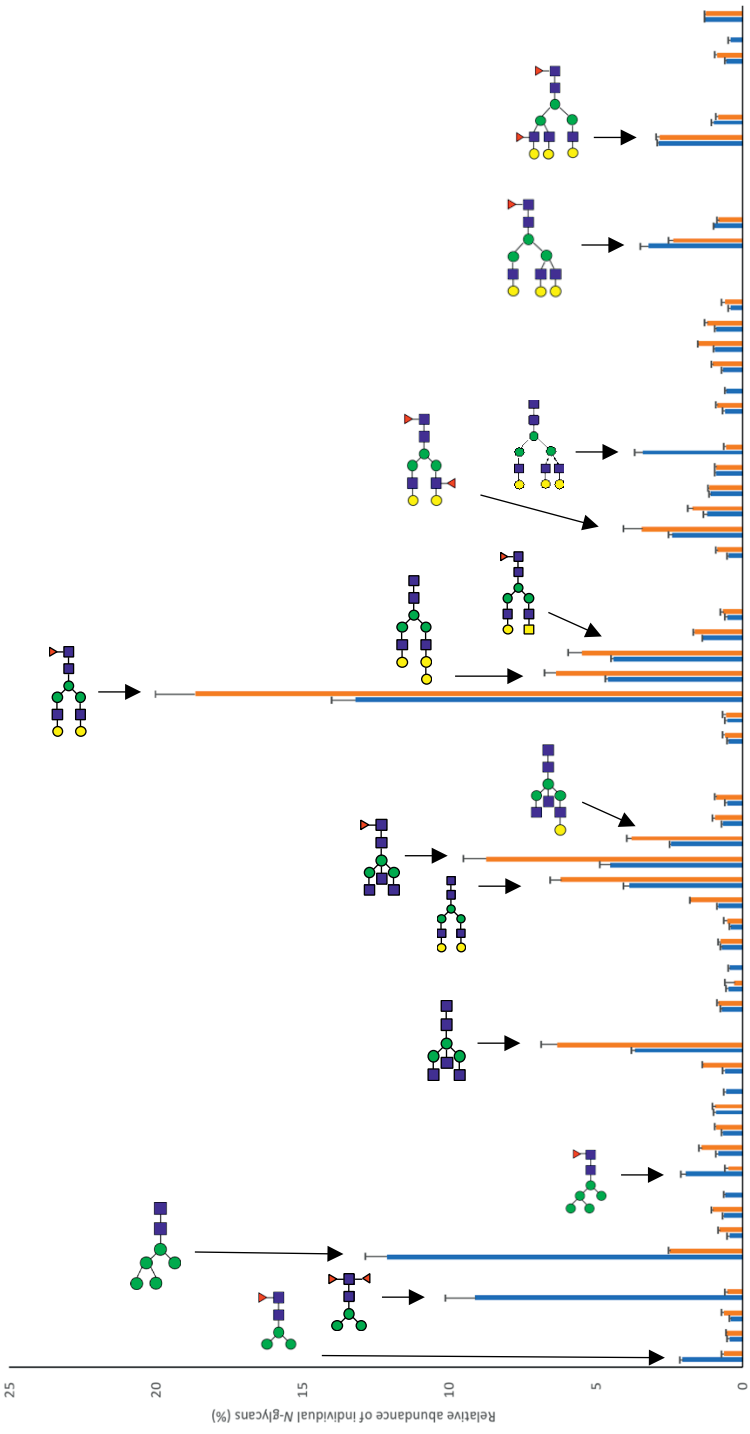


Figure 7. Individual *N*-glycan profiles in colostrum (blue bars) and mature milk (orange bars) of Chinese mother 4, as measured using MALDI-TOF-MS. Mother 4 was grouped in the $Se^{-}Le^{+}$ group. Numbers on the x-axis indicate the *N*-glycans displayed in Table 1. Biological replicates (N = 2). The structures displayed highlight the 15 most abundant *N*-glycans.

It might be that structures 54, 57-59, 63, which lack an α1,2-linked Fuc residue, are not present due to the absence of the FUT2 enzyme for the $Se^{-}Le^{+}$ mother. These latter structures were only present in the milk from the $Se^{+}Le^{+}$ mothers (Table S1). The structures 15, 17 and 62 only contain α1,6-linked Fuc residues due to the enzyme FUT8. The *N*-glycans 30 and 46 were only present in the milk from the $Se^{+}Le^{+}$ mothers (Table S1), which suggest that these Fuc residues might be α1,2-linked to a Gal residue instead of a α1,3-linked to peripheral GlcNAc residue (Figure 2). No distinction between isomers could be made by MALDI-TOF-MS.

The levels of the neutral fucosylated *N*-glycans 25, 33, 35 decreased from colostrum to mature milk for the $Se^{+}Le^{+}$ mother, whereas the neutral *N*-glycan 16 became more dominant over time. In contract to the milk from the $Se^{+}Le^{+}$ mother, the levels of neutral *N*-glycans 6, 45 decreased over time for the $Se^{-}Le^{+}$ mother, while the levels of the neutral fucosylated *N*-glycans 25, 33, 35 increased over time. As a consequence, other neutral fucosylated *N*-glycans 1, 10, 33, 55, 60 might explain why the total neutral fucosylated concentrations ends up being the same for both genetic groups (Figure 4). The milk of 1 $Se^{+}Le^{+}$ mother and 1 $Se^{-}Le^{+}$ mother thus mainly differed based on neutral fucosylated *N*-glycans. The same top 15 in *N*-glycans was found in colostrum and mature milk for both $Se^{-}Le^{+}$ mothers (Table S2). A similar top 15 in *N*-glycans can be found for all $Se^{+}Le^{+}$ mothers over lactation. The profiles of the top 15 *N*-glycans for the $Se^{+}Le^{+}$ mothers 3, 5 and 7 have more in common than the $Se^{+}Le^{+}$ mothers 2 and 6 (Table S2). The patterns (increase/decrease in levels) of the individual 15 *N*-glycans behave differently over time for the $Se^{+}Le^{+}$ mothers 2 and 6 (Table S2) and the $Se^{+}Le^{+}$ mothers 3, 5 and 7. The 15 most abundant *N*-glycans covered in levels >72% and >65% of the total *N*-glycan content in colostrum and mature milk, respectively (Table S2). In contrast to the most abundant serum proteins[11] and HMOs[28] in human milk, a much larger variety in type and levels can be observed for the top 15 *N*-glycans among mothers and over lactation. This indicates that low abundant *N*-glycans should deserve the same attention as the relative highly abundant *N*-glycans. Overall, *N*-glycans share several building blocks with HMOs although the latter does not have Man and GalNAc residues. In addition, most *N*-glycan and HMO structures are fucosylated. Based on this study, it can be concluded that fucosylation of serum *N*-glycans was associated with the mother's secretor status.

Conclusions

This study aimed to fill a gap in literature by investigating the *N*-glycan profiles in milk of 7 mothers over lactation individually. For this purpose, an accurate and reproducible method was needed. The procedure to remove HMOs was efficient, addition of 2-BME improved denaturation of serum proteins, and the incubation time and amount of the PNGase F was optimized. After method optimization and validation, a larger set of human milk samples was used. Acidic *N*-glycans do not belong to the 15 most abundant *N*-glycans, as mainly neutral

fucosylated and nonfucosylated *N*-glycans can be found in colostrum and mature milk, for all 7 mothers. The difference between secretor status was mainly based on the neutral fucosylated *N*-glycans.

References

1. Kunz, C.; Rodriguez-Palmero, M.; Koletzko, B.; Jensen, R. Nutritional and biochemical properties of human milk, part I: General aspects, proteins, and carbohydrates. Clin. Perinatol. **1999**, 26, 307–33.
2. Labbok, M.; Clark, D.; Goldman, A. Breastfeeding: Maintaining an irreplaceable immunological resource. Nat. Rev. Immunol. **2004**, 4, 565–72.
3. Ballard, O.; Morrow, A. Human milk composition: Nutrients and bioactive factors. Pediatr. Clin. North Am. **2013**, 60, 49–74.
4. Bode, L. Recent advances on structure, metabolism, and function of human milk oligosaccharides. J. Nutr. **2006**, 136, 2127−30.
5. De Leoz, M.; Wu, S.; Ninonuevo, M.; Gaerlan, S.; Mirmiran, M.; German, J.; Mills, D.; Lebrilla, C.; Underwood, M. A quantitative and comprehensive method to analyse human milk oligosaccharide structures in the urine and faeces of infants. Anal. Bioanal. Chem. **2013**, 405, 4089–105.
6. Hettinga, K.A.; Van Valenberg, H.; De Vries, S.; Boeren, S.; Van Hooijdonk, T.; Van Arendonk, J.; Vervoort, J. The host defense proteome of human and bovine milk. PLoS One, **2011**, 6, 1–8.
7. Froehlich, J.; Dodds, E.; Barboza, M.; McJimpsey, E.; Seipert, R.; Francis, J.; An, H.; Freeman, S.; German, J; Lebrilla, C. Glycoprotein expression in human milk during lactation. J. Agric. Food Chem. **2015**, 58, 6440–8.
8. Peterson, R.; Cheah, W.; Grinyer, J.; Packer, N. Glycoconjugates in human milk: Protecting infants from disease. Glycobiology, **2013**, 23, 1425–38.
9. Karav, S.; Le Parc, A.; Bell, J.; Frese, S.; Kirmiz, N.; Block, B.; Barile, D.; Mills, D. Oligosaccharides released from milk glycoproteins are selective growth substrates for infant-associated bifidobacteria. Appl. Environ. Microbiol. **2016**, 82, 3622–30.
10. Kirmiz, N.; Robinson, R.; Shah, I.; Barile, D.; Mills, D. Milk glycans and their interaction with the infant-gut microbiota. Annu Rev. Food Sci. Technol. **2018**, 9, 429–50.
11. Elwakiel, M.; Boeren, S.; Hageman, J.A.; Szeto, I.M.; Schols, H.A.; Hettinga, K.A. Variability of serum proteins in Chinese and Dutch human milk during lactation. Nutrients, **2018**, 11, 499–513.
12. Lönnerdal, B. Bioactive proteins in human milk: Health, nutrition, and implications for infant formulas. J. Pediatr. **2016**, 173, 4–9.
13. Picariello, G.; Ferranti, P.; Mamone, G.; Roepstorff, P.; Addeo, F. Identification of *N*-linked glycoproteins in human milk by hydrophilic interaction liquid chromatography and mass spectrometry. Proteomics, **2008**, 8, 3833–47.
14. Dallas, D.; Martin, W.; Strum, J.; Zivkovic, A.; Smilowitz, J.; Underwood, M.; Affolter, M.; Lebrilla, C.; German, B. *N*-linked glycan profiling of mature human milk by high performance microfluidic chip liquid chromatography time of flight tandem mass spectrometry. J. Agric. Food Chem. **2011**, 59, 4255–63.

15. Nwosu, C.; Aldredge, D.; Lerno, L.; Zivkovic, A.; German, B.; Lebrilla, C. Comparison of the human and bovine milk *N*-glycome via high-performance microfluidic chip liquid chromatography and tandem mass spectrometry. J. Proteome Res. **2012**, 11, 2912–24.
16. Lu, L.; Liu, J.; Jia, Y.; Yang, Y.; Chen, Q.; Sun, L.; Song, S.; Huang, L.; Wang, Z. Mass spectrometry analysis of changes in human milk *N/O*-glycopatterns at different lactation stages. J. Agric. Food Chem. **2019**, 67, 10702–12.
17. Van Leeuwen, S.; Schoemaker, R.; Timmer, C.; Kamerling, J.; Dijkhuizen, L. *N*-and *O*-glycosylation of a commercial bovine whey protein product. J. Agric. Food Chem. **2012**, 60, 12553–64.
18. Takimori, S.; Shimaoka, H.; Furukawa, J.; Yamashita, T.; Fujitani, N.; Takegawa, Y.; Hammarstrom, L.; Kacskovics, I.; Shinohara, Y.; Nishimura, S. Alteration of the *N*-glycome of bovine milk glycoproteins during early lactation. FEBS J. **2011**, 278, 3769–81.
19. Wang, W.; Wang. L.; Du, Y.; Wu.; Yu, X.; Ye, K.; Li, C.; Jung, Y.; Qian, Y.; Voglmeir, J.; Liu, L. Comparison of antipathogenic activities of the human and bovine milk *N*-glycome: Fucosylation is a key factor. Food Chem. **2017**, 235, 167–74.
20. Landberg, E.; Huang, Y.; Stromgvist, M.; Mechref, Y.; Hansson, L.; Lundblad, A.; Novotny, M.; Pahlsson, P. Changes in glycosylation of human bile salt-activated lipase during lactation. Arch. Biochem. Biophys. **2000**, 337, 246–52.
21. Le Parc, A.; Karav, S.; Rouquie, C.; Maga, E.; Bunyatratchata, A.; Barile, D. Characterization of recombinant human lactoferrin *N*-glycans expressed in the milk of transgenic cows. PLoS One, **2017**, 12, 1–15.
22. Barboza, M.; Pinzon, J.; Wickramasinghe, S.; Froehlich, J.; Moeller, I.; Smilowitz, J.; Ruhaak, L.; Lönnerdal, B.; German, J.; Medrano, J.; Weimer, B.; Lebrilla, C. Glycosylation of human milk lactoferrin exhibits dynamic changes during early lactation enhancing its role in pathogenic bacteria host interactions. MCP, **2012**. 11, 1–10.
23. Royle, L.; Roos, A.; Harvey, D.; Wormald, M.; Van Gijlswijk-Janssen, D.; Redwan, E.; Wilson, I.; Daha, M.; Dwek, R.; Rudd, P. Secretory IgA *N*- and *O*-glycans provide a link between the innate and adaptive immune systems. J. Biol. Chem. **2003**, 278, 20140–53.
24. Bai, Y.; Tao, J.; Zhou, J.; Fan, Q.; Liu, M.; Hu, Y.; Xu, Y.; Zhang, L.; Yuan, Y.; Li, W.; Ze, X.; Malard, P.; Guo, Z.; Yan, J.; Li, M. Fucosylated human milk oligosaccharides and *N*-glycans in the milk of Chinese mothers regulate gut microbiome of their breast-fed infants during different lactation stages. mSystems, **2018**, 3, 1–19.
25. Li, M.; Bai, Y.; Zhou, J.; Huang, W.; Yan, J.; Tao, J.; Fan, Q.; Liu, Y.; Mei, D.; Yan, Q.; Yuan, J.; Malard, P.; Wang, Z.; Gu, J.; Tanigchi, N.; Li, W. Core fucosulation of maternal milk of *N*-glycans evokes B cell activation by selectively promoting the L-fucose metabolism of gut *Bifidobacterium* spp. and *Lactobacillus* spp. mBio, **2019**, 10, 1–19.
26. Holst, S.; Wilding, J.; Koprowska, K.; Rombouts, Y.; Wuhrer, M. *N*-glycomic and transcriptomic changes associated with CDX1 mRNA expression in colorectal cancer cell lines. Cell, **2019**, 8, 273–94.
27. Lis-Kuberka, J.; Kątnik-Prastowska, I.; Berghausen-Mazur, M.; Orczyk-Pawilowicz, M. Lectin-based analysis of fucosylated glycoproteins of human skim milk during 47 days of lactation. Glycoconj. J. **2015**, 32**,** 665–74.

28. Elwakiel, M.; Hageman, J.A.; Wang, W.; Szeto, I.M.; van Goudoever, J.B.; Hettinga, K.A.; Schols, H.A. Human milk oligosaccharides in colostrum and mature milk of Chinese mothers: Lewis positive secretor subgroups. J. Agric. Food Chem. **2018**, 66, 7036–43.
29. Morelle, W.; Michalski, J. Analysis of protein glycosylation by mass spectrometry. Nat. Protoc. **2007**, 2, 1585–602.
30. Chu, C.; Ninonuevo, M.; Clowers, B.; Perkins, P.; An, H.; Yin, H.; Killeen, K.; Miyamoto, S.; Grimm, R.; Lebrilla, C. Profile of native *N*-linked glycan structures from human serum using high performance liquid chromatography on a microfluidic chip and time-of-flight mass spectrometry. Proteomics, **2009**, 9, 1939–51.
31. Aldredge, D.; An, H.; Tang, N.; Waddell, K.; Lebrilla, C. Annotation of a serum *N*-glycan library for rapid identification of structures. J. Proteome Res. **2012**, 11, 1958–68.
32. Kita, Y.; Miura, Y.; Furukawa, J.; Nakano, M.; Shinohara, Y.; Ohno, M.; Takimoto, A.; Nishimura, S. Quantitave glycomics of human whole serum glycoproteins based on the standardized protocol for liberating *N*-glycans. MCP, **2007**, 6, 1437–45.
33. Albrecht, S.; van Muiswinkel, G.C.; Schols, H.A.; Voragen, A.G.; Gruppen, H. Introducing capillary electrophoresis with laser-induced fluorescence detection for the characterization of konjac glucomannan oligosaccharides and their *in vitro* fermentation behavior. J. Agric. Food Chem. **2009**, 13, 3867–76.
34. Tiemeyer, M.; Paulson, J.; York, W.; Lisacek, F.; Campbell, M.; Fijiti, A.; Shinmachi, D.; Yamada, I.; Ranzinger, R.; Aoki-Kinoshita, K. GlyTouCan: An accessible glycan structure repository. Glycobiology, **2017**, 27, 915–9.
35. Vliegenthart, J.; Dorland, L.; Van Halbeek, H. High resolution,'H-nuclear magnetic resonance spectroscopy as a tool in the structural analysis of carbohydrates related to glycoproteins. Adv. Carbohydr. Chem. Biochem. **1983**, 41, 209–374.
36. Plaza-Diaz, J.; Fontana, L.; Gil. Human milk oligosaccharides and immune system development. Nutrients, **2018**, 10, 1038–55.

Supporting information

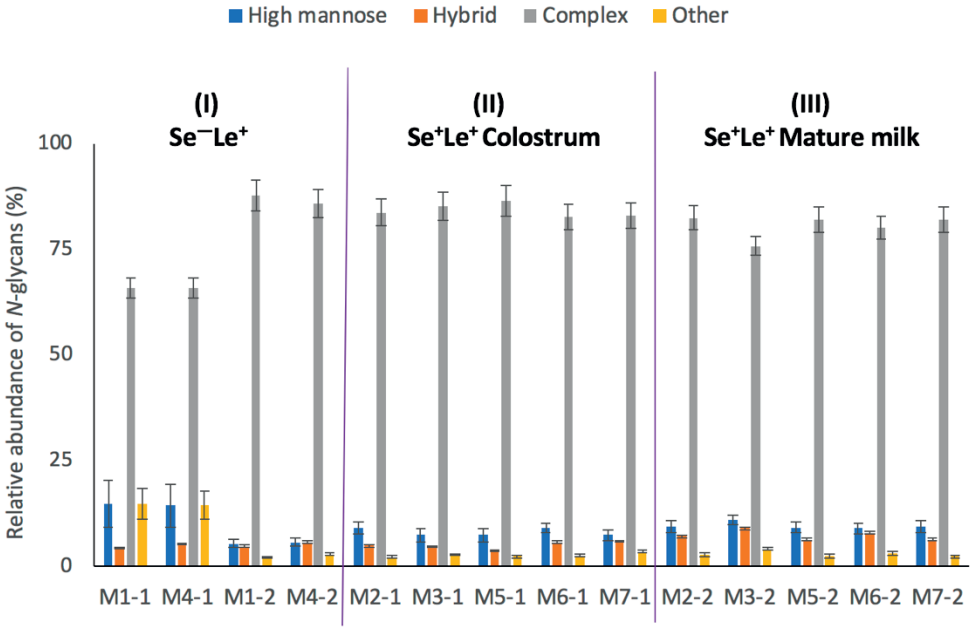


Figure S1. Complex, high mannose, hybrid, and other *N*-glycans in milk of 7 mothers over 2 different lactation periods. The numbers after the hyphen per mother indicates colostrum (1) and mature milk (2). Mother 1 and 4 can be assigned to the $Se^{-}Le^{+}$ milk group, whereas mother 2, 3, 5, 6 and 7 belong to the $Se^{+}Le^{+}$ group. Roman numerals (I-III) refer to the groups in the PCA plot in Figure 3.

Table S1. Relative amounts of *N*-glycans in colostrum (week 1) and mature milk (week 4) per mother (M1–7). Numbers indicate the *N*-glycans displayed in Table 1 and Figure 1. The numbers after the hyphen per mother indicates colostrum (1) and mature milk (2).

No.	Mass	M1-1	M1-2	M2-1	M2-2	M3-1	M3-2	M4-1	M4-2	M5-1	M5-2	M6-1	M6-2	M7-1	M7-2
N1	1056,4	2,1	0,5	0,4	0,7	0,3	0,7	2,1	0,7	0,2	0,7	0,5	0,8	0,4	0,6
N2	1072,3	0,4	0,4	0,6	0,8	0,6	0,8	0,5	0,6	0,6	0,7	0,7	0,9	0,7	0,7
N3	1113,3	0,3	0,5	0,6	0,6	0,5	0,8	0,4	0,7	0,5	0,6	0,7	0,7	0,6	0,5
N4	1202,4	9,6	0,4	0,0	0,0	0,3	0,5	9,1	0,5	0,2	0,0	0,0	0,0	0,4	0,0
N5	1218,4	0,0	0,0	0,0	0,0	0,4	0,5	0,0	0,0	0,3	0,0	0,0	0,0	0,5	0,0
N6	1234,3	12,9	2,5	3,2	3,4	3,6	4,1	12,1	2,5	3,8	3,4	3,1	3,2	3,3	3,5
N7	1259,3	0,4	0,7	0,4	0,0	0,4	0,7	0,5	0,8	0,3	0,0	0,5	0,0	0,5	0,0
N8	1275,4	0,6	0,9	0,6	0,7	0,5	0,5	0,7	1,0	0,4	0,6	0,7	0,7	0,6	0,6
N9	1316,3	0,5	0,0	1,0	1,0	1,1	1,7	0,6	0,0	1,1	0,9	1,1	1,0	1,2	0,9
N10	1380,6	1,9	0,3	0,0	0,0	0,0	0,0	1,9	0,5	0,0	0,0	0,0	0,0	0,0	0,0
N11	1396,3	0,8	1,3	3,8	3,5	1,7	2,3	0,8	1,4	1,7	3,4	3,6	3,3	1,7	3,5
N12	1421,4	0,6	0,8	0,8	0,8	0,5	1,0	0,7	0,9	0,4	0,7	0,9	0,8	0,6	0,7
N13	1437,4	0,9	0,8	0,9	0,9	0,5	1,0	0,9	0,9	0,5	0,8	0,9	0,9	0,6	0,8
N14	1462,5	0,5	0,0	1,2	1,5	1,1	1,6	0,6	0,0	1,1	1,4	1,2	1,5	1,1	1,4
N15	1478,4	0,5	1,3	1,5	1,8	1,6	1,9	0,6	1,4	1,6	1,8	1,5	1,8	1,6	1,8
N16	1519,4	3,8	6,7	14,4	17,3	13,1	6,0	3,7	6,3	13,9	17,4	13,3	15,6	11,8	18,0
N17	1542,4	0,0	0,0	0,6	0,6	0,5	0,8	0,0	0,0	0,5	0,5	0,7	0,7	0,6	0,5
N18	1558,4	0,6	0,7	1,3	1,5	0,9	1,7	0,7	0,8	0,8	1,5	1,3	1,5	0,9	1,5
N19	1566,9	0,4	0,2	0,0	0,5	0,6	0,7	0,5	0,3	0,5	0,5	0,0	0,6	0,7	0,4
N20	1583,4	0,3	0,0	0,5	0,7	0,4	0,9	0,4	0,0	0,3	0,7	0,6	0,8	0,5	0,6
N21	1599,5	0,6	0,6	0,6	1,2	0,4	1,3	0,7	0,8	0,4	1,2	0,7	1,2	0,6	1,1
N22	1607,3	0,3	0,4	0,3	0,6	0,3	0,6	0,4	0,5	0,2	0,5	0,5	0,6	0,4	0,5

Table S1 (continued)

No.	Mass	M1-1	M1-2	M2-1	M2-2	M3-1	M3-2	M4-1	M4-2	M5-1	M5-2	M6-1	M6-2	M7-1	M7-2
N23	1624,5	0,8	1,8	1,8	1,2	1,7	1,6	0,8	1,8	1,8	1,1	1,8	1,2	1,7	1,1
N24	1640,4	4,0	6,6	2,4	2,5	2,6	2,2	3,9	6,2	2,6	2,5	2,3	2,4	2,5	2,5
N25	1665,5	4,7	9,4	12,3	7,9	16,8	9,6	4,5	8,8	17,5	7,8	11,4	7,2	15,4	8,1
N26	1681,5	2,5	4,0	6,6	6,8	6,7	8,4	2,5	3,8	7,1	6,6	6,1	6,2	6,2	6,9
N27	1720,4	0,6	0,9	0,7	1,4	0,6	1,5	0,7	1,0	0,6	1,3	0,8	1,4	0,7	1,3
N28	1722,5	0,4	0,8	0,5	1,0	0,5	1,1	0,5	0,9	0,4	0,9	0,6	1,0	0,6	0,9
N29	1728,1	0,0	0,0	0,0	0,6	0,4	0,6	0,0	0,0	0,3	0,6	0,0	0,7	0,5	0,5
N30	1729,0	0,0	0,0	0,4	0,4	0,4	0,6	0,0	0,0	0,3	0,4	0,5	0,5	0,5	0,3
N31	1769,5	0,4	0,5	0,4	0,5	0,4	0,7	0,5	0,6	0,3	0,5	0,5	0,6	0,5	0,4
N32	1770,8	0,4	0,4	0,5	0,6	0,4	0,7	0,5	0,6	0,4	0,6	0,6	0,7	0,5	0,5
N33	1786,5	14,0	20,3	6,0	5,4	6,0	5,4	13,2	18,7	6,2	5,5	5,6	4,9	5,6	5,6
N34	1802,5	4,8	6,8	1,6	2,3	1,7	2,0	4,6	6,4	1,7	2,3	1,6	2,1	1,7	2,3
N35	1827,5	4,6	5,8	9,3	5,0	10,3	5,0	4,4	5,5	10,9	5,0	8,6	4,6	9,5	5,1
N36	1843,5	1,3	1,6	2,6	1,8	2,6	1,8	1,4	1,7	2,7	1,8	2,5	1,7	2,5	1,8
N37	1874,4	0,4	0,5	0,3	0,9	0,4	1,1	0,5	0,7	0,3	0,8	0,4	0,9	0,5	0,8
N38	1882,5	0,0	0,0	0,5	0,0	0,4	1,3	0,0	0,0	0,3	0,0	0,6	0,0	0,5	0,0
N39	1884,5	0,0	0,0	0,6	0,0	0,5	1,0	0,0	0,0	0,4	0,0	0,7	0,0	0,6	0,0
N40	1931,9	0,4	0,7	0,6	0,0	0,5	0,8	0,5	0,9	0,4	0,0	0,6	0,0	0,6	0,0
N41	1932,9	2,4	3,6	3,2	2,2	2,6	2,1	2,4	3,4	2,6	2,1	3,0	2,1	2,5	2,1
N42	1948,5	1,2	1,7	1,2	1,8	0,7	1,4	1,2	1,7	0,6	1,7	1,3	1,7	0,8	1,7
N43	1972,7	1,0	1,0	0,3	0,6	0,3	0,7	1,1	1,1	0,2	0,6	0,4	0,7	0,4	0,5
N44	1989,5	0,8	0,8	1,7	1,5	1,8	1,9	0,9	0,9	1,8	1,5	1,7	1,5	1,8	1,5

Table S1 (continued)

No.	Mass	M1-1	M1-2	M2-1	M2-2	M3-1	M3-2	M4-1	M4-2	M5-1	M5-2	M6-1	M6-2	M7-1	M7-2
N45	2005,6	3,5	0,4	0,6	1,1	0,7	1,2	3,4	0,6	0,7	1,0	0,7	1,1	0,8	1,0
N46	2014,5	0,0	0,0	0,5	0,0	0,4	0,4	0,0	0,0	0,4	0,0	0,6	0,0	0,5	0,0
N47	2036,6	0,5	0,7	0,4	0,7	0,4	0,4	0,6	0,9	0,3	0,6	0,5	0,8	0,5	0,6
N48	2060,9	0,5	0,0	0,1	0,6	0,0	0,7	0,6	0,0	0,0	0,6	0,3	0,7	0,0	0,5
N49	2077,5	0,6	0,9	0,4	0,6	0,4	0,8	0,7	1,0	0,3	0,5	0,5	0,7	0,5	0,5
N50	2078,6	0,9	1,4	1,5	2,0	1,5	1,2	1,0	1,5	1,5	1,9	1,5	1,9	1,5	2,0
N51	2094,6	0,8	1,1	0,6	1,6	0,6	1,7	0,9	1,2	0,6	1,6	0,7	1,6	0,7	1,6
N52	2118,7	0,3	0,5	0,5	0,9	0,4	0,6	0,4	0,6	0,4	0,8	0,6	1,0	0,5	0,8
N53	2134,7	0,0	0,0	0,3	0,4	0,2	0,6	0,0	0,0	0,2	0,3	0,4	0,5	0,4	0,3
N54	2135,5	0,0	0,0	0,4	0,5	0,3	0,8	0,0	0,0	0,3	0,5	0,5	0,6	0,5	0,4
N55	2151,6	3,3	2,4	1,1	1,4	0,8	1,0	3,2	2,4	0,7	1,3	1,1	1,4	0,9	1,3
N56	2167,6	0,9	0,7	0,3	0,6	0,3	0,7	1,0	0,8	0,2	0,5	0,5	0,7	0,4	0,5
N57	2176,7	0,0	0,0	0,4	0,3	0,4	0,4	0,0	0,0	0,3	0,3	0,5	0,4	0,5	0,2
N58	2224,6	0,0	0,0	0,7	0,7	0,6	0,9	0,0	0,0	0,6	0,7	0,7	0,8	0,7	0,6
N59	2240,7	0,0	0,0	0,4	0,9	0,3	1,1	0,0	0,0	0,2	0,9	0,5	0,9	0,4	0,8
N60	2297,7	2,9	2,9	1,1	1,7	1,0	2,0	2,9	2,9	1,0	1,7	1,2	1,7	1,1	1,7
N61	2313,6	0,9	0,7	0,7	0,7	0,6	0,8	1,0	0,9	0,6	0,7	0,7	0,8	0,7	0,6
N62	2354,8	0,0	0,0	0,2	0,6	0,5	0,0	0,0	0,0	0,4	0,6	0,3	0,7	0,6	0,5
N63	2370,8	0,0	0,0	0,8	0,7	0,5	0,0	0,0	0,0	0,4	0,7	0,8	0,8	0,6	0,6
N64	2443,8	0,5	0,8	1,0	1,4	0,7	0,9	0,6	0,9	0,6	1,3	1,0	1,4	0,8	1,3
N65	2459,8	0,3	0,0	0,3	0,7	0,3	0,5	0,4	0,0	0,2	0,6	0,4	0,8	0,4	0,6
N66	2500,6	1,3	1,2	1,6	2,0	1,6	1,5	1,3	1,3	1,6	1,9	1,6	1,9	1,5	2,0

Table S2. Top 15 most abundant *N*-glycans in colostrum (week 1) and mature milk (week 4) per Chinese mother (M1–7). Numbers indicate the *N*-glycans displayed in Table 1 and Figure 1. The numbers after the hyphen per mother indicates colostrum (1) and mature milk (2).

$Se^{-}Le^{+}$ mother			$Se^{+}Le^{+}$ mother			$Se^{+}Le^{+}$ mother			$Se^{-}Le^{+}$ mother			$Se^{+}Le^{+}$ mother			$Se^{+}Le^{+}$ mother			$Se^{+}Le^{+}$ mother		
No.	**M1-1**	**M1-2**	**No.**	**M2-1**	**M2-2**	**No.**	**M3-1**	**M3-2**	**No.**	**M4-1**	**M4-2**	**No.**	**M5-1**	**M5-2**	**No.**	**M6-1**	**M6-2**	**No.**	**M7-1**	**M7-2**
33	14,0	20,3	**16**	14,4	17,3	**25**	16,8	9,6	**33**	13,2	18,7	**25**	17,5	7,8	**16**	13,3	15,6	**25**	15,4	8,1
6	12,9	2,5	**25**	12,3	7,9	**16**	13,1	6,0	**6**	12,1	2,5	**16**	13,9	17,4	**25**	11,4	7,2	**16**	11,8	18,0
4	9,6	0,4	**35**	9,3	5,0	**35**	10,3	5,0	**4**	9,1	0,5	**35**	10,9	5,0	**35**	8,6	4,6	**35**	9,5	5,1
34	4,8	6,8	**26**	6,6	6,8	**26**	6,7	8,4	**34**	4,6	6,4	**26**	7,1	6,6	**26**	6,1	6,2	**26**	6,2	6,9
25	4,7	9,4	**33**	6,0	5,4	**33**	6,0	5,4	**25**	4,5	8,8	**33**	6,2	5,5	**33**	5,6	4,9	**33**	5,6	5,6
35	4,6	5,8	**11**	3,8	3,5	**6**	3,6	4,1	**35**	4,4	5,5	**6**	3,8	3,4	**11**	3,6	3,3	**6**	3,3	3,5
24	4,0	6,6	**6**	3,2	3,4	**41**	2,6	2,1	**24**	3,9	6,2	**36**	2,7	1,8	**6**	3,1	3,2	**41**	2,5	2,1
16	3,8	6,7	**41**	3,2	2,2	**24**	2,6	2,2	**16**	3,7	6,3	**24**	2,6	2,5	**41**	3,0	2,1	**24**	2,5	2,5
45	3,5	0,4	**36**	2,6	1,8	**36**	2,6	1,8	**45**	3,4	0,6	**41**	2,6	2,1	**36**	2,5	1,7	**36**	2,5	1,8
55	3,3	2,4	**24**	2,4	2,5	**44**	1,8	1,9	**55**	3,2	2,4	**44**	1,8	1,5	**24**	2,3	2,4	**44**	1,8	1,5
60	2,9	2,9	**23**	1,8	1,2	**23**	1,7	1,6	**60**	2,9	2,9	**23**	1,8	1,1	**23**	1,8	1,2	**23**	1,7	1,1
26	2,5	4,0	**44**	1,7	1,5	**11**	1,7	2,3	**26**	2,5	3,8	**11**	1,7	3,4	**44**	1,7	1,5	**11**	1,7	3,5
41	2,4	3,6	**66**	1,6	2,0	**34**	1,7	2,0	**41**	2,4	3,4	**34**	1,7	2,3	**66**	1,6	1,9	**34**	1,7	2,3
1	2,1	0,5	**34**	1,6	2,3	**15**	1,6	1,9	**1**	2,1	0,7	**66**	1,6	1,9	**34**	1,6	2,1	**15**	1,6	1,8
10	1,9	0,3	**50**	1,5	2,0	**66**	1,6	1,5	**10**	1,9	0,5	**15**	1,6	1,8	**50**	1,5	1,9	**66**	1,5	2,0
Total	77,2	72,7	**Total**	71,9	64,6	**Total**	74,4	55,8	**Total**	73,9	69,1	**Total**	77,5	64,2	**Total**	67,9	59,8	**Total**	69,3	65,6

Table S3. Missing *N*-glycans in this study as compared to Dallas et al. (2011)[14], Nwosu et al. (2012)[15], and Lu et al. (2019)[16].

Mass	Composition				Type
	Hex	HexNAc	Fuc	NeuAc	II
1405,3 [16]	3	3	2	0	NF
1567,6 [14]	4	3	2	0	NF
1712,6 [14, 15, 16]	4	3	1	0	NF
1745,6 [14]	6	3	1	0	NF
1890,6 [14, 15]	6	3	0	1	A
1915,7 [14]	4	4	1	1	AF
2030,7 [15]	4	6	1	0	NF
2217,8 [14]	3	7	2	0	NF
2222,8 [14, 15]	5	4	0	2	A
2223,8 [14, 15]	5	4	2	1	AF
2280,8 [14]	5	5	1	1	AF
2363,8 [14]	3	7	3	0	NF
2368,8 [14, 16]	5	4	1	2	AF
2369,9 [14, 15]	5	4	3	1	AF

Table S3 (continued)

Mass	Composition				Type
	Hex	HexNAc	Fuc	NeuAc	II
2516,9 [14]	7	6	1	0	NF
2588,9 [14, 16]	6	5	2	1	AF
2589,9 [14]	6	5	4	0	NF
2605,9 [14]	7	5	3	0	NF
2735,0 [14, 15]	6	5	3	1	AF
2736,0 [15]	8	7	0	0	N
2809,0 [14]	7	6	3	0	NF
2880,0 [14]	6	5	2	2	AF
2881,1 [15]	6	5	4	1	AF
2882,1 [16]	8	7	1	0	NF
2955,1 [14]	7	6	4	0	NF
3100,1 [14]	7	6	3	1	AF
3246,2 [14]	7	6	4	1	AF

Chapter 5

Human milk oligosaccharides in colostrum and mature milk of Chinese mothers: Lewis positive secretor subgroups

Published as: Elwakiel, M.; Hageman, J.A.; Wang, W.; Szeto, I.M.; van Goudoever, J.B.; Hettinga, K.A.; Schols, H.A. J. Agric. Food Chem. **2018**, 66, 7036–43.

Abstract

To study the variability in HMO composition of Chinese human milk over a 20-week lactation period, HMO profiles of 30 mothers were analysed using CE-LIF. This study showed that total HMO concentrations in Chinese human milk decreased significantly over a 20-week lactation period, independent of the mother's SeLe status, although with individual variations. In addition, total acidic and neutral HMO concentrations in Chinese human milk decreased over lactation, and their levels are driven by their mother's SeLe status. Analysis showed that total neutral fucosylated HMO concentrations in Chinese human milk were higher for the two secretor groups as compared to the Lewis positive nonsecretor group. On the basis of the total neutral fucosylated HMO concentrations in Chinese human milk, HMO profiles within the Se^+Le^+ group can be divided into two subgroups. HMOs that differed in level between these Se^+Le^+ subgroups were 2'FL, DF-L, LNFP I, and F-LNO. HMO profiles in Dutch human milk also showed Se^+Le^+ subgroup division, with 2'FL, LNT, and F-LNO as the driving force.

Keywords

Carbohydrates, variability, lactation stage, genetic polymorphisms

Introduction

Human milk is the natural food for infants after birth, providing not only nutrition but also protection against infectious diseases[1]. Human milk contains a variety of milk components like proteins, lipids, carbohydrates, which support the healthy growth and development of infants[2]. Specific protective components like oligosaccharides (HMOs) and immune-active proteins in human milk are present in higher concentrations in early lactation than in late lactation, while other nutritional components like lactose and fatty acids increase over lactation[3].

Lactose and HMOs are both part of the carbohydrate fraction in human milk[4,5]. Lactose is a disaccharide formed by a β1,4 linkage between galactose and glucose, and its concentrations in human milk range from 56 to 69 g/L over lactation[3-5], although with large individual variation. The enzyme lactase is present in the small intestine, and breaks down lactose into glucose and galactose[6-8], although lactose may end up in the colon at early life. HMOs are complex lactose-based glycans synthesized in the mammary gland throughout lactation[9-11]. HMOs are composed of five monosaccharides; glucose, galactose, *N*-acetylglucosamine, fucose, and *N*-acetylneuramic acid. During the synthesis of HMOs, lactose can be elongated with lacto-*N*-biose (β1,3 linkage) or with *N*-acetyllactosamine (β1,6 linkage), and these core HMO structures can be further decorated with fuc or sialic acid NeuAc residues[9-11]. HMOs and lactose are resistant to gastric and duodenal digestion, able to modulate the immune system of the intestinal mucosa, and influence the composition of the gut microbiome[12-17]. The size, structure, and function differ between HMOs[18]. More than 100 different structures have been identified and characterized in human milk, including many isomers[19]. Total HMO concentrations in human milk ranged from 5 to 25 g/L over a 6 months lactation period[19]. HMOs can be classified as neutral or acidic HMOs, with acidic oligosaccharides generally being present at a 10-fold lower concentration than neutral oligosaccharides[20,21].

The type and amount of HMOs present in human milk depend on the genetic profile of the mother, resulting in four major milk-types[22-25]. Fucosyltransferase (FUT) 2 is encoded by the Secretor (Se) gene and determines the presence of α1,2-fucosylated oligosaccharides in human milk. On the basis of the Lewis (Le) blood group system, FUT3 is encoded by the Le gene, which determines the presence of α1,4-fucosylated oligosaccharides in human milk. Women with an active Se locus are classified as secretor (Se^+), whereas women with an active Le locus are classified as Lewis positive (Le^+). Women without FUT2 or FUT3 activity are classified as nonsecretors (Se^-) or Lewis negative (Le^-), lacking α1,2-fucosylated or α1,4-fucosylated oligosaccharides, respectively. A large variation in HMO composition within the four major milk-type groups has been reported[26], and might be explained by mutations in the Se and Le genes[27]. Additional Se and Le phenotypes have been reported, the so-called weak Se and Le phenotype, respectively, mostly found in the Asian population[27], and less common in European population. Weak Se and Le phenotypes are probably able to produce

FUT2- and FUT3-mediated oligosaccharides, respectively, with fucosylated HMO levels lower than typically found in regular milk of Se and Le phenotypes[19]. For example, it has been reported that FUT2- and FUT3-mediated oligosaccharides, such as 2'fucosyllactose (2'FL) and 3FL, respectively, can be present in human milk in lower amounts[19]. HMO profiles were also shown to be different within and between breastfeeding populations from >10 countries[28-30]. Although human milk of most individuals can be grouped into four SeLe groups, there exists a large variation in HMO levels within SeLe groups[26,28,31], but none of these studies so far tried to find patterns in HMO profiles within the four milk-type groups.
The main objective of this study was to investigate the level and type of HMOs in Chinese human milk over a 20-week lactation period. HMO profiles of 30 mothers over the course of lactation were investigated using capillary electrophoresis-laser-induced fluorescence (CE-LIF). To investigate whether the observed clustering in HMO composition is typical for only Chinese mothers, HMO profiles of 28 Dutch mothers were determined 4 weeks after delivery.

Materials and methods

Setup of study and sample collection
Chinese participants were recruited between August 2014 and November 2015. The Yili Innovation Center (Hohhot, CN) took care of the human milk collection. Women living in the Hohhot region collected milk samples using a human milk pump. For every time point, a volume of 10 mL was collected in a polypropylene bottle. Milk bottles were shaken gently, aliquoted into 1 mL Eppendorf tubes, and stored at –20 °C. Milk samples of 30 mothers were assessed in week 1, 2, 4, 8, 12, and 20. Human milk collection was approved by the Chinese Ethics Committee of Registering Clinical Trials (ChiECRCT-20150017). Written informed consent was obtained for all of the Chinese participants. Dutch participants were recruited between September 2015 and June 2016. Human milk samples of women who gave birth at the obstetric department of the VU Medical Center in Amsterdam were collected by the Dutch Human Milk Bank. A volume of 10 mL was collected in a polypropylene bottle and stored at –20 °C. Milk of 28 Dutch mothers was collected, after 4 weeks of delivery. Human milk collection was approved by the VU Medical Center institutional committee and written informed consent was obtained from all mothers.

Sample preparation, labeling, and data analysis
HMOs were isolated and extracted from human milk, as described previously[20]. Defatting of the human milk samples was followed by protein precipitation and the pellet obtained after centrifugation containing denatured proteins was removed. HMOs present in the supernatant were isolated via solid phase extraction on graphitized carbon cartridges

(Alltech, Deerfield, U.S.). Subsequently, the isolated HMOs were labeled with fluorescent 9-aminopyrene-1,4,6-trisulfonate (APTS), as described previously[20]. During derivatization, oligosaccharides are linked in a molar ratio of 1:1 to the negatively charged label APTS. After labeling of the HMOs, the samples were analysed using CE-LIF, as described previously[20]. Samples were measured in triplicate, and xylose was used as the internal standard. HMOs were identified using commercially available standards, and the elution behavior of HMOs was identified in existing literature[20]. Quantification was done using the molar response factor of APTS labeled xylose, and concentrations compared nicely with known quantities of available HMOs measured. HMO standards 3'- and 6'-sialyllactose (SL) were bought from Sigma-Aldrich (St. Louis, U.S.). The HMO standards, 2'FL and 3FL, sialyllacto-*N*-tetraose (S-LNT), lacto-*N*-fucopentaose (LNFP) I–III, lacto-*N*-difucosylhexaose (LNDFH) I, fucosyllacto-*N*-hexaose (F-LNH) III, and lacto-*N*-hexaose (LNH) were purchased from Dextra (Reading, UK). Difucosyllactose (DF-L) was provided by Elicityl OligoTech (Crolles, FR), while lacto-*N*-tetraose (LNT) and disialyllacto-*N*-tetraose (DS-LNT) were purchased from Carbosynth (Berkshire, UK). The isomers LNFP I and III co-elute with CE-LIF and quantification was performed using the LNFP I standard, since this isomer is the dominant LNFP isomer present. Isomer LNFP II almost co-elutes on CE with DS-LNT and only the latter is shown in tables, since LNFP II is only a minor compound and elutes just before the DS-LNT peak. For data analysis, Chromeleon 7.1 (Thermo Fisher Scientific, Waltham, U.S.) was used. CE-LIF peak areas were converted to the corresponding HMO concentration in nanomoles g/L.

Statistical analysis

Total HMO concentrations in Chinese human milk over lactation were compared and correlated with maternal characteristics (age, parity, body mass index) and socioeconomic indicators (employment status and educational background) using SPSS (IBM Corp., New York. U.S.). The scales for educational background, as well as for employment status, were made from items of a three-point Likert scale. The scale for parity consisted of two. Participants did not have missing values for the categorical items in this study. Distributional aspects of the quantitative variables (age, body mass index, total HMO concentrations) were assessed by histograms (Gaussian distribution), QQ plots (normal distribution), Kolmogorov—Smirnov test (normal distribution), and by asymmetry and kurtosis values (between —3 and 3). The values of mother 8 at week 12 postpartum were excluded from analysis. The quantitative variables were assessed before regression analysis for linearity, univariate and bivariate outliers, and homoscedasticity, using scatterplot matrices, box plots, and residue plots, respectively. For statistical analysis, a *t*-test for independent samples, ANOVA, and multiple linear regression were used. The significance level was set at $\alpha = 0.05$. Human milk was assigned to their mother's SeLe status using 2'FL, LNFP I, LNDFH I, and LNT, as described previously[20]. The first three structures exclusively qualified the $Se^{+}Le^{+}$, $Se^{-}Le^{+}$,

and Se^+Le^- groups. In addition, average concentrations of LNT make a clear distinction between Se^+ and Se^- groups, which can be used as extra information next to the absence of α1,2-fucosylated or α1,4-fucosylated oligosaccharides in the Se^-Le^- group. Interpretation of the HMO profiles in human milk was facilitated by hierarchical clustering using R (Lucent Technologies, New York, U.S.), with Euclidean distance measure and Ward's linkage method. Hierarchical clustering was performed to detect and identify SeLe subgroups based on total, acidic, neutral, and individual HMO concentrations in Chinese human milk over a 20-week lactation period. HMO concentrations in Dutch human milk were evaluated in a similar way. The total HMO concentrations are based on 14 HMOs identified in this study, which are expected to present about 90% of all oligosaccharides present in human milk.

Results and discussion

Total lactose and HMO concentrations

To investigate the variability of lactose and HMOs in Chinese human milk over a 20-week lactation period, lactose and HMO profiles of 30 mothers were analysed using CE-LIF. Lactose concentrations were ranging from 40 to 85 g/L over a 20-week lactation period (Figure 1). Lactose levels in Chinese human milk increased in the first 4 weeks of lactation, then started to decline. Total HMO concentrations in Chinese human milk, as a sum of all individual HMOs (data not shown), were ranging from 8 to 23 g/L over lactation (Figure 1). The 14 HMOs identified in this study represent about 90% of the oligosaccharides present in human milk (data not shown). Human milk in early lactation (week 1 and 2) contained higher total HMO concentrations than in intermediate (week 4 and 8) and late lactation (week 12 and 20).

The lactose and total HMO concentrations in Chinese human milk over lactation (Figure 1) match with these observed in earlier studies, with average values of 56–69 g/L[3] and 5–25 g/L[18], respectively, with large individual variation. A change in lactose levels is expected in the first 2 weeks of lactation due to the general increase in nutritional components in milk[4,5]. Transition milk is produced from a couple of days up to 2 weeks postpartum, supporting the growth and development of the rapidly growing infant. It has been previously reported that levels of lactose levels are low in colostrum, increases in transition milk, and then remains constant in mature milk[5] , however lactose levels may be more variable in mature milk[4]. Milk becomes fully mature between 4 to 6 weeks postpartum, and contains higher amounts of nutrients as compared to bioactive components[4,5]. In early life, infants have an immature intestinal immune system, making them more vulnerable to infection by opportunistic pathogens in early lactation[1,2]. The high HMO level in colostrum may provide protection to the infant in this sensitive stage of its development[10,13].

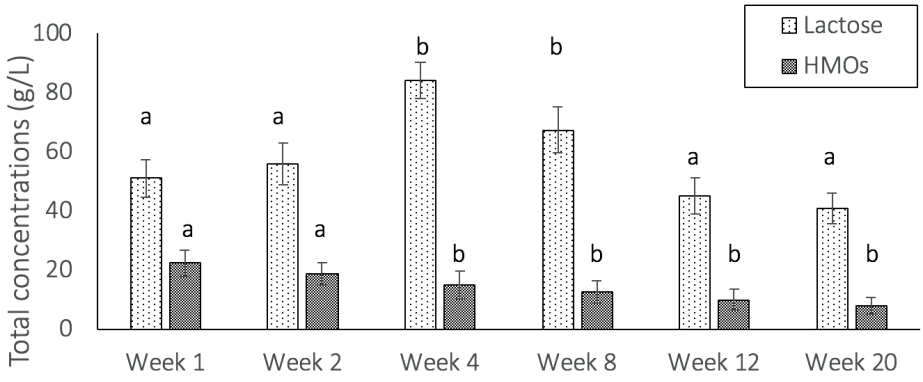


Figure 1. Total lactose and HMO concentrations (g/L) in Chinese human milk of 30 mothers over a 20-week lactation period. Error bars indicate the standard deviation. (a—b) Different alphabet letters indicate different lactose and HMO concentrations in human milk (two-tailed *t*-test, $\alpha < 0.05$) between different time points in lactation.

Total HMO concentrations in human milk of Chinese mothers over a 20-week lactation period, as presented in Figure 1, varied significantly among mothers (Figure 2). Although total HMO concentrations were always higher in early lactation than in intermediate and late lactation, the rate of decline varied among mothers. The total HMO concentrations, for example, mother 11 and 25 both started around 26 g/L, although showing the lowest (38%) and highest (85%) decline over lactation (Figure 2). As shown in Figure 2, the lowest concentrations in colostrum (9.9 g/L) and mature milk (3.7 g/L) were linked to mother 4, whereas the highest concentrations in colostrum (33.4 g/L) and mature milk (25.4 g/L) were found for mother 28 and 11, respectively.

Information collected from individual Chinese mothers and their total HMO concentrations are provided as supplementary information (Table S1). No correlation could be found by ANOVA and multiple regression analysis between the maternal characteristics (age, parity, and socioeconomic status) and the total HMO concentrations up to 20 weeks (results not shown). Body mass index seems to be positively correlated with total HMO concentrations at week 1 and 2, whereas no significant relationship was observed at later time points (results not shown). Total HMO concentrations were lower for mothers with a low body mass index in week 1 and 2. Several studies have suggested that mother's body mass index might influence the total HMO composition in human milk composition, especially colostrum[15,16,28], but the underlying mechanism is not yet clear.

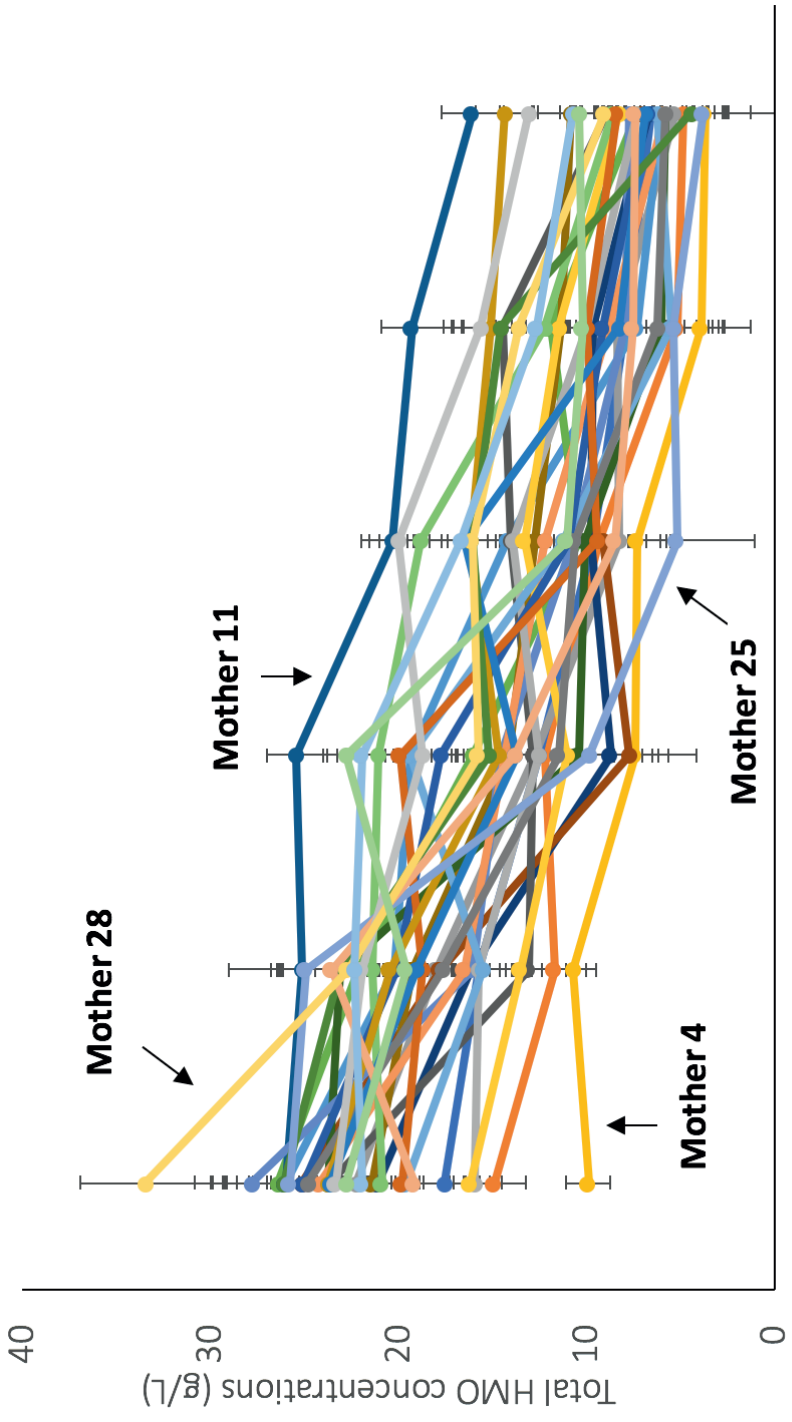


Figure 2. Total HMO concentrations (g/L) in Chinese human milk of 30 individual mothers over a 20-week lactation period. Error bars indicate the standard deviation.

Secretor and Lewis histo-blood group system

Milk samples were assigned to their mother's SeLe status (Table S1 and Figure S1). Twenty-two out of the 30 Chinese mothers can be assigned to the Se^+Le^+ group (73%), while 6 and 2 out of the 30 mothers were assigned to the Se^-Le^+ (20%) and Se^+Le^- (7%) groups, respectively. Milk samples from Se^-Le^- mothers were not present in this study. Distributions of these phenotypes vary among populations, and the frequency of the secretor phenotype in the Chinese population was previously estimated to be between 50% and 70%[29-30], which match with the findings in this study. Subsequently, total HMO concentrations in Chinese human milk for the three SeLe groups decreased over a 20-week lactation period (Figure 3), independent of the mother's SeLe status. The total HMO concentrations in Chinese human milk over a 20-week lactation period for the Se^+Le^+, Se^-Le^+, and Se^+Le^- group were 8.1–23.0 g/L, 6.5–20.0 g/L, and 9.4–23.5 g/L, respectively, and fall within the range of all combined SeLe groups over lactation[18].

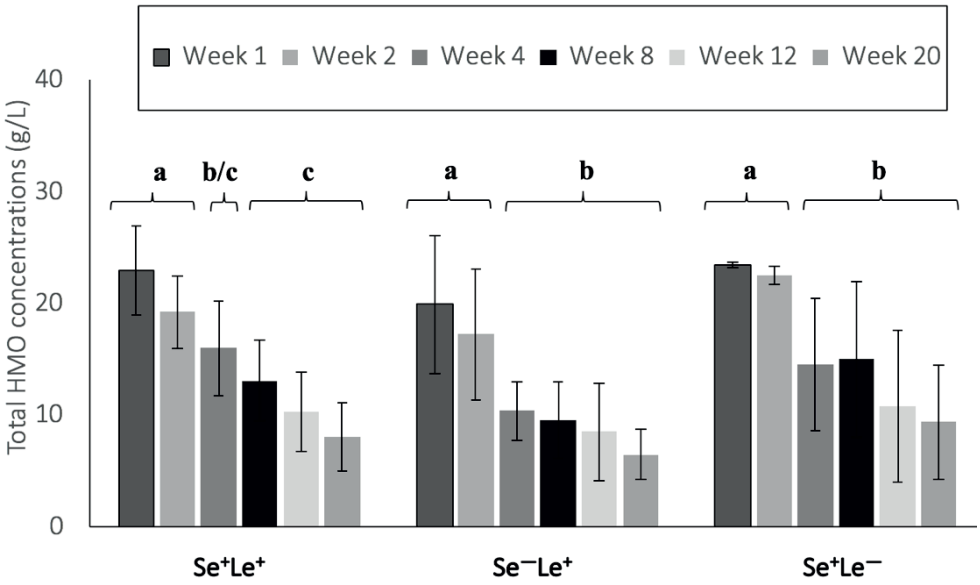


Figure 3. Total HMO concentrations (g/L) in Chinese human milk of 30 mothers over a 20-week lactation period categorized per SeLe group. Error bars indicate the standard deviation. Se^+Le^+ milk-type group n = 22, Se^-Le^+ milk-type group n = 6, and Se^+Le^- milk-type group n = 2. (a—c) Different alphabet letters indicate different HMO concentrations in human milk (two-tailed *t*-test, $\alpha < 0.05$) between different time points in lactation per SeLe group.

Acidic and neutral HMO concentrations

The total acidic and total neutral HMO fractions in human milk per mother and per time point in lactation are available in Table S2. For both the Se^+Le^+ (n = 22) and the Se^+Le^- (n = 2) groups (Figure 4A and C), the concentrations for the total neutral fucosylated HMO fraction decreased with 10.3 and 7.2 g/L over lactation, respectively, while the concentrations of the total acidic and neutral nonfucosylated HMO fractions even decreased relatively faster over lactation. For the Se^-Le^+ (n = 6) milk type, the concentrations for the total neutral nonfucosylated HMO fraction were decreasing the most with 7.6 g/L over lactation (Figure 4B). For the Se^+ groups, higher amounts were found for the total neutral fucosylated HMO fraction as compared to the Se^- group. Despite the absence of the FUT2 enzyme for the Se^-Le^+ group, and different profiles of three groups of HMOs in Chinese human milk over a 20-week lactation period, concentrations of the total neutral nonfucosylated HMO fraction might function as compensation, which possibly explains why the total HMO concentration ends up being the same for all genetic groups (Figure 3). However, having very few individuals in the Se^-Le^+ and Se^+Le^- milk-type groups complicates comparison between groups. The concentrations of the three groups of HMOs expressed in percentages in Chinese human milk for the Se^+Le^+ and Se^-Le^+ milk-type groups over lactation can be found in Figure S2.

The data of the Se^+Le^- milk-type group are not displayed in Figure S2, because it showed identical patterns over time with the Se^+Le^+ milk-type. The ratios between total acidic and total neutral HMO concentrations for the Se^+Le^+ milk-type and Se^-Le^+ milk-type group were ranging from 13:87 to 12:88 and from 28:72 to 40:60 over lactation (Figure S2), respectively, indicating that acidic HMOs over time might be relatively more dominant in the Se^-Le^+ milk-type group than in the Se^+Le^+ milk-type group. Overall, total acidic and total neutral HMO concentrations in Chinese human milk per SeLe group vary over the course of lactation, with overall higher total neutral HMO concentration in all groups.

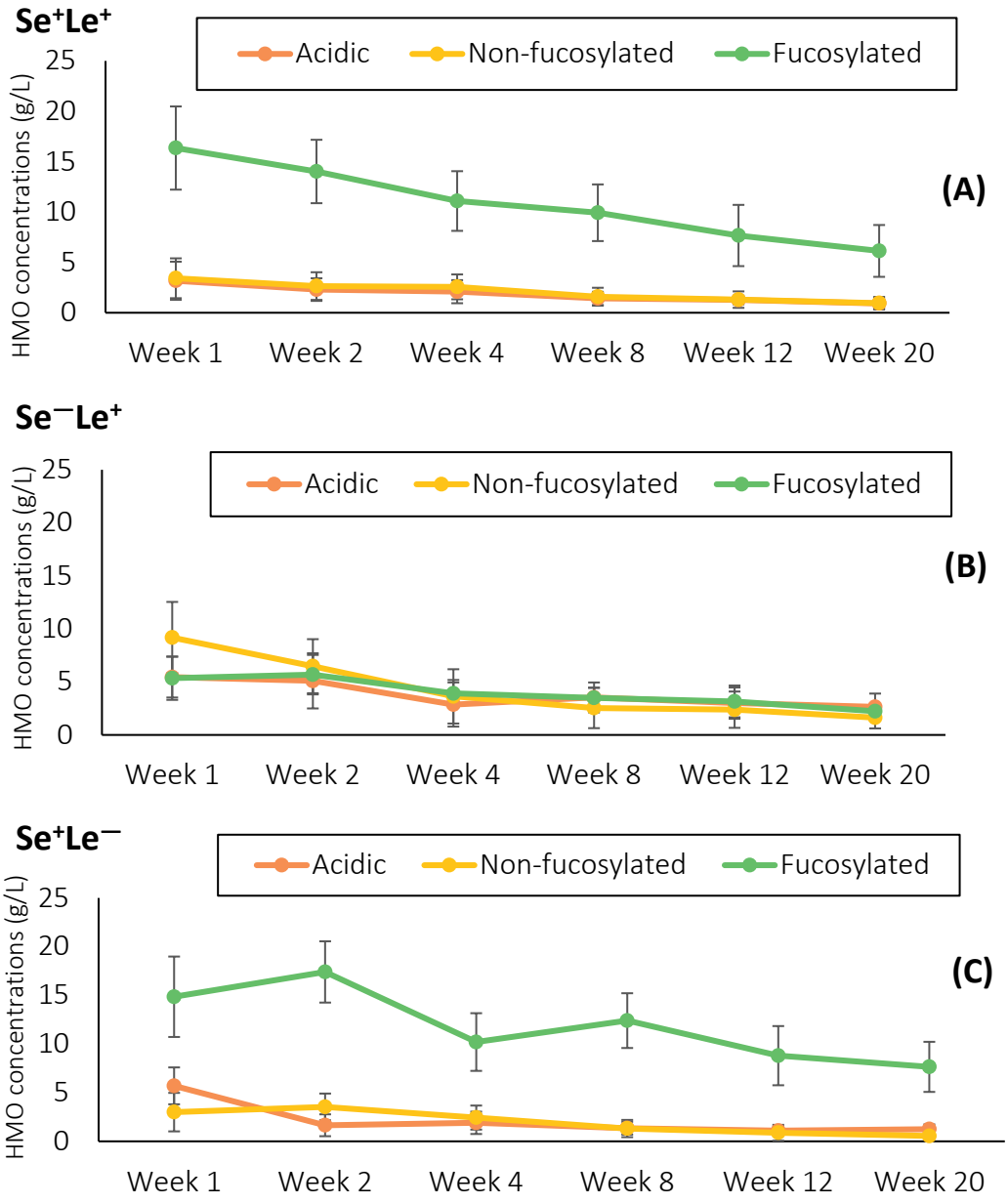


Figure 4. Concentrations of total acidic and total neutral (nonfucosylated and fucosylated) HMO fractions in Chinese human milk over a 20-week lactation period for **(A)** $Se^{+}Le^{+}$ milk-type group n = 22, **(B)** $Se^{-}Le^{+}$ milk-type group n = 6, and **(C)** $Se^{+}Le^{-}$ milk-type group n = 2.

Se^+Le^+ subgroups in Chinese human milk

To investigate the observed variability in total acidic and total neutral (nonfucosylated and fucosylated) HMO concentrations in human milk of Chinese mothers for the three SeLe groups over a 20-week lactation period, total acidic and total neutral HMO concentrations per mother were examined by clustering analysis. Statistical analysis confirmed the clear difference that exists between Se^- and Se^+ groups (Figure 5, cluster I/II versus III). However with concentrations of the total acidic and total neutral (nonfucosylated and fucosylated) HMO fractions in human milk per mother, Se^+Le^+ mothers could be divided into two subgroups (Figure 5, cluster I and II). The size of the Se^-Le^+ (n = 6) and Se^+Le^- (n = 2) groups was too small to detect any subgroups. Cluster III consisted only of Se^-Le^+ mothers. Milk from the two mothers having Se^+Le^- could not be clustered and end up in the Se^+Le^+ group (Figure 5).

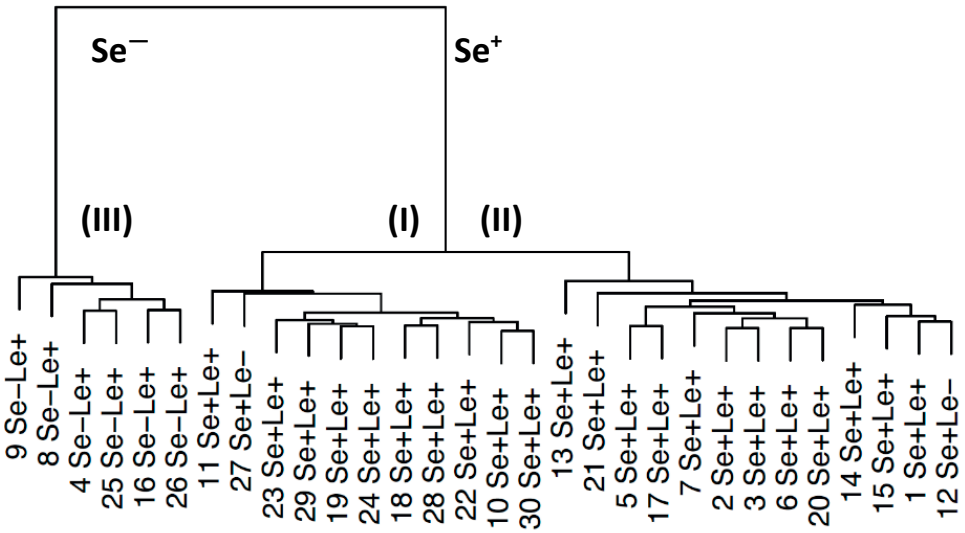


Figure 5. Hierarchical cluster analysis of total acidic and total neutral (nonfucosylated and fucosylated) HMO concentrations (g/L) in Chinese human milk per mother over a 20-week lactation period. **(I)** Se^+Le^+ milk-type group I, **(II)** Se^+Le^+ milk-type group II, and **(III)** Se^-Le^+ milk-type.

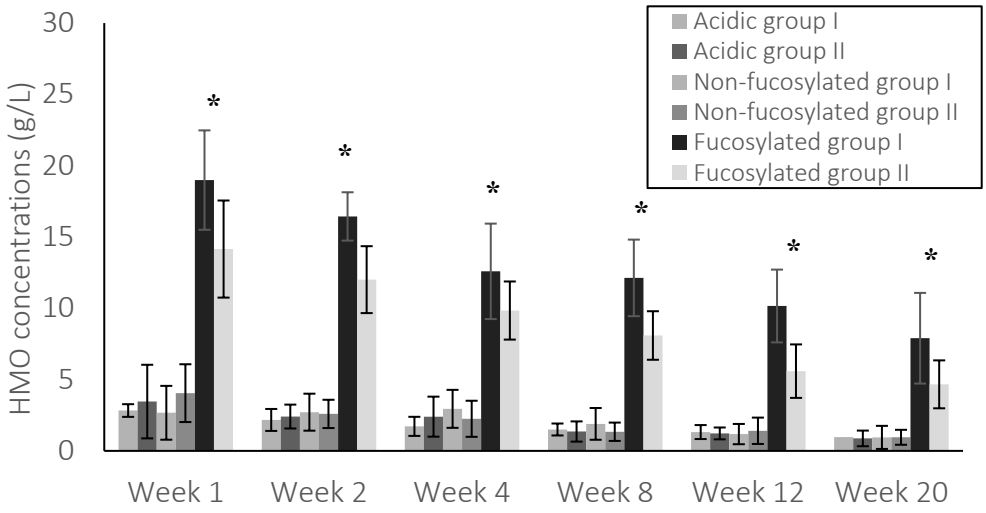


Figure 6. Concentrations of the total acidic and total neutral (fucosylated and nonfucosylated) HMO fractions for the two Se^+Le^+ subgroups in Chinese human milk over a 20-week lactation period. The Se^+Le^+ milk-type group (22 of the 30 mothers, 73%) can be divided into group I = 12 (40%) and group II = 10 (33%). (*) indicates significant differences (two-tailed *t*-test, $\alpha < 0.05$).

The 2 Se^+Le^+ subgroups, displayed in Figure 5, seem to be distinguished by their total neutral fucosylated HMO fraction (Figure 6). The concentrations of the total neutral fucosylated HMO fraction are significantly different and are almost 20% higher in subgroup I than in subgroup II (Figure 6). The concentrations of the total acid and total neutral nonfucosylated HMO fraction did not differ significantly between the 2 Se^+Le^+ subgroups.

The phenomena of the Se^+Le^+ subgroup formation might be a consequence of the observation that Se and Le genes can contain mutations[27]. Besides the full absence of FUT2 and FUT3, two different phenotypes have been found, so-called weak Se and Le phenotypes, respectively[27]. Because of modifications in the amino acid sequence, the activity of the FUT2 or FUT3 enzyme can be reduced, thereby possibly leading to a decrease in the synthesis of HMOs in one of the subgroups[19]. From the table containing all individual HMO concentrations (data not shown) it could be deducted that HMOs that differed between Se^+Le^+ subgroups were 2'FL, DF-L, LNFP I, and fucosyllacto-*N*-octaose (F-LNO) (Figure 7), having in common α1,2-fucosylated linkages to the core HMO structures.

The reason for the variation in these specific HMOs in this study (Figure 7) is not yet clarified. However, previous studies have reported that levels of 2'FL and LNFP I were found below normal ranges in human milk from a small group of Chinese participants[19]. Gene mutations

are not limited to the FUT2 enzyme activity, because various mutations have also been reported in the Le gene encoding for the FUT3 enzyme. In human milk collected from the Chinese mothers, variation based on FUT3-mediated oligosaccharides in Se^+Le^+ and Se^-Le^+ group could not be seen. Additionally, it has been noticed that 3FL, as indicator for the FUT3 enzyme, was removed in the pretreatment step, and therefore 3FL was not able to be detected in the samples of this study.

Se^+Le^+ subgroups in Dutch human milk

To investigate whether the observed differentiation in Se^+Le^+ subgroups in Chinese human milk also applies to other populations, HMO profiles from 28 Dutch mothers were collected 4 weeks after delivery and analysed. Total HMO concentrations measured in Dutch human milk, as a sum of the 14 HMOs (data not shown), ranged from 4 to 27 g/L 4 weeks postpartum (Table S3), independent of the mother's SeLe status and body mass index (results not shown). Milk samples were also assigned to their mother's SeLe status (Table S3). Fourteen Dutch mothers can be assigned to the Se^+Le^+ group (50%), while 11 and 3 mothers are identified as belonging to the Se^-Le^+ (39%) and Se^+Le^- (11%) groups, respectively. Milk samples from Se^-Le^- mothers were not present in this study. The distribution over the four SeLe groups for the Dutch mothers did not correspond with previously reported numbers reporting 70–80% being Se^+Le^+ for the European population[21-25] and 80% for the Dutch population[20]. These unexpected proportions should not affect the analysis, as the group for Se^+Le^+ mothers was large enough to perform cluster analysis, and the observed uncommon ratio between SeLe groups made it even possible to investigate the Se^-Le^+ group in more detail. After clustering analysis, HMO levels in milk of 28 Dutch mothers also showed Se^+Le^+ subgroup division (Figure 8). Two Se^+Le^+ mothers (8%) did not fall in either the Se^+Le^+ subgroup I or II (Figure 8). The Se^-Le^+ milk-type (39%) can be roughly divided in group III = 6 (21%), group IV = 2 (11%), and group V = 3 (7%). As shown in Figure 8, there is a lot of variation in HMO concentrations in Dutch human milk for the Se^-Le^+ group (cluster III—V). However, no significant difference could be found in concentrations of the total acidic and total neutral (nonfucosylated and fucosylated) HMO fractions between the Se^-Le^+ subgroups (data not shown).

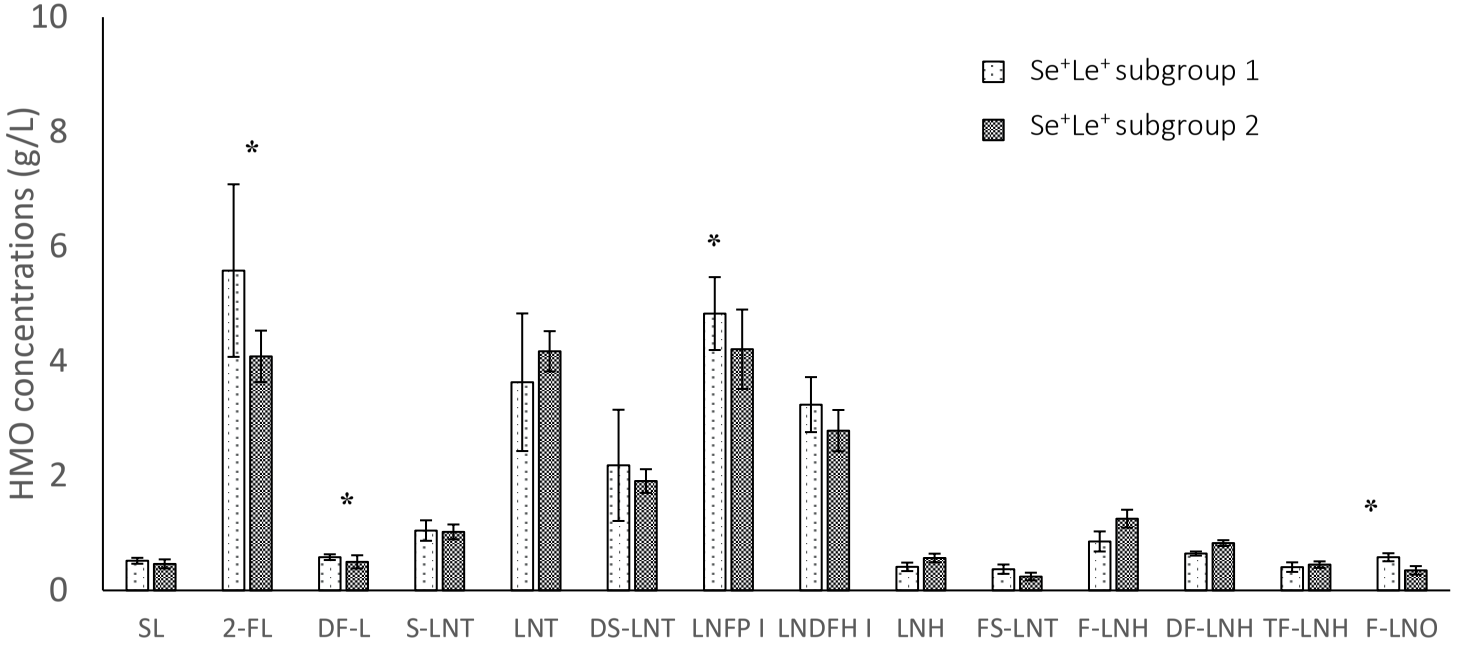


Figure 7. HMO concentrations (g/L) in Chinese human milk over a 20-week lactation period per Se^+Le^+ subgroup. **(I)** Se^+Le^+ milk-type group I, **(II)** Se^+Le^+ milk-type group II. (*) indicates significant differences (two-tailed *t*-test, $\alpha < 0.05$). Other than 2'FL, LNFP I, and LNDFH I, isomers are not further specified. Fucosyl-sialyllacto-*N*-tetraose (FS-LNT), difucosyllacto-*N*-hexaose (DF-LNH), trifucosyllacto-*N*-hexaose (TF-LNH).

Milk of the Dutch mothers categorized in the Se^+Le^+ group (Figure 8) can be divided into two subgroups (I and II) on the basis of the concentrations of the neutral fucosylated HMO fraction (Figure 9), like it was done for Se^+Le^+ group in Chinese human milk (Figure 6), however, with 2'FL, LNT, and F-LNO contributing to the differentiation more than the other HMOs (Figure 10).

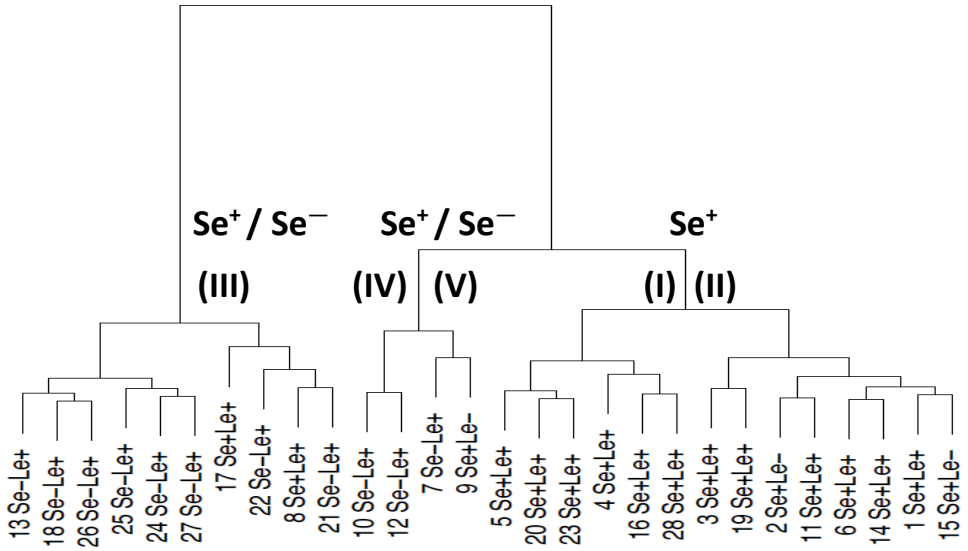


Figure 8. Hierarchical clustering analysis based on concentrations of the total acidic and total neutral (nonfucosylated and fucosylated) HMO fractions (g/L) in Dutch human milk per mother collected after 4 weeks of delivery.

Although the levels of HMOs, like DF-L and LNFP I, do not differ significantly between the Se^+Le^+ subgroups in Dutch human milk (Figure 10), a trend was visible that concentrations were slightly higher for Se^+Le^+ subgroup I as compared to subgroup II, which was much more clear for Chinese human milk. Levels of DF-L and LNFP I were significantly different between the Se^+Le^+ subgroups in Chinese human milk, also higher in Se^+Le^+ subgroup I than in subgroup II (Figure 7). Subsequently, concentrations for LNT in Dutch human milk were significantly higher for Se^+Le^+ subgroup II than subgroup I (Figure 10), such a trend could also be observed in Chinese human milk (Figure 7), although not significantly different. Overall, FUT2-mediated HMO structures play a key role in the differentiation between the subgroups in both Chinese and Dutch human milk, indicating that enzyme activity may be reduced for the FUT2 enzyme due to polymorphism.

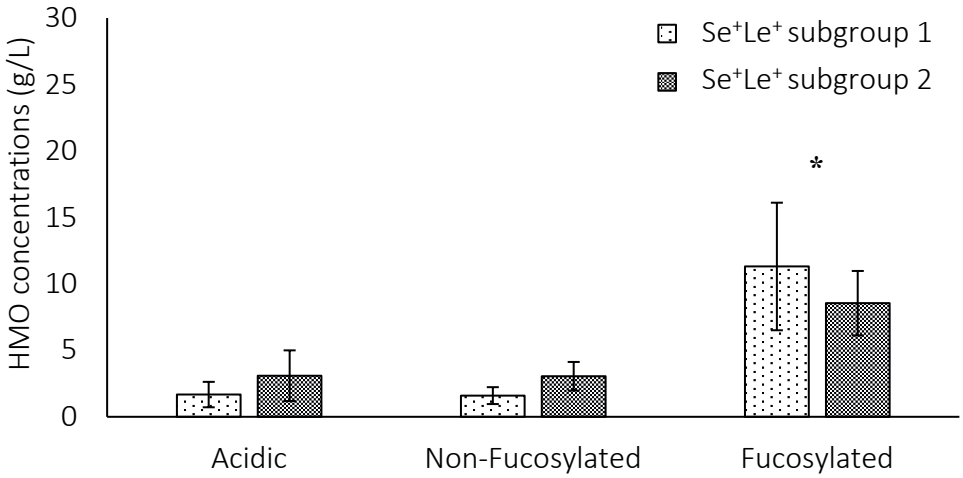


Figure 9. Concentrations of the total acidic and total neutral (fucosylated and nonfucosylated) HMO fractions for the two Se^+Le^+ subgroups in Dutch human milk collected 4 weeks postpartum. The Se^+Le^+ milk-type group (14 of the 28 mothers, 50%) can be divided into group I = 6 (21%), group II = 6 (21%), and group IV = 2 (8%). The 2 Se^+Le^+ mothers (8%), which could not be grouped into Se^+Le^+ subgroup I and II, were excluded from comparison. (*) indicates significant differences (two-tailed *t*-test, $\alpha < 0.05$).

This study tried to fill a gap in the literature by trying to recognize subgroups with statistics and highlight the variability in HMO composition in Chinese human milk of 30 mothers over a 20-week lactation period. This study showed that total HMO concentrations in Chinese human milk are not driven by their mother's SeLe status, but ratios of the total acidic and total neutral HMO fractions in human milk of Chinese mothers are responsible for the clustering. On the basis of the neutral fucosylated HMO fraction, Se^+Le^+ subgroups were recognized. To investigate whether the observed variability in HMO composition is typical for only Chinese mothers, HMO profiles of 28 Dutch mothers 4 weeks postpartum were investigated, which resulted in Se^+Le^+ subgroups, based on the concentrations of the neutral fucosylated HMO fraction, although with distinctive HMOs having a different concentration for the two subgroups.

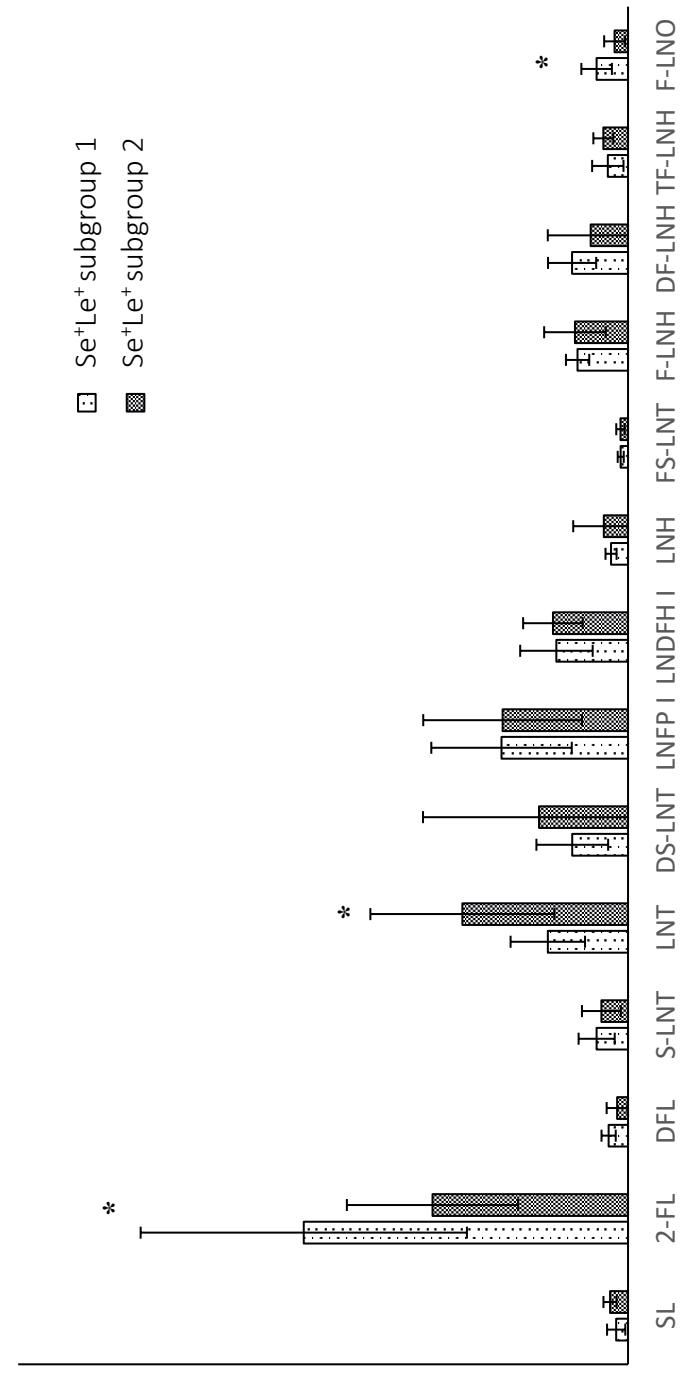


Figure 10. HMO concentrations (g/L) of the two Se^+Le^+ subgroups in Dutch human milk after 4 weeks of delivery. **(I)** Se^+Le^+ milk-type group I, **(II)** Se^+Le^+ milk-type group II. (*) indicates significant differences (two-tailed *t*-test, $\alpha < 0.05$). Other than 2'FL, LNFP I, and LNDFH I, isomers are not further specified.

References

1. Leung, A.; Sauve, R. Breast is best for babies. J. Natl. Med. Assoc. **2005,** 97, 1010–9.
2. Labbok, M.; Clark, D.; Goldman, A. Breastfeeding: Maintaining an irreplaceable immunological resource. Nat. Rev. Immunol. **2004**, 4, 565–72.
3. Coppa, G.; Gabrielli, O.; Pierani, P.; Catassi, C.; Carlucci, A.; Giorgi, P. Changes in carbohydrate composition in human milk over 4 months of lactation. Pediatrics, **1993**, 91, 637–41.
4. Newburg, D. Glycobiology of human milk. Biochemistry, **2013**, 78, 771–85.
5. Ballard, O.; Morrow, A. Human milk composition: Nutrients and bioactive factors. Pediatr. Clin. North Am. **2013**, 60, 49–74.
6. Schaafsma, G. Lactose and lactose derivatives as bioactive ingredients in human nutrition. Int. Dairy. J. **2008**, 18, 458–65.
7. Tran, H.; Moreno, R.; Hinkle, E.; Bundy, J.; Walter, J.; Burkey, T.; Miller, P. Effect of lactose and yeast-dried milk on growth performance, fecal microbiota, and immune parameters of nursery pigs. J. Anim. Sci. **2012**, 90, 3049–59.
8. Cederlund, A.; Kai-Larsen, Y.; Printz, G.; Yoshio, H.; Alvelius, G.; Lagercrantz, H.; Strömberg, R.; Jörnvall, H.; Gudmundsson, G.; Agerberth, B. Lactose in human breast milk an inducer of innate immunity with implications for a role in intestinal homeostasis. PLoS One, **2013**, 8, 1–12.
9. Stahl, B.; Thurl, S.; Zeng, J.; Karas, M.; Hillenkamp, F.; Steup, M.; Sawatzki, G. Oligosaccharides from human milk as revealed by matrix-assisted laser desorption/ionization mass spectrometry. Anal. Biochem. **1994**, 223, 218–26.
10. Kunz, C.; Rudloff, S.; Baier, W.; Klein, N.; Strobel S. Oligosaccharides in human milk: Structural, functional, and metabolic aspects. Annu. Rev. Nutr. **2000**, 20, 699–722.
11. Boehm, G.; Stahl, B. Oligosaccharides from milk. J. Nutr. **2007**, 137, 847–9.
12. Newburg, D. Oligosaccharides in human milk and bacterial colonization. J. Pediatr. Gastroenterol. Nutr. **1999**, 30, 8–17.
13. Kunz, C.; Rudloff, S. Health promoting aspects of milk oligosaccharides. Int. Dairy J. **2006**, 16, 1341–6.
14. Bode, L.; Jantscher-Krenn, E. Structure-function relationships of human milk oligosaccharides. Adv. Nutr. **2012**, 3, 383–91.
15. Lewis, Z.; Totten, S.; Smilowitz, J.; Popoviv, M.; Parker, E.; Lemay, D.; Van Tassell, M.; Miller, M.; Jin, Y.; German, B.; Lebrilla, C.; Mills, D. Maternal fucosyltransferase 2 status affects the gut *bifidobacterial* communities of breastfed infants. Microbiome, **2015**, 11, 1–21.
16. Sprenger N.; Odenwald H.; Kukkonen A.K.; Kuitunen M.; Savilahti E.; Kunz C. FUT2-dependent breast milk oligosaccharides and allergy at 2 and 5 years of age in infants with high hereditary allergy risk. Eur. J. Nutr. **2017**, 56, 1293–1301.
17. Seppo, A.; Austran, C.; Bode, L.; Järvinen, K.; Human milk oligosaccharides and development of cow's milk allergy in infants. J. Allergy Clin. Immunol. **2017**, 139, 708–11.
18. Bode, L. Human milk oligosaccharides: Every baby needs a sugar mama. Glycobiology, **2012**, 22, 1147–62.
19. Austin, S.; De Castro, C.; Benet, T.; Hou, Y.; Sun, H.; Thakkar, S.; Vinyes-Pares, G.; Zhang, Y.; Wang, P. Temporal change of the content of 10 oligosaccharides in the milk of Chinese urban mothers. Nutrients, **2016**, 8, 346–68.

20. Albrecht, S.; Schols, H.A.; van den Heuvel, E.; Voragen, A.G.; Gruppen, H. Occurrence of oligosaccharides in faeces of breastfed babies in their first six months of life and the corresponding breast milk. Carbohydr. Res. **2011**, 346, 2540–50.
21. Thurl, S.; Munzert, M.; Henker, J.; Boehm, G.; Muller-Werner, B; Jelinek, J.; Stahl, B. Variation of human milk oligosaccharides in relation to milk groups and lactational periods. Br. J. Nutr. **2010**, 104, 1261–71.
22. Grollman, E.; Kobata, A.; Ginsburg, V. An enzymatic basis for Lewis blood types in man. J. Clin. Invest. **1969**, 48, 1489–94.
23. Thurl, S.; Henker, J.; Siegel, M.; Tovar, K.; Sawatzki, G. Detection of four human milk groups with respect to Lewis blood group dependent oligosaccharides. Glycoconjugate. J. **1997**, 14, 795–9.
24. Newburg, D. Are all human milks created equal? Variation in human milk oligosaccharides. J. Pediatr. Gastroenterol. Nutr. **2000**, 30, 131–3.
25. Blank, D.; Dotz, V.; Geyer, R.; Kunz, C. Human milk oligosaccharides and Lewis blood group: Individual high-throughput sample profiling to enhance cconclusions from functional studies. Adv. Nutr. **2012**, 3, 440–9.
26. Prieto, P. Profiles of human milk oligosaccharides and production of some human milk oligosaccharides in transgenic animals. Adv. Nutr. **2012**, 3, 456–64.
27. Schenkel-Brunner, H. Human blood groups: Chemical and biochemical basis of antigen specificity. Springer: Vienna, Austria, **2000**, pp. 205–12. (Ed.)
28. McGuire, M. What's normal? Oligosaccharide concentrations and profiles in milk produced by healthy women vary geographically. Am. J. Clin. Nutr. **2017**, 105, 1086–100.
29. Erney, R.; Malone, W.; Skelding, M.; Marcon, A.; Kleman-Leyer, K.; O'Ryan, M.; Ruiz-Palacios, G.; Hilty, M.; Pickering, L.; Prieto, P. Variability of human milk neutral oligosaccharides in a diverse population. J. Pediatr. Gastroenterol. Nutr. **2000**, 30, 181–92.
30. Castanys-Muoz, E.; Martin, M; Prieto, P. 2'fucosyllactose: An abundant, genetically determined soluble glycan present in human milk. Nutr. Rev. **2013**, 71, 773–89.
31. Sprenger, N.; Lee, L.; De Castro, C.; Steenhout, P.; Thakkar, K. Longitudinal change of selected human milk oligosaccharides and association to infants' growth, an observatory, single center, longitudinal cohort study. PLoS One, **2017**, 12, 1–15.

Supporting information

Table S1.

Total HMO concentrations (g/L) in human milk of 30 Chinese mothers over a 20-week lactation period.

Mother No.	Total HMO concentrations (g/L) (weeks)						SeLe status*	Age	Body mass index kg/m^2	Interpretation**	Educational background	Employment status	Parity
	1	2	4	8	12	20							
1	14.9	11.7	12.2	9.3	5.2	4.8	Se^+Le^+	29	22.3	Normal	Master	Middle	1
2	17.5	15.7	12.5	9.6	7.9	6.5	Se^+Le^+	31	25.4	Pre-obese	Junior high school	Low	2
3	22.2	18.0	12.7	8.2	8.4	5.3	Se^+Le^+	24	23.6	Overweight	Senior high school	Low	1
4	9.9	10.7	7.4	7.3	3.9	3.7	Se^-Le^+	27	22.4	Normal	College	Middle	1
5	26.0	20.2	19.4	14.2	7.3	6.0	Se^+Le^+	25	20.3	Normal	Senior high school	Low	2
6	26.4	22.1	16.0	10.4	11.9	7.3	Se^+Le^+	28	21.3	Normal	Master	Middle	1
7	21.1	16.2	8.7	9.7	9.6	6.6	Se^+Le^+	35	20.5	Normal	Technical school	Middle	2
8	25.1	17.8	7.7	9.1	-	7.0	Se^-Le^+	27	18.7	Normal	Bachelor	Middle	1
9	23.6	13.1	12.8	14.0	14.5	8.9	Se^-Le^+	29	20.0	Normal	Master	Middle	1
10	21.4	19.3	14.1	12.7	11.4	10.8	Se^+Le^+	33	23.6	Overweight	Senior high school	Low	2
11	26.0	25.1	25.4	20.3	19.3	16.1	Se^+Le^+	28	17.6	Underweight	Bachelor	Middle	1
12	23.7	23.1	10.4	10.1	6.0	5.8	Se^+Le^-	30	19.6	Normal	Junior high school	Low	2
13	27.8	16.1	14.6	10.7	7.9	7.7	Se^+Le^+	30	22.7	Normal	Bachelor	High	1
14	24.2	16.5	14.3	12.2	8.8	5.8	Se^+Le^+	30	21.2	Normal	College	Middle	1
15	15.9	15.7	12.5	13.9	10.0	7.3	Se^+Le^+	28	22.9	Normal	Bachelor	Low	1

Table S1 (continued)

Mother No.	Total HMO concentrations (g/L) (weeks) 1	2	4	8	12	20	SeLe status*	Age	Body mass index kg/m^2	Interpretation**	Educational background	Employment status	Parity
16	16.2	13.5	11.0	13.3	11.4	8.2	$Se^{-}Le^{+}$	36	23.0	Overweight	Primary school	Low	2
17	19.6	15.5	19.4	11.2	5.5	6.2	$Se^{+}Le^{+}$	30	18.8	Normal	Junior high school	Low	1
18	20.9	21.4	21.1	18.8	12.2	8.6	$Se^{+}Le^{+}$	40	28.0	Pre-obese	Primary school	Low	2
19	25.0	20.1	17.8	11.0	9.2	6.7	$Se^{+}Le^{+}$	29	20.4	Normal	College	Middle	1
20	19.8	18.7	19.9	9.4	10.0	8.4	$Se^{+}Le^{+}$	33	25.4	Pre-obese	Bachelor	Middle	1
21	24.8	17.6	11.5	10.6	6.2	5.8	$Se^{+}Le^{+}$	28	18.4	Underweight	Bachelor	High	1
22	23.7	20.5	14.7	16.2	15.1	14.3	$Se^{+}Le^{+}$	31	29.3	Pre-obese	Technical school	Middle	1
23	23.6	19.0	13.8	16.4	8.2	6.8	$Se^{+}Le^{+}$	29	18.0	Underweight	Bachelor	Middle	1
24	26.1	23.2	15.2	16.1	14.5	4.4	$Se^{+}Le^{+}$	28	20.4	Normal	Bachelor	Middle	1
25	25.8	24.9	9.8	5.2	5.4	3.8	$Se^{-}Le^{+}$	32	23.3	Overweight	Master	Middle	1
26	19.2	23.6	13.8	8.5	7.5	7.4	$Se^{-}Le^{+}$	27	18.3	Underweight	Senior high school	Middle	1
27	23.4	22.0	18.7	20.0	15.6	13.0	$Se^{+}Le^{-}$	32	26.4	Pre-obese	Junior high school	Low	2
28	33.4	22.7	15.8	16.1	13.6	9.1	$Se^{+}Le^{+}$	23	19.5	Normal	Senior high school	Low	1
29	22.0	22.4	22.0	16.7	12.7	10.7	$Se^{+}Le^{+}$	28	25.0	Pre-obese	Junior high school	Middle	2
30	22.7	19.6	22.7	11.1	10.2	10.4	$Se^{+}Le^{+}$	28	23.4	Overweight	Bachelor	High	1

* Secretor, Lewis positive ($Se^{+}Le^{+}$) n = 22; non-secretor, Lewis positive ($Se^{-}Le^{+}$) n = 6; secretor, Lewis negative ($Se^{+}Le^{-}$) n = 2; non-secretor, Lewis negative ($Se^{-}Le^{-}$) n = 0. ** International standard for body mass index adapted for Asian people: < 18.5 kg/m^2 (underweight), 18.5 - 22.9 kg/m^2 (normal), 22.9 - 24.9 kg/m^2 (overweight), ≥ 25.0 kg/m^2 (pre-obese), and ≥ 30 kg/m^2 (obese).

Table S2.

Total acidic and total neutral HMO concentrations (g/L) in human milk of 30 Chinese mothers over a 20-week lactation period.

		Acidic and neutral HMO concentrations**																	
		week 1			week 2			week 4			week 8			week 12			week 20		
Mother No.	**SeLe status***	**A**	**B**	**C**	**A**	**B**	**C**	**A**	**B**	**C**	**A**	**B**	**C**	**A**	**B**	**C**	**A**	**B**	**C**
1	Se^+Le^+	1.5	2.1	11.0	1.2	2.1	8.4	1.1	3.3	7.8	0.8	1.9	6.7	0.3	0.7	4.3	0.5	0.4	3.9
2	Se^+Le^+	1.4	5.2	10.7	1.4	2.3	12.0	1.6	1.2	9.7	1.1	1.3	7.1	1.1	0.9	5.9	0.6	0.9	5.0
3	Se^+Le^+	3.0	5.3	14.0	2.5	2.7	12.9	1.7	1.6	9.5	1.4	1.3	5.6	1.4	1.3	5.8	1.1	1.1	3.2
4	Se^-Le^+	2.6	5.0	2.3	3.4	3.2	4.1	2.5	1.9	3.1	2.5	1.9	2.9	1.4	0.9	1.6	1.3	0.9	1.5
5	Se^+Le^+	4.7	3.1	18.2	2.0	1.1	17.1	4.1	3.6	11.7	2.5	1.5	10.3	1.2	0.8	5.3	1.0	0.5	4.5
6	Se^+Le^+	4.9	3.3	18.3	4.0	4.6	13.6	2.9	2.0	11.2	1.7	1.1	7.6	1.6	1.6	8.7	1.2	0.8	5.3
7	Se^+Le^+	3.1	2.5	15.6	1.8	2.0	12.4	1.3	1.0	6.5	1.3	1.1	7.3	1.3	1.1	7.2	1.9	1.9	2.8
8	Se^-Le^+	7.3	9.7	8.2	3.9	8.4	5.5	1.0	1.1	5.5	4.9	0.8	3.5	-	-	-	4.3	0.7	2.0
9	Se^-Le^+	4.0	13.8	5.9	2.1	6.9	4.1	2.3	8.4	2.0	3.7	6.0	4.3	3.9	5.1	5.5	2.4	3.3	3.2
10	Se^+Le^+	2.3	2.7	16.4	1.5	1.9	15.9	1.3	1.9	10.9	1.2	1.7	9.7	1.1	0.8	9.5	1.3	0.6	9.0
11	Se^+Le^+	2.1	7.8	16.1	2.8	5.9	16.4	3.1	5.3	17.0	2.0	3.8	14.5	2.2	2.7	14.4	1.3	2.9	12.0
12	Se^+Le^-	1.7	4.9	17.1	1.5	6.3	15.4	1.3	2.9	6.2	0.6	2.1	7.5	1.0	0.9	4.1	0.5	0.5	4.9
13	Se^+Le^+	10.1	8.0	9.6	1.3	1.0	13.8	5.0	3.6	6.0	0.5	1.7	8.5	2.0	2.3	3.6	0.0	1.9	5.8
14	Se^+Le^+	2.7	6.1	15.4	1.7	2.3	12.5	1.2	0.9	12.3	0.5	0.5	11.2	1.4	3.6	3.8	0.6	0.7	4.5
15	Se^+Le^+	1.2	1.7	13.0	1.0	1.7	13.0	2.1	1.1	9.3	1.7	2.9	9.3	1.3	2.3	6.4	0.2	1.2	6.0

Table S2 (continued)

		Acidic and neutral HMO concentrations**																	
		week 1			week 2			week 4			week 8			week 12			week 20		
Mother No.	**SeLe status***	**A**	**B**	**C**	**A**	**B**	**C**	**A**	**B**	**C**	**A**	**B**	**C**	**A**	**B**	**C**	**A**	**B**	**C**
16	$Se^{-}Le^{+}$	4.9	7.3	4.0	5.1	3.5	4.9	3.2	3.2	4.6	5.1	3.3	4.9	4.9	3.1	3.5	3.6	2.4	2.2
17	$Se^{+}Le^{+}$	3.8	4.0	11.8	2.4	1.6	11.5	4.2	4.4	10.8	2.4	1.2	7.6	1.2	0.8	3.5	1.4	0.7	4.1
18	$Se^{+}Le^{+}$	2.1	1.8	17.0	1.9	2.2	17.2	2.0	1.8	17.3	1.9	1.2	15.6	1.2	0.4	10.6	1.0	0.5	7.1
19	$Se^{+}Le^{+}$	2.4	1.7	21.0	2.3	2.7	15.1	2.1	3.3	12.4	1.0	1.9	8.1	0.9	0.9	7.4	0.8	0.6	5.3
20	$Se^{+}Le^{+}$	2.9	4.9	12.1	3.5	3.8	11.4	2.2	2.5	15.2	1.6	1.0	6.8	1.4	1.2	7.3	1.5	0.9	6.0
21	$Se^{+}Le^{+}$	2.2	2.4	20.1	6.1	6.0	5.5	1.4	1.9	8.1	0.8	0.8	9.1	0.5	0.4	5.3	0.5	0.3	4.9
22	$Se^{+}Le^{+}$	2.5	3.1	18.1	1.9	2.7	15.9	0.8	2.6	11.3	1.8	1.4	13.0	1.0	1.9	12.2	0.6	1.5	12.2
23	$Se^{+}Le^{+}$	4.4	2.3	16.9	3.1	1.5	14.5	1.0	2.8	10.0	0.3	3.6	12.5	1.0	0.9	6.4	0.3	0.9	5.7
24	$Se^{+}Le^{+}$	3.2	2.3	20.6	1.9	3.5	17.7	1.1	3.2	11.6	2.0	1.4	12.7	1.6	1.2	11.8	0.6	0.3	3.5
25	$Se^{-}Le^{+}$	7.0	12.3	6.5	6.8	9.2	9.0	1.4	3.8	4.7	1.4	1.5	2.3	1.7	1.4	2.2	1.2	0.9	1.6
26	$Se^{-}Le^{+}$	6.9	7.1	5.2	9.3	7.7	6.5	6.8	3.5	3.6	3.6	1.8	3.1	3.1	1.4	3.0	3.1	1.5	2.9
27	$Se^{+}Le^{-}$	9.7	1.1	12.6	1.8	0.8	19.4	2.5	2.0	14.2	2.1	0.6	17.3	1.2	0.9	13.5	2.0	0.7	10.4
28	$Se^{+}Le^{+}$	4.0	2.1	27.3	2.2	2.7	17.8	2.8	1.9	11.1	2.0	0.7	13.4	1.8	0.7	11.1	1.5	0.5	7.1
29	$Se^{+}Le^{+}$	4.0	1.0	16.9	2.3	2.7	17.4	1.6	4.8	14.6	2.2	2.5	12.0	1.4	1.7	9.6	1.4	1.2	8.1
30	$Se^{+}Le^{+}$	1.3	1.9	19.6	1.8	1.4	16.5	1.4	2.1	9.7	0.6	0.7	9.8	1.0	0.7	8.6	0.9	0.5	9.0

* Secretor, Lewis positive ($Se^{+}Le^{+}$) n = 22; non-secretor, Lewis positive ($Se^{-}Le^{+}$) n = 6; secretor, Lewis negative ($Se^{+}Le^{-}$) n = 2; non-secretor, Lewis negative ($Se^{-}Le^{-}$) n = 0. ** Symbols: A = acidic HMO concentrations in g/L; B = neutral nonfucosylated HMO concentrations in g/L; C = neutral fucosylated HMO concentrations in g/L.

Table S3.

Total HMO concentrations, total acidic HMO concentrations, and total neutral HMO concentrations (g/L) in human milk of 28 Dutch mothers collected at week 4 postpartum.

					Acidic and neutral HMO concentrations (g/L)**		
Mother No.	**SeLe status***	**Total HMO concentrations (g/L)**	**Body mass index kg/m²**	**Age**	**A**	**B**	**C**
1	Se^+Le^+	12.9	21.2	24	1.1	1.4	10.3
2	Se^+Le^-	15.2	21.8	33	0.5	2.1	12.6
3	Se^+Le^+	20.5	22.9	27	1.9	1.4	17.2
4	Se^+Le^+	4.2	27.6	28	0.5	0.6	3.2
5	Se^+Le^+	14.7	27.6	28	2.6	4.0	8.0
6	Se^+Le^+	14.5	20.7	28	2.3	1.9	10.3
7	Se^-Le^+	16.0	18.7	33	1.7	7.1	15.1
8	Se^+Le^+	26.9	39.0	28	8.4	7.7	10.8
9	Se^+Le^-	26.6	21.1	31	1.3	13.5	11.9
10	Se^-Le^+	10.1	31.0	31	0.8	4.9	5.0
11	Se^+Le^+	15.4	24.9	35	1.1	1.4	12.9
12	Se^-Le^+	5.7	21.4	34	0.5	2.5	2.9
13	Se^-Le^+	13.9	21.6	33	5.8	3.6	4.6
14	Se^+Le^+	13.5	19.5	42	1.7	2.0	9.8

Table S3 (continued)

					Acidic and neutral HMO concentrations (g/L)**		
Mother No.	**SeLe status***	**Total HMO concentrations (g/L)**	**Body mass index kg/m²**	**Age**	**A**	**B**	**C**
15	$Se^{+}Le^{-}$	9.3	22.1	27	0.8	1.2	7.3
16	$Se^{+}Le^{+}$	9.6	23.1	34	2.5	2.0	5.1
17	$Se^{+}Le^{+}$	15.4	23.3	35	7.0	2.4	10.9
18	$Se^{-}Le^{+}$	10.1	22.4	32	4.4	2.4	3.3
19	$Se^{+}Le^{+}$	21.3	23.9	26	3.2	2.6	15.5
20	$Se^{+}Le^{+}$	15.4	34.9	24	2.2	3.7	9.5
21	$Se^{-}Le^{+}$	20.3	37.2	27	6.6	6.4	7.3
22	$Se^{-}Le^{+}$	19.7	25.9	29	4.6	8.8	6.3
23	$Se^{+}Le^{+}$	17.7	25.9	31	2.2	4.3	11.2
24	$Se^{-}Le^{+}$	15.7	21.0	31	6.4	5.1	4.1
25	$Se^{-}Le^{+}$	10.3	39.0	28	3.9	3.5	2.9
26	$Se^{-}Le^{+}$	10.9	35.6	33	4.5	2.9	3.6
27	$Se^{-}Le^{+}$	13.9	20.7	32	5.3	5.1	3.4
28	$Se^{+}Le^{+}$	10.6	25.0	27	2.1	1.9	6.6

* Secretor, Lewis positive ($Se^{+}Le^{+}$) n = 14; non-secretor, Lewis positive ($Se^{-}Le^{+}$) n = 11; secretor, Lewis negative ($Se^{+}Le^{-}$) n = 3; non-secretor, Lewis negative ($Se^{-}Le^{-}$) n = 0.
** Symbols: A = acidic HMO concentrations in g/L; B = neutral nonfucosylated HMO concentrations in g/L; C = neutral fucosylated HMO concentrations in g/L.

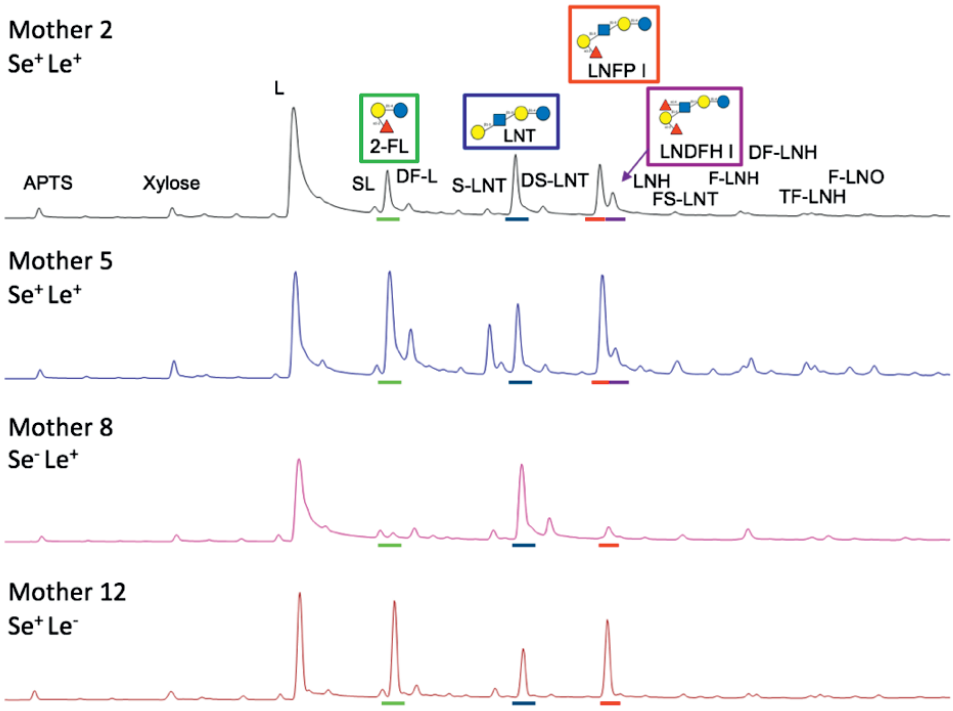


Figure S1. CE-LIF electropherogram of Chinese human milk of 4 mothers at week 1 postpartum. HMOs used to discriminate between the 3 SeLe major milk-types groups were 2'FL, LNT, LNFP I, and LNDFH I.

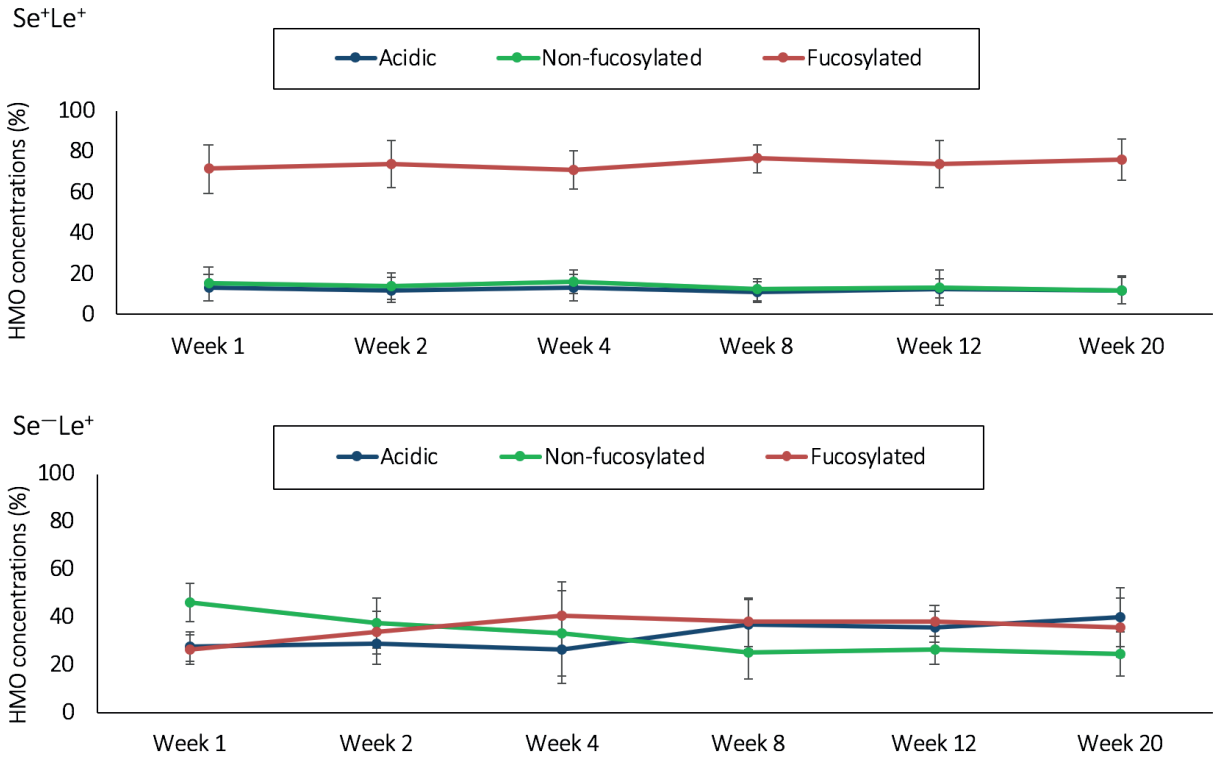


Figure S2. Concentrations of three groups of HMOs in Chinese human milk expressed in percentages over a 20-week lactation period for **(A)** $Se^{+}Le^{+}$ milk-type group n = 22 and **(B)** $Se^{-}Le^{+}$ milk-type group n = 6. The three groups of HMOs are acidic HMO concentrations, neutral nonfucosylated HMO concentrations, and neutral fucosylated HMO concentrations expressed in g/L.

Chapter 6

General discussion

Human milk as golden standard for infant nutrition

Infants who are exclusively breastfed for 6 months are among others less subjected to gastrointestinal infectious diseases[1]. Because of observed health benefits to newborns, the World Health Organization advices mothers to exclusively breastfeed babies for the first 6 months of life[2]. Although human milk is the gold standard for the infant's healthy development, not all mothers can breastfeed their infants for 6 months[3]. For women who are unable to breastfeed, infant formula is a helpful substitute to feed their child. However, infant formula does not offer the same benefits as breastfeeding, e.g. with regard to immune development[3]. Many improvements in infant formula composition have been implemented to make infant formula more similar to human milk in terms of composition and functionality. For example, the ratio between serum proteins and caseins was changed[4], but also bovine milk proteins (β-casein, α-lactalbumin, lactoferrin) and galacto- and fructooligosaccharides can be supplemented in infant formula[5]. Up to now, there is still a tendency to apply a one-size-fits-all approach when it comes to infant formula, but it may require a shift towards a more tailored infant approach in the future. Therefore, a better understanding of the human milk composition is needed, and more specifically on the immune-active human milk components.

The work described in this PhD thesis provides advanced knowledge on the variability in type and levels of human milk serum proteins in milk of Chinese and Dutch mothers over lactation (**Chapter 2**), and elaborates on the enzymatic digestion of proteins from colostrum and mature milk of Chinese mothers in an *in vitro* infant (0–3 months) digestion model (**Chapter 3**). In addition, it also provides insight in the variability in type and levels of serum protein *N*-glycans (**Chapter 4**) and human milk oligosaccharides (HMOs) in milk of Chinese mothers over time (**Chapter 5**). The HMO profiles in milk of Dutch mothers at week 4 was used for comparison (**Chapter 5**).

Proteins in Chinese and Dutch human milk over lactation

To better understand the variability in type and level of serum proteins in human milk, the milk serum proteome of 7 Chinese mothers over a 20-week lactation period was investigated using proteomic approaches (**Chapter 2**). A similar dataset earlier collected from milk of Dutch mothers was used for comparison.

The results showed that total milk serum protein concentrations in Chinese human milk decreased significantly over a 20-week lactation period, although with variation between mothers in rate of decrease. It was also observed that the starting milk serum protein concentrations differed among mothers. Variation was also found in the composition of serum proteins in both colostrum and mature milk, although the group of immune-active proteins, enzymes, and transport proteins were the most abundant for all mothers at all time points. These three protein groups encompass most of the 15 most abundant proteins in

Chinese human milk. A comparison between the milk of Chinese and Dutch mothers, resulted in 166 common serum proteins. The common 166 serum proteins in Chinese and Dutch human milk, represented more than 95% of the total amount of proteins in milk serum. The milk serum proteome of Chinese and Dutch mothers was similar in relative abundance of different functional groups as well as the most abundant proteins per group.
The ratio between protease inhibitors and proteolytic enzymes in colostrum of Chinese and Dutch mothers was approximately 10:1, with levels of proteases remaining relatively consistent over lactation. Colostrum contained a higher level of protease inhibitors compared to mature milk. No pepsin inhibitors can be found in human milk. Furthermore, a correlation higher than 0.5 was found between protease inhibitors and immune-active proteins in both Chinese and Dutch human milk. The data showed an even stronger correlation (0.8) for serine protease inhibitors and immunoglobulins, specifically. For the immune-active serum proteins (e.g. immunoglobulins, lactoferrin) to have an immune-activating role in the infant, they might be protected against the serine protease trypsin in human milk. It was also suggested that protease inhibitors (e.g. α_1-antichymotrypsin and α_1-antitrypsin) may reduce overall proteolysis during digestion, which will be further addressed in the next section.

Infant *in vitro* digestion of proteins in colostrum and mature milk of Chinese mothers

The variation in levels of protease inhibitors and total protein concentrations were investigated in relation to the degree of protein hydrolysis after *in vitro* digestion of colostrum and mature milk of 7 Chinese mothers (**Chapter 3**). A new *in vitro* digestion model was developed representing 0 to 3-month-old infant's digestion. The level of undigested protein, as marker for completeness of digestion was analysed by a combination of sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) and liquid chromatography–tandem mass spectrometry (LC–MS/MS).
No infant (0–3 months) *in vitro* digestion model was available before the start of this research. The *in vitro* digestion model was specifically developed to better comprehend the digestion of human milk proteins from colostrum and mature milk in this early life stage. SDS-PAGE is only able to detect some of the most abundant human milk proteins, therefore, LC-MS-MS was chosen to give a complete picture of the protein composition after digestion.
During data analysis, a static *in vitro* model was independenly developed and published by INFOGEST for infants born term and aged 28 days[6]. This model was based on literature[7] in combination with *in vivo* data of infants[8] and piglets[9]. The incubation time for each digestion stage was 60 min, and a pH of 5.3 and 6.6 were used during gastric and intestinal digestion, respectively[6]. The enzyme activities selected were 268 U/mL and 19 U/mL in the gastric phase for pepsin and lipase, respectively[6]. A comparison was made between the INFOGEST infant and adult model with regards to the digestive kinetics of the same food, i.e. a

commercial infant formula, composed of 1.18 g of proteins (ratio whey proteins: caseins w/w: 30:70) and 3.45 g of lipids for 100 mL. This INFOGEST infant model showed that α-lactalbumin, β-lactoglobulin, and caseins were not easily digested during gastric digestion[6]. After 60 min of infant gastric digestion, circa 62% of intact proteins were found[6]. No intact proteins could be found after intestinal digestion[6]. Although the parameters in this published model are not identical to ours, the conditions of the INFOGEST infant model did not deviate much from our model. Despite the differences in pepsin activity (200 U/mL), slightly different pH in the gastric (5.0) and intestinal phase (7.0) in our model, it was found that the total intact protein content, as measured by the bicinchoninic acid protein assay, was 52% in the gastric phase for all mothers averaged. It could also be observed that α-lactalbumin and caseins were completely digested after intestinal digestion. β-lactoglobulin is the most abundant protein in bovine milk serum[10-13], but cannot be found in human milk. Based on our experiments with bovine milk, β-lactoglobulin was also resistant against digestion in the gastric phase, and was completely digested after intestinal digestion (data not shown). Overall, the characteristics of protein breakdown in this INFOGEST study were similar to our results mentioned in **Chapter 3**.

The results showed that serum proteins and caseins were not completely digested by pepsin during digestion in the gastric phase. During the gastric phase, caseins clot or precipitate under the gastric pH[10]. At the end of intestinal digestion, caseins were almost completely digested *in vitro* by trypsin and chymotrypsin. Most milk serum proteins were more resistant against intestinal proteases than caseins. It also became clear that more than 40 serum proteins, including immune-active proteins (e.g. lactoferrin, immunoglobulins) and protease inhibitors (e.g. α_1-antichymotrypsin and α_1-antitrypsin), were still partially present intact after *in vitro* intestinal digestion. Although only present in small quantities, these proteins might be involved in supporting the infant's digestive tract against colonization of pathogens. Besides caseins, the human milk serum proteins polymeric immunoglobulin receptor (PIGR), osteopontin and clusterin also do not have an extensive tertiary structure[14]. These proteins might be more susceptible to proteases in both human milk and during digestion. Their structural features might explain why these proteins were almost completely digested at the end of *in vitro* intestinal digestion, in contrast to more tightly folded serum proteins like lactoferrin and immunoglobulins that showed less degradation (**Chapter 3**). In addition, a variety of studies have shown that specific serum proteins like lactoferrin, immunoglobulins, and α_1-antitrypsin in human milk can be found intact in the stool of breastfed infants, showing that those specific proteins were able to partially survive the digestion in the infant's digestive tract[10,12,13,].

For the first time, information was collected on the role of protease inhibitors and total protein content during *in vitro* infant digestion on level of protein degradation. It was hypothesized in **Chapter 2** that protease inhibitors might protect proteins against

degradation by proteases in the infant's gastrointestinal tract, because protease inhibitors might block trypsin and other serine proteases during intestinal digestion[14,15]. Reported concentrations of α_1-antichymotrypsin and α_1-antitrypsin are ranging from 0.4 to 0.7 g/L and 0.1 to 0.4 g/L, respectively, with a subsequent decline as lactation progresses[15]. Although it is unlikely that these protease inhibitors completely inhibit pancreatic proteolytic activity, they may allow enough of these milk serum proteins to survive and influence infant development[15]. It was also hypothesized that digestion of proteins could be influenced by the total protein content in colostrum, due to its effect on the protein-to-enzyme ratio (**Chapter 2**). It was found that colostrum contains a higher total protein content compared to mature milk (**Chapter 3**). Therefore, both the variation in levels of protease inhibitors, as well as total protein concentrations, were investigated in relation to the level of undigested protein. This investigation showed that both these factors were not correlated to the level of protein degradation. In addition, it was shown that the total undigested milk protein content decreased from the start to the end of digestion with large variation between mothers, especially in the gastric phase. Colostrum and mature milk of 7 Chinese mothers were digested after intestinal digestion to a similar extent. Overall, the higher levels of protease inhibitors and total protein in colostrum did not reduce the overall proteolysis during digestion.

Enzymatic hydrolysis is protein specific in human milk and in the infant's digestive tract

To support the work described in **Chapter 3**, which showed that most milk serum proteins are more resistant against proteolytic degradation than caseins, additional results were obtained on the presence of peptides in human milk after the *in vitro* digestion. In addition, information was gathered on the functionality of peptides formed during digestion from colostrum and mature milk.

In total, 1765 unique peptides were found for milk of 7 Chinese mothers from 2 different lactation periods after digestion (data not shown). This number includes a whole series of peptides that differed in length by one amino acid, which explains the high number of peptides. Of these peptides, 745 peptides were found in the milk prior to digestion of 7 Chinese mothers, and 716 and 180 peptides were released upon digestion in the gastric and intestinal phase, respectively (data not shown). The length of all the identified peptides were ranging between 8 and 25 amino acids. It was not possible to detect peptides shorther than 8 amino acids using MaxQuant, which explains absence of short peptides after intestinal digestion. Peptides found in colostrum and mature milk before digestion were mainly originating from the dominant proteins β-casein, PIGR, osteopontin, clusterin, κ-casein, and α_{S1}-casein (data not shown). The formation of peptides from other highly abundant human milk proteins (e.g. lactoferrin, serum albumin, immunoglobulins, and α-lactalbumin),

occurred mainly during digestion (data not shown). The relative levels of the total peptide content in both human milk and digesta can be found in Figure 1.

Material and Methods

Human milk of 7 Chinese mothers from 2 different lactation periods (week 1, colostrum; week 4, mature milk) was used. Samples before, during and after *in vitro* infant (0–3 months) digestion, were obtained and prepared as described in **Chapter 3**, and used for peptide isolation after trichloroacetic acid precipitation of the proteins[16]. The peptides were purified by reversed phase solid phase extraction, prior to analysis by LC-MS/MS[16]. MS/MS spectra from each run were obtained and analysed using MaxQuant with the Andromeda search engine[17]. Peptides and their native proteins were quantified based on label free quantification with a minimum ratio count of two peptides per protein[16]. Filtering and statistics on the MaxQuant output was done with Perseus[18], based on peak intensities of each single peptide as main input. The peak intensities of the peptides were $\log_2$ transformed, grouped per lactation period (colostrum and mature milk), and digestion stage (present in milk before digestion, gastric and intestinal digestion). Peptides in human milk and in digesta were grouped to their parent protein and their cleavage sites in the protein sequence were determined, and illustrated using the online software package Peptigram[19]. The function of the peptides were predicted with the milk bioactive peptide database[20].

It has been reported that plasminogen in human milk can bind to casein micelles, where it is activated to plasmin by urokinase-type and tissue-type plasminogen activators[21]. Due to the presence of activated plasmin, caseins are easier digested in milk itself than most serum proteins in human milk[21]. Most serum proteins resist digestion in human milk, because of a more tightly packed structure. The structural properties explain the low number and total intensity of serum protein derived peptides in human milk. The lower number and intensity of α_{S1}-casein peptides as compared to β-casein peptides might be explained by plasmin activity, which is known to be lower for α_{S1}-casein[21]. For κ-casein the least peptides and lowest total intensity in human milk was found, which might be explained by resistance of the N-terminus segment against native milk proteases[21].

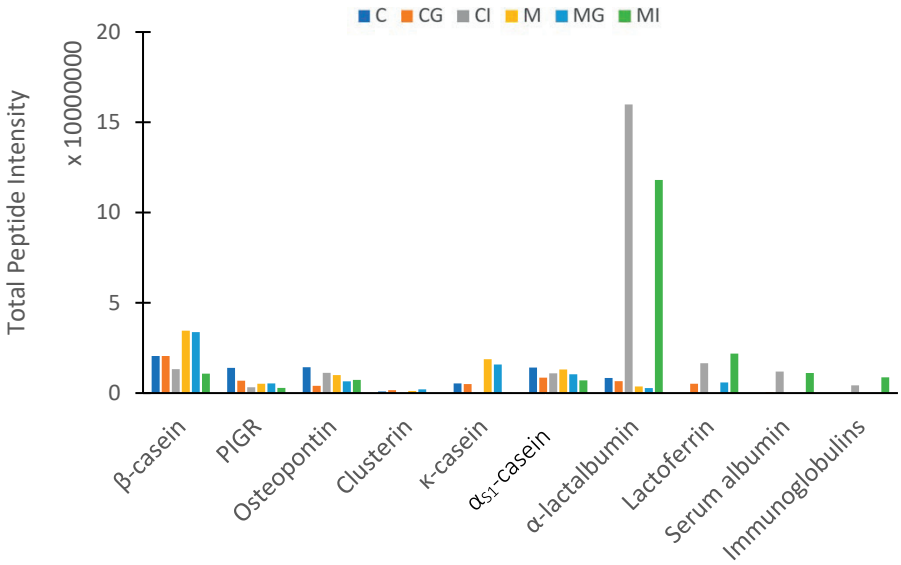


Figure 1. Total peptide intensities in milk of 7 Chinese mothers and in digesta (gastric and intestinal phase). The abbreviations indicate colostrum (C), mature milk (M), before digestion (B), gastric digestion (G), intestinal (I) digestion.

In total, 662 unique functional peptides were identified (37.5% of the 1765 peptides) in human milk before and after digestion using the milk bioactive peptide database[20]. β-casein as parent protein was the major supplier of functional peptides (630 of the 662 functional peptides, 95%) in human milk and during infant digestion (data not shown), with the majority released after gastric digestion. A schematic representation of the main functional peptides of β-casein can be found in Figure 2. The functional peptides of β-casein (Figure 2) in all samples mainly originated from the amino acid sequence 112–135 (LKSPTIPFFDPQIPKLTD) and amino acids 197–221 (LLNQELLLNPTHQIYPVTQPLAPVH) of the protein sequence. The peptide sequence LLNQELLLNPTHQIYPVTQPLAPVH was previously identified for having antimicrobial activity, including shorter parts of this peptide[23-26]. The similarity in patterns of the main bioactive peptides in Figure 2 also indicates that proteins from colostrum and mature milk are being digested similarly, as mentioned in **Chapter 3**. In addition, a comparitive analysis with the intensities of the common (>90%) peptides after intestinal digestion from colostrum and mature milk of 7 Chinese mothers showed no significant difference (data not shown).

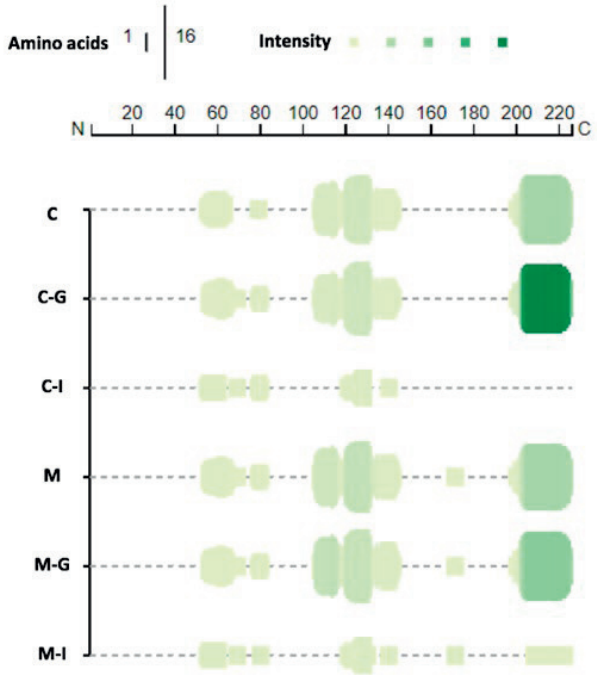


Figure 2. Peptigram comparing the main functional peptides originated from β-casein in colostrum (C) and mature milk (M) of 7 Chinese mothers, and across gastric (G) and intestinal (I) digestion. The position of the green color indicates origin of the peptide from the intact protein sequence, with the colors ranging from dark green to light green indicating high to low peptide intensity. Above the peptigram the range of amino acids in the human β-casein sequence (from the N-terminus to the C-terminus) are depicted.

In summary, the results indicate that enzymatic protein hydrolysis in human milk and after infant digestion differed between proteins, and that β-casein is the major supplier of functional peptides, independent of lactation stage. It also showed that caseins in milk are predigested more by native milk proteases than most serum proteins, although variation was found in type and levels of peptides in milk and digesta (gastric and intestinal phase) among mothers. However, it could be clearly observed that proteins from colostrum and mature milk of 7 Chinese mothers were digested to a similar extent after intestinal digestion. Although the presence of peptides in human milk and in digesta has been extensively studied[16,23-26], here, for the first time peptidomics data was collected on the digestion of proteins from colostrum and mature milk of mothers individually. Although, the composition of peptides found after intestinal digestion for colostrum and mature milk was quite similar

in numbers and levels, variation between mothers could be observed, especially in the gastric phase. This large variation was also found for the levels of undigested proteins after gastric digestion (**Chapter 3**).

Focus for future research on human milk proteins and peptides

Human milk contains a wide range of bioactive proteins and peptides[27]. There are several immune-active proteins present in human milk, such as lactoferrin, osteopontin, clusterin, immunoglobulins, and β_2-microglobulin, which are all present in high abundance (**Chapter 2**). These serum proteins may protect neonates from infections, when the infant's digestive tract and immune system is not yet fully developed[27]. For example, lactoferrin and immunoglobulins have direct bactericidal properties[27], and were still partially present after *in vitro* intestinal digestion (**Chapter 3**). Caseins do not have antimicrobial properties in their intact state, however, from caseins, many peptides as formed upon digestion may exert immunomodulatory and antimicrobial activity[23]. In addition, lactoferricin and lactoferrampin are two potent antimicrobial peptides that are generated from lactoferrin upon digestion, as well as the neurotensin-related peptide (kinetensin) from serum albumin[25]. These peptides were also found in our experiments in the digesta samples, for all mothers. A few studies have investigated the physical and structural properties of the different human milk proteins, peptides, and peptides formed upon digestion, as well as their biological activity[28-30]. In future research, the peptides derived from β-casein, κ-casein, lactoferrin, serum albumin could be synthesized and then further tested for their bacteriostatic properties. It would be helpful for the development of new ingredients for functional dairy products, to estimate the bacteriostatic properties of the main human milk proteins, both intact and upon enzymatic hydrolysis using a bacteriostatic assay. With such follow-up studies, the effect of high and low infant protein digestion levels can be monitored, to establish the biological importance of the difference between intact proteins in relation to the digesta. It would also be interesting to further explore the protein and peptide human milk composition before and after infant digestion for any correlations in regard to the diet and other characteristics of the mothers.

Serum protein *N*-glycans in Chinese human milk during lactation

Glycans can be found in human milk as free oligosaccharides or as glycoconjugates[31]. Two types of protein glycosylation exists in human milk; *N*-glycosylation to the amide nitrogen of asparagine side chains, *O*-glycosylation to the hydroxyl groups of serine and threonine side chains[31]. Different methods were tested for the release and profiling of serum protein *N*-glycans in human milk in **Chapter 4**. After method development and validation of the final method, *N*-glycans were successfully measured in human milk of 7 Chinese mothers over lactation using matrix assisted laser desorption ionization time of flight mass spectrometry

(MALDI-TOF-MS). The milk of the 7 mothers was assigned to the mother's Secretor (Se) Lewis (Le) histo-blood group. The SeLe histo-blood group system is based upon the fucosyltransferase (FUT) 2 secretor gene and the FUT3 Lewis gene.

In total, 66 different *N*-glycans were found in colostrum (week 1) and mature milk (week 4) of 7 Chinese mothers, with 15 structures covering >65% of the total *N*-glycan content. Based on the levels of the individual *N*-glycans in combination with principal component analysis, for the first time, a clear difference was observed with respect to milk of 5 Se^+Le^+ and 2 Se^-Le^+ mothers. In addition, it was also observed that milk of the Se^+Le^+ mothers were grouped based on lactation stage. This latter might also be the case for Se^-Le^+ mothers, however, a larger sample size is needed for confirmation. The milk of the Se^+Le^+ mothers and Se^-Le^+ mothers mainly differed based on the type and levels of neutral fucosylated *N*-glycans. The relative levels of the total neutral fucosylated *N*-glycans in milk from the 5 Se^+Le^+ mothers decreased over lactation, while the relative levels of the total acidic *N*-glycans remained constant, and the relative levels of total neutral nonfucosylated *N*-glycans increased. This pattern could not be observed for the milk from the 2 Se^-Le^+ mothers. The profiles of these three *N*-glycan groups stayed stable in milk over lactation for the 2 Se^-Le^+ mothers. It was found that the relative amounts of neutral *N*-glycans covered >90% of the total *N*-glycan content, for all 7 mothers.

The top 15 serum proteins cover 95% of the total serum protein content (**Chapter 2**), which was much lower for the top 15 *N*-glycans, where it was ranging between 65-72%. *N*-glycans and HMOs partly consist of similar monosaccharides, and the distribution of acidic and neutral *N*-glycans in **Chapter 4** matched with HMOs in human milk (**Chapter 5**). HMOs can be classified as neutral or acidic HMOs, with neutral oligosaccharides generally being present at a 10-fold higher concentration than acidic oligosaccharides (**Chapter 5**). Similar to HMOs, *N*-glycans containing sialic acid residues might be a good source of nutrients for the infant's brain development[32]. The fucosylation of HMOs depend on the mother's SeLe status, which was previously determined by others using different methods and analytical techniques[33]. It has been reported that >75% of the HMOs are fucosylated. The fucosylated *N*-glycans covered in levels >55% of the total *N*-glycan content in **Chapter 4**. The FUT2 and FUT3 enzymes are responsible for the presence of α1,2- and α1,4-fucosylated HMOs, respectively. Based on the mother's SeLe status, human milk can be classified in four groups for the HMOs. The milk of Se^+Le^+ contains α1,2-, α1,3-, α1,4-fucosylated HMOs, whereas the milk of Se^-Le^+ contain only α1,3-, α1,4-fucosylated HMOs[33].

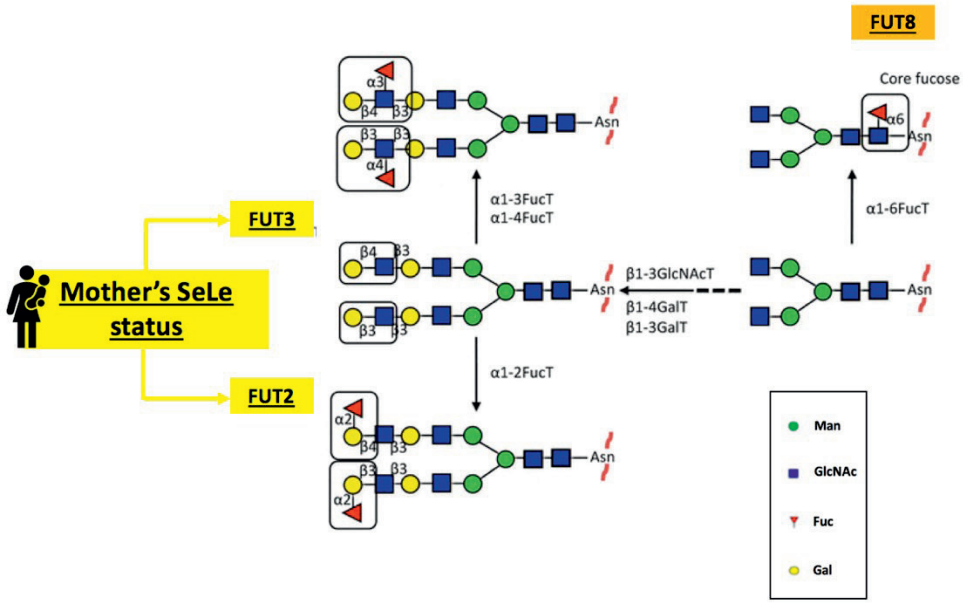


Figure 3. Hypothetical route of fucosylation of *N*-glycans in human milk. The core fucosylation level is regulated by FUT8, which is responsible for α1,6-linked fucosylated *N*-glycans[32], whereas FUT3 may be involved in the biosynthesis of fucosylated *N*-glycans.

The milk of the $Se^{+}Le^{-}$ mothers do contain α1,2-fucosylated HMOs[33]. The milk of the $Se^{-}Le^{-}$ mothers do not contain α1,2- nor α1,4-fucosylated HMOs[33]. The milk-types, $Se^{+}Le^{-}$ and $Se^{-}Le^{-}$, may also contain α1,3-fucosylated HMOs. A similar pathway may play a role of *N*-glycans (Figure 3). However, no *N*-glycans containing α1,4-linked fucosylated residues have been identified yet in human milk. Overall, this was the first study that provided detailed information on serum protein *N*-glycans in milk of individual mothers and over time, and provided evidence that *N*-fucosylation of serum proteins in human milk is associated with the mother's secretor status.

Focus for future research on *N*-glycans in human milk during lactation

N-glycans might not only affect the folding and stability of serum proteins in human milk and during infant digestion[34], but also influence the infant's microbiota[35-38]. These *N*-glycans might arrive in the colon as glycoprotein or as glycopeptides from serum proteins after digestion in the infant's small intestine. These *N*-glycans might have a similar function in the colon as HMOs. It has been reported that HMO profiles in the infant's faeces can be similar to that of ingested milk due to a specific fermentation of all HMOs, but HMOs can also

selectively be metabolized and with a different rate of fermentation[33]. Such information has not been shown for the *N*-glycans.

The relative amounts of the *N*-glycans were provided in **Chapter 4**, as measured by MALDI-TOF-MS. The different fucosyl-linkages for the *N*-glycans could not be determined in **Chapter 4**, which requires other analytical techniques. Further research should be undertaken to investigate the α1,2-, α1,3-, and α1,6-linked fucosylated *N*-glycans. A pilot study was conducted in the lab, trying to absolutely quantify *N*-glycans using ultra-high performance liquid chromatography mass spectrometry (UPLC-MS/MS) after labeling. Although it was confirmed that the abundant *N*-glycans as found by MALDI-TOF-MS were also the main *N*-glycans detected by UPLC-MS/MS (data not shown), the labeling method and the efficiency were not yet sufficient to accurately identify and quantify *N*-glycans in human milk. Overall, more knowledge about the structure of *N*-glycans (isomers), abundance, and their functionality is required. This should be further explored in a follow-up study. This could be complemented by an *in vitro* fermentation study to study their fate in the digestive tract. The method described in **Chapter 4** can be used to further study the type and levels of individual *N*-glycans in milk of mothers from different geographical and ethnic origin by MALDI-TOF-MS. In terms of directions for further research, it would also be interesting to investigate whether the α1,4-fucosylated *N*-glycans can be found in human milk, using milk from Se^+Le^- mothers in combination with milk of Se^+Le^+ or Se^-Le^+ mothers.

HMOs in Chinese and Dutch human milk over lactation

The variability of the lactose and HMO concentrations in human milk of Chinese and Dutch mothers was investigated using capillary electrophoresis laser-induced fluorescence (CE-LIF) (**Chapter 5**). This research investigated the level and type of HMOs in Chinese human milk of 30 mothers over a 20-week lactation period. For comparison, HMO profiles of 28 Dutch mother milk were determined 4 weeks after delivery and related to the HMO composition of milk from Chinese mothers.

Lactose and HMOs are the main components of the carbohydrate fraction in human milk (**Chapter 5**). Colostrum and mature milk generally contains 22 g/L and 13 g/L HMOs, respectively. It was observed that large variation exists between total HMO concentrations, which were ranging from 5 to 25 g/L. Total HMO concentrations in milk varied among Chinese mothers in colostrum, independent of the mother's SeLe status. The total HMO concentrations decreased with increasing time postpartum, although with individual variation in rate of decrease. Total acidic and neutral HMO concentrations in Chinese human milk varied over lactation, and neutral fucosylated HMO concentrations were driven by the mother's SeLe status. The obtained HMO data also showed a large variation between Chinese and Dutch mothers in total content of acidic and neutral HMOs at week 4 (data not

shown), as well on the levels of individual HMOs (data not shown). There was no significant difference observed on total HMO content (Figure 4).
Concentrations differ for the individual HMOs in literature, as well as the total HMO concentrations[39]. Different methods have been used to isolate and analyse HMOs. In a few cases, lower total HMO concentrations in milk have been reported, as compared to **Chapter 5**, however, with large variation in total HMO concent in milk between mothers. The reported values reported in **Chapter 5** can be found in the most HMO articles, making HMOs the third largest part in human milk, exceeding the protein content in human milk.

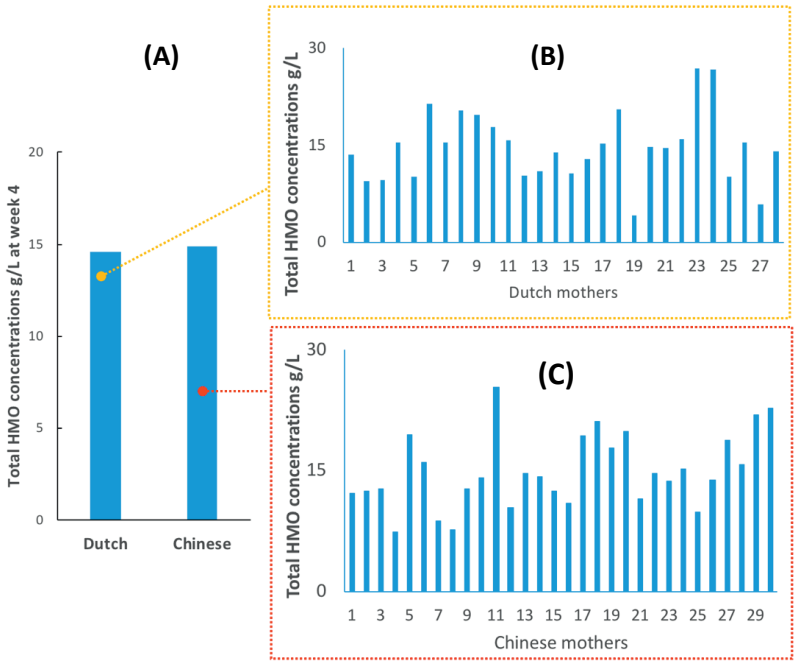


Figure 4. (A) Average total HMO concentrations (g/L) in Dutch and Chinese human milk at week 4 postpartum and total HMO concentrations (g/L) per individual mother; **(B)** Dutch human milk, 28 mothers; **(C)** Chinese human milk, 30 mothers).

For the first time, $Se^{+}Le^{+}$ subgroups were recognized in both the Chinese and Dutch population, based on the total concentrations of the neutral fucosylated HMO concentrations in human milk. No subgroups could be detected for the $Se^{-}Le^{+}$ groups in the milk of Chinese and Dutch mothers. The HMO 3-fucosyllactose (FL) was removed in the pretreatment step and not analysed (**Chapter 5**), and generally covers circa 5% of the total

HMO content. The HMO 3-FL is frequently not measured[39]. In addition, structural isomers co-elute using CE-LIF, like 2'FL and 3-FL, but also lacto-*N*-tetraose (LNT) and lacto-*N*-neotetraose, will co-elute[39]. In some rare cases, the structural isomers are separated on the column, for example lacto-*N*-fucopentaose (LNFP) II and LNFP III from LNFP I[39]. The HMO 3-FL might be one of the indicators to investigate the possibilities of subgroups for the $Se^{-}Le^{+}$ group, which contains α1,4-linked fucose residue. The distribution of the 3 other SeLe groups worldwide are 20% ($Se^{-}Le^{+}$ group), 10% ($Se^{+}Le^{-}$ group), <1% ($Se^{-}Le^{-}$ group). Based on the FUT2 and FUT3 enzymes, it would be expected that subgroups exist in the milk-types of $Se^{-}Le^{+}$ and $Se^{+}Le^{-}$ mothers. For example, milk of the $Se^{+}Le^{-}$ mothers contain HMOs with α1,2-linked fucose residues, and the enzyme activity may be reduced for the FUT2 enzyme due to polymorphism.

The 14 structures covered 90% of the total HMO content. Several studies have shown that only a few HMOs dominate the entire total HMO content[40-43]. Another recent study investigated the presence of the different binding-types (α1,2-, α1,3-, and α1,4-linked fucose residues) for the HMOs[43]. They observed that within each of the four milk groups, the relative levels of the different fucosylation epitopes greatly differed, and also showed the presence of two $Se^{+}Le^{+}$ group subgroups[43]. This grouping was based on the activity of a third unidentified Fuc enzyme, which is responsible for the synthesis of α1,3-linked fucose containing structures in FUT3 deficient individuals[43]. The SeLe status of an individual is the result of nonsense or missense single-nucleotide polymorphism in the FUT2 and FUT3 gene, and the observed Fuc X deficiency might lead to a non-functional allele in a similar manner[44], which may explain the variability in HMOs among mothers with the same SeLe status.

Focus for future research on *in vitro* fermentation of *N*-glycans and HMOs by infant microbiota

HMOs reach the colon intact and are fermented[33]. The manner and timing by which the gut is colonized in early life has a lasting impact on the microbiome and contributes largely to the variation in microbiota observed between individuals[31]. In the first year, the infant's microbiota evolves from relatively simple but rapidly increasing in diversity to an adult state that is more complex and stable[31].

Some HMOs can be found intact in the faeces of breastfed infants during the first 8-10 weeks after birth, whereas other HMOs are partly or completely utilized in the colon by microorganisms[44]. Such a monitoring of the utilization of single HMOs by the immature infant's microbiota is enhanced by recent techniques making it possible to separate and identify individual HMO structures after *in vitro* and *in vivo* fermentation by infant inoculum[45]. The preference of microorganisms and the fermentation rate of individual HMOs might be linked to the size, type of decorations, and type of linkages present in the oligosaccharide. However, the microbiome is found to be not directly correlated with the

HMO composition[44]. It is not clear whether health beneficial effects originate from the HMO structures directly, or from degradation and metabolisation products produced upon fermentation[33]. To get a better understanding of the functionality of HMOs, detailed information on the fate of these oligosaccharides during *in vitro* fermentation could be obtained.

Besides HMOs, also some serum proteins can be found intact in the faeces of breastfed infants during the first weeks after birth[22]. It can also be observed in **Chapter 3** that more than 40 serum proteins were found intact after *in vitro* intestinal digestion, altough some in small quantities, and these serum proteins may thus still have antibacterial activity in the gut. Most (>75%) of the serum proteins are *N*-glycosylated[46]. *N*-glycans of the glycosylated serum proteins in human milk are not yet extensively studied, but are considered to make a significant contribution to the development of the infant's immune system. The *N*-glycans in human milk consists of similar buildings blocks as the *N*-glycans in bovine milk. *N*-glycans released from bovine serum proteins, for example, were already tested *in vitro*[38]. In this latter study, the researchers used a cell wall-associated endo-ß-*N*-acetylglucosaminidase found in the infant's stool to release a range of intact bovine serum protein *N*-linked glycans. The *N*-glycans from concentrated bovine colostral milk did support the growth of *Bifidobacterium infantis*, which is a species known to grow well on HMOs. However, these *N*-glycans did not support the growth of *Bifidobacterium lactis*[38]. HMOs also did not support the growth of *Bifidobacterium lactis*[38]. Interestingly, *Bifidobacterium infantis* did not grow on intact glycoproteins. Mass spectrometry-based profiling of the *N*-glycans revealed that *Bifidobacterium infantis* consumed 73% of neutral and 92% of sialylated *N*-glycans, while *B. lactis* degraded only 11% of neutral and virtually no sialylated *N*-glycans[38]. This study supports the idea that *N*-glycans serve as selective substrates, and that released *N*-glycans were better fermentation substrates than intact glycoproteins[38]. However, it should be kept in mind, that the total content for *N*-glycans is typically 200-fold lower than for HMOs in human milk[46], and that only a low amount of *N*-glycans thus reach the infant's colon. Some *N*-glycans might be traced back intact in the infant's stool in early infancy, while others might serve as nutrients for colonic bacteria and preferentially support the growth of *bifidobacteria*. So far, it is not known whether intact *N*-glycans or glycosylated peptides can be traced back in the infant's stool, although glycosylated proteins are being digested in the infant's gastrointestinal tract, which leads to peptides, and the *N*-glycans attached to these peptides might prevent these peptides from being absorped in the small intestine. It is also not yet known whether the *N*-glycans from bovine and human milk are similar in their functions and structures[47]. In addition, the *N*-acetylneuraminic acid (NeuAc) residue was observed in bovine and human milk, whereas the *N*-glycolylneuraminic acid residue was only observed in bovine milk and marks a major difference between human and bovine milk[47]. Human milk analysis revealed that 57% and 43% of the structures were acidic and neutral *N*-

glycans, respectively, and 75% of the *N*-glycans were fucosylated. Bovine milk analysis showed that 68% and 32% of the structures were acidic and neutral *N*-glycans, respectively, and 31% of the *N*-glycans were fucosylated[47].

Several *in vitro* fermentation studies have been done for HMOs[48-52]. It would be interesting to continue the work presented in **Chapter 4** for *N*-glycans and **Chapter 5** for HMOs with *in vitro* fermentation of *N*-glycans and HMOs by infant microbiota, with the main objective to characterize *N*-glycans and HMOs before and after *in vitro* fermentation with pooled infant faecal inoculum using UPLC-MS/MS, and then to determine whether smaller *N*-glycans and HMOs are easier fermented compared to larger *N*-glycans and HMOs. It would also be interesting to include whether fucosylation of these structures affects the fermentation speed. Isomeric structure information of *N*-glycans and HMOs relating to fermentation behaviour could give more insight in the preferences of infant microbiota towards specific structures.

Selection of key components from human milk for follow-up studies, and possible future addition to infant formula

As mentioned earlier, a better understanding of variability in human milk composition may be used to improve infant formulas. Different approaches can be followed for the different immune-active components described throughout this thesis. Specifically for HMOs, the next step would be to obtain and supplement larger HMOs into infant formula. For HMOs, it is still a bottleneck to produce large quantities of these HMOs by organic synthesis[53]. Different approaches should be used to obtain HMOs in large quantities[53]. The shift from chemical synthesis to biotechnological manufacturing might be the preferred way to get HMOs in accessible in large quantities and at prices within reach for commercial applications, including infant formula[53]. The effect of heat processing on serum proteins in general, and on immune-active proteins in particular would be of interest, ensuring that these proteins stay intact, before adding serum proteins to infant formula, usually in the form of a whey protein ingredient. Unfolding of proteins by heating might effect their functional properties. The most abundant serum proteins (**Chapter 2**), *N*-glycans (**Chapter 4**), and HMOs (**Chapter 5**) in human milk were investigated, and variation in type and levels of these three components was established. The HMOs and *N*-glycans displayed in Figure 5, might be interesting for all infant's in infant formula. HMOs like lacto-*N*-hexaose (LNH) and LNT are not fucosylated, and highly abundant in human milk. The *N*-glycans 25 and 33 consists of an α1,6-linked fucose residue, which are structures that are always present in human milk, independent of the SeLe status (**Chapter 4**). Recently, 2'FL is supplemented to infant infant[54], consisting of α1,2-fucosylated linkage to the core HMO structure, whereas 3-FL is being subjected for approval, which contains a α1,3-linked fucose residue[54,55]. However, it is still not yet known, for example, what would happen if you give a specific HMO to an infant non-

matching with their mother's SeLe blood group, which might be an interesting follow-up study. For 2'FL and 3-FL these findings might be unpublished, but these results certainly would be of interest.

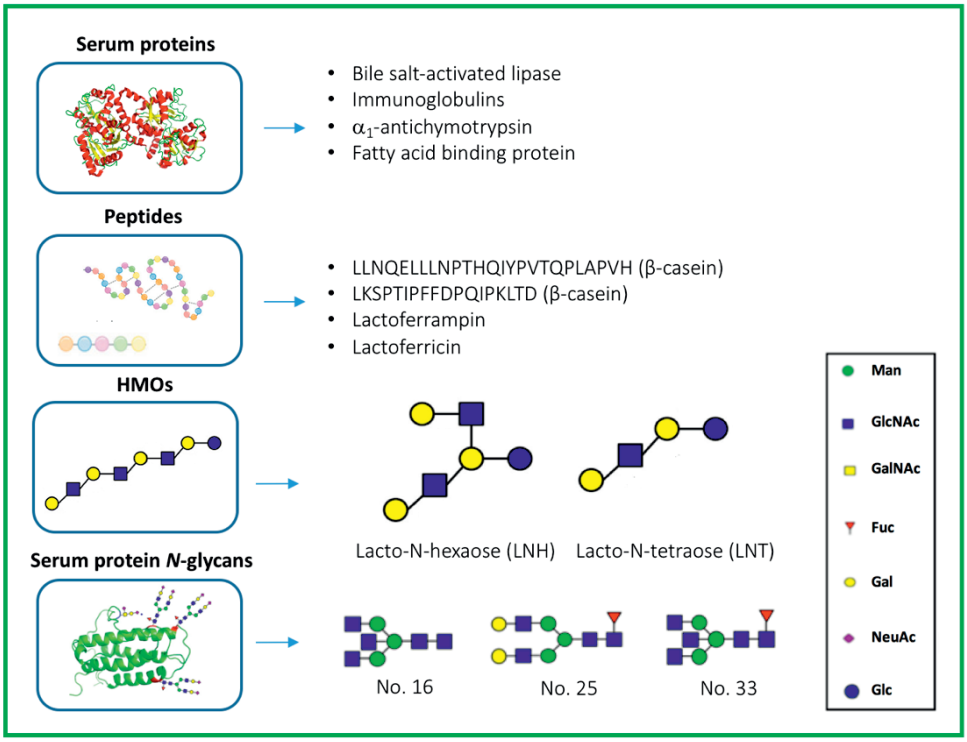


Figure 5. Proposed key components in human milk for follow-up studies, and possible future addition to infant formula. The *N*-glycans 25 and 33 (numbers as reported in **Chapter 4**, Table 1) consists of an α1,6-linked fucose residue. LNH, and LNT indicates lacto-*N*-hexaose and lacto-*N*-tetraose, respectively. Symbols indicate mannose (Man), *N*-acetylglucosamine (GlcNAc), *N*-acetylgalactosamine (GalNAc), fucose (Fuc), galactose (Gal), NeuAc and glucose (Glc).

Lactoferrin, β-casein, and α-lactalbumin have previously been added to infant formula[5], and the serum proteins mentioned in Figure 5 might be important as well, as these proteins partly survive digestion (**Chapter 3**) and can be found in the infant's stool up to six months[22], and have known physiological benefits in infants. Intact immune-active proteins, like lactoferrin and immunoglobulins, might support the infant's intestine against pathogens. Caseins are overall more easily degraded than serum proteins by proteases during infant

digestion, and might be a valuable nutritional source providing essential amino acids, peptides, and minerals for the infant's growth and development. Although β-casein is already added to infant formula, it should be monitored whether functional peptides are being formed. A comparison study between human milk and infant formula showed that β-casein is the largest supplier of peptides in human milk, whereas in infant formula the release of peptides was derived from a more broaden range of proteins[26]. It should be mentiond that bovine and human milk proteins differ not only in amino acids, but also in their digestibility[56,57]. The addition of peptides to infant formula might be benefical for formula-fed infants, especially the peptides from human β-casein and lactoferrin, as they may have antibacterial capacity and may thereby reduce gastrointestal infections.

Concluding remarks

Human milk composition is not constant throughout lactation and changes based on the infant's nutritional and developmental needs. Human milk consists of many high and low abundant immune-related and nutritional components, making the composition unique. The top 15 serum proteins in human milk generally cover >95% of the protein content, whereas for HMOs the 14 main structures cover >90% of the HMO content. The most abundant HMOs and serum proteins in human milk might be important for the healthy development of newborns. In early life, infants have an immature intestinal immune system, making them more vulnerable to infection by opportunistic pathogens. This does not mean that low abundant structures not play a key role in the infant's healthy development. The low abundant serum proteins and HMOs may stay relatively more intact during digestion and fermentation, respectively. The large variability in structures and levels of both the most abundant serum proteins and HMOs in human milk makes it complicated to design a single infant formula enriched with these components for all babies. Nevertheless, the 15 most abundant HMOs and serum proteins in human milk might be a good starting point to make infant formula more similar to human milk in terms of composition and functionality.

References

1. Newburg, D. Bioactive components of human milk: Evolution, efficiency and protection. Adv. Exp. Med. Bio. **2001**, 501, 3–10.
2. Kramer, M.; Kakuma, R. The optimal duration of exclusive breastfeeding: A systematic review. Adv. Exp. Med. Bio. **2004**, 554, 63–77.
3. Clark, D.; Goldman, A. Breastfeeding: Maintaining an irreplaceable immunological resource. Nat. Rev. Immunol. **2004**, 4, 565–72.
4. Ballard, O.; Morrow, A. Human milk composition: Nutrients and bioactive factors. Pediatr. Clin. North Am. **2013**, 60, 49–74.

5. Moro, G.; Minoli, I.; Mosca, M.; Fanaro, S.; Jelinek, J.; Stahl, B.; Boehm, G. Dosage-related bifidogenic effects of galacto- and fructooligosaccharides in formula-fed term infants. J. Pediatr. Gastrointestinal Nutr. **2002**, 34, 291–5.
6. Ménard, O.; Bourlieu, C.; De Oliveira, S.; Dellarosa, N.; Laghi, L.; Carriere, F.; Capozzi, F.; Dupont, D.; Deglaire, A. A first step towards a consensus static *in vitro* model for stimulating full-term infant digestion. Food Chem. **2018**, 240, 338–45.
7. Bourlieu, C.; Menard, O.; Bouzerzour, K.; Mandalari, G.; Macierzanka, A.; Mackie, A.; Dupont, D. Specificity of infant digestive conditions: Some clues for developing relevant *in vitro* models. Crit. Rev. Food Sci. Nutr. **2014**, 54, 1427–57.
8. De Oliveira, S.; Bellanger, A.; Ménard, O.; Pladys, P.; Le Gouar, Y.; Dirson, E.; Kroell, F.; Dupont, D.; Deglaire, A.; Bourlieu, C. Impact of human milk pasteurization on gastric digestion in preterm infants: A randomized controlled trial. Am. J. Clin. Nutr. **2017**, 105, 379–90.
9. Bouzerzour, K.; Morgan, F.; Cuinet, I.; Bonhomme, C.; Jardin, J.; Le Huerou-Luron, I.; Dupont, D. *In vivo* digestion of infant formula in piglets: Protein digestion kinetics and release of bioactive peptides. Br. J. Nutr. **2012**, 108, 2105–14.
10. Dupont, D.; Tome, D. Milk proteins: Digestion and absorption in the gastrointestinal tract. In milk proteins; Boland, M., Singh, H., Thompson, A., Eds.; Academic Press: London, **2014**, pp. 557–69.
11. Zhang, L.; de Waard, M.; Verheijen, H.; Boeren, S.; Hageman, J.A.; van Hooijdonk, T.; Vervoort, J.; van Goudoever, J.B.; Hettinga, K.A. Changes over lactation in breast milk serum proteins involved in the maturation of immune and digestive system of the infant. J. Proteomics, **2016**, 147, 40–7.
12. Lindberg, T.; Engberg, S.; Sjoberg, L.; Lönnerdal, B. *In vitro* digestion of proteins in human milk fortifiers and in preterm formula. J. Pediatr. Gastroenterol. Nutr. **1998**, 27, 30–6.
13. Chatterton, D.; Rasmussen, J.; Heegaard, C.; Sørensen, E.; Petersen, T. *In vitro* digestion of novel milk protein ingredients for use in infant formulas: Research on biological functions. Trends Food Sci. Technol. **2004**, 15, 373–83.
14. Dallas, D.; Murray, N.; Gan, J. Proteolytic systems in milk: Perspectives on the evolutionary function within the mammary gland and the infant. J. Mammary Gland Biol. Neoplasia. **2015**, 20, 133–47.
15. Lindberg, T. Protease inhibitors in human milk. Pediatr. Res. **1982**, 16, 479–83.
16. Dingess, K.; de Waard, M.; Boeren, S.; Vervoort, J.; Lambers, T.; van Goudoever, J.B.; Hettinga, K.A. Human milk peptides differentiate between the preterm and term infant and across varying lactational stages. Food Funct. **2017**, 8, 3769–82.
17. Cox, J.; Mann, M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. Nature, **2008**, 26, 1367–72.
18. Tyanova, S.; Temu, T.; Sinitcyn, P.; Carlson, A.; Hein, M.; Geiger, T.; Mann, M; Cox, J. The Perseus computational platform for comprehensive analysis of proteomics data. Nat. Methods, **2016**, 13, 731–40.
19. Manguy, J.; Jehl, P.; Dillon, E.; Davey, N.; Shields, D.; Holton, T. Peptigram: A web-based application for peptidomics data visualization. J. Proteome Res. **2016**, 16, 712–9.
20. Nielsen, D.; Beverly, R.; Qu, Y.; Dallas, D. Milk Bioactive Peptide Database: A comprehensive database of milk protein-derived bioactive peptides and novel visualization. Food Chem. **2017**, 232, 673–82.

21. Dallas, D.; Smink, C.; Robinson, R.; Tian, T.; Guerro, A.; Parker, E.; Smilowitz, J.; Hettinga, K.A.; Underwood, M.; Lebrilla, C.; German, B.; Barile, D. Endogenous human milk peptide release is greater after preterm birth than term birth. J. Nutr. **2015**, 145, 425–33.
22. Davidson, L.; Lönnerdal B. Persistence of human milk proteins in the breastfed infant. Acta Pædiatr. **1987**, 76, 733–40.
23. Gan, J.; Robinson, R.; Wang, J.; Krishnakumar, N.; Manning, C.; Lor, Y.; Breck, M.; Barile, D.; German, B. Peptidomic profiling of human milk with LC-MS/MS reveals pH-specific proteolysis of milk proteins. Food Chem. **2019**, 274, 766–74.
24. Su, M.; Broadhurst, M.; Liu, C.; Gathercole, J.; Cheng, W.; Qi, X. Comparative analysis of human milk and infant formula derived peptides following *in vitro* digestion. Food Chem. **2017**, 221, 1895–903.
25. Dall'Asta, C.; Florio, P.; Lammardo, A.; Prandi, B.; Mazzeo, T.; Budelli, A.; Pellegrini, N. Development of an *in vitro* digestive model for studying the peptide profile of breast milk. Int. J. Food Sci. Nutr. **2015**, 66, 409–15.
26. Hernandez-Ledesma, B.; Quiros, A.; Amigo, L.; Recio, I. Identification of bioactive peptides after digestion of human milk and infant formula with pepsin and pancreatin. Int. Dairy J. **2007**, 17, 42–9.
27. Lönnerdal, B. Bioactive proteins in human milk: Mechanisms of action. J. Pediatr. **2010**, 156, 26–30.
28. Korhonen, H.; Pihlanto-Leppäla, A.; Rantamäki, P.; Tupasela, T. Impact of processing on bioactive proteins and peptides. Trends Food Sci. Technol. **1998**, 9, 307–19.
29. Christen, L.; Lai, C.; Hartmann, B.; Hartmann, P.; Geddes, D. The effect of UV-C pasteurization on bacteriostatic properties and immunological proteins of donor human milk. PLoS One, **2013**, 8, 1–9.
30. Recio, I.; Lopez-exposito, I. Antibacterial activity of peptides and folding variants from milk proteins. Int. Diary J. **2006**, 16, 1294–305.
31. Kirmiz, N.; Robinson, R.; Shah, I.; Barile, D.; Mills, D. Milk glycans and their interaction with the infant-gut microbiota. Annu Rev. Food Sci. Technol. **2018**, 9, 429–50.
32. Wang, W.; Wang. L.; Du, Y.; Wu.; Yu, X.; Ye, K.; Li, C.; Jung, Y.; Qian, Y.; Voglmeir, J.; Liu, L. Comparison of antipathogenic activities of the human and bovine milk *N*-glycome: Fucosylation is a key factor. Food Chem. **2017**, 235, 167–74.
33. Albrecht, S.; Schols, H.A.; van den Heuvel, E.; Voragen, A.G.; Gruppen, H. Occurrence of oligosaccharides in faeces of breastfed babies in their first six months of life and the corresponding breast milk. Carbohydr. Res. **2011**, 346, 2540–50.
34. Vasconcelos-Dos-Santos, A.; Oliveira, I.; Lucena, M. Biosynthetic machinery involved in aberrant glycosylation: Promising targets for developing of drugs against cancer. Front Oncol. **2015**, 5, 1–23.
35. Yang, Q.; Wang, L. Mammalian a1,6-fucosyltransferase (FUT8) is the sole enzyme responsible for the *N*-acetylglucosaminyltransferase I-independent core fucosylation of high-mannose *N*-glycans. J. Biol. Chem. **2016**, 291, 11064–71.
36. Becker, D.; Lowe, J. Fucose: Biosynthesis and biological functions in mammals. Glycobiology, **2003**, 291, 41–53.

37. Dallas, D.; Martin, W.; Strum, J.; Zivkovic, A.; Smilowitz, J.; Underwood, M.; Affolter, M.; Lebrilla, C.; German, B. *N*-linked glycan profiling of mature human milk by high performance microfluidic chip liquid chromatography time of flight tandem mass spectrometry. J. Agric. Food Chem. **2011**, 59, 4255–63.
38. Karav, S.; Le Parc, A.; Bell, J.; Frese, S.; Kirmiz, N.; Block, B.; Barile, D.; Mills, D. Oligosaccharides released from milk glycoproteins are selective growth substrates for infant-associated bifidobacteria. Appl. Environ. Microbiol. **2016**, 82, 3622–30.
39. Van Leeuwen, S. Challenges and pitfalls in human milk oligosaccharide analysis. Nutrients, **2019**, 11, 1–21.
40. Tonin, K.; de Morais, M.; Abrao, A.; Miranda, A. Maternal and infant factors associated with human milk oligosaccharide concentrations according to secretor and Lewis phenotypes. Nutrients, **2019**, 11, 1–16.
41. Austin, S.; De Castro, C.; Sprenger, N.; Binia, A.; Affolter, M.; Garcia-Rodenas, C.; Beauport, L.; Tolsa, J.; Fumeaux, C. Human milk oligosaccharides in the milk of mothers delivering term versus preterm infants. Nutrients, **2019**, 11, 1–23.
42. Paganini, D.; Uyoga, M.; Kortman, G.; Boekhorst, J.; Schneeberger, S.; Karanja, S.; Hennet, T.; Zimmermann, M. Maternal human milk oligosaccharide profile modulates the impact of an intervention with iron and galacto-oligosaccharides in Kenyan infants. Nutrients, **2019**, 11, 1–17.
43. Van Leeuwen, S.; Stoutjesdijk, E.; ten Kate, G.; Schaafsma, A.; Dijck-Brouwer, J.; Muskiet, F.; Dijkhuizen, L. Regional variations in human milk oligosaccharides in Vietnam suggest FucTx activity besides FucT2 and FucT3. Sci. Rep. **2018**, 8, 16790–801.
44. Borewicz, K.; Gu, F.; Saccenti, E.; Hechler, C.; Beijers, R.; de Weerth, C.; van Leeuwen, S.; Schols, H.; Smidt. The association between breastmilk oligosaccharides and faecal microbiota in healthy breastfed infants at two, six, and twelve weeks of age. Sci. Rep. **2020**, Manuscript accepted for publication.
45. Mantovani, V.; Galeotti, F.; Maccari, F.; Volpi, N. Recent advances on separation and characterization of human milk oligosaccharides. Electrophoresis, **2016**, 37, 1514–24.
46. Froehlich, J.; Dodds, E.; Barboza, M.; McJimpsey, E.; Seipert, R.; Francis, J.; An, H.; Freeman, S.; German, J; Lebrilla, C. Glycoprotein expression in human milk during lactation. J. Agric. Food Chem. **2015**, 58, 6440–8.
47. Nwosu, C.; Aldredge, D.; Lerno, L.; Zivkovic, A.; German, B.; Lebrilla, C. Comparison of the human and bovine milk *N*-glycome via high-performance microfluidic chip liquid chromatography and tandem mass spectrometry. J. Proteome Res. **2012**, 11, 2912–24.
48. Gotoh, A.; Urashima, T.; Harata, G.; Okuda, S.; Katayama, T. Sharing of human milk oligosaccharides degradants within bifidobacterial communities in faecal cultures supplemented with *Bifidobacterium bifidum*. Sci. Rep. **2018**, 8, 1–14.
49. Marcobal, A.; Barboza, M.; Froehlich, J.; Block, D.; German, B.; Lebrilla, C.; Mills, D. Consumption of human milk oligosaccharides by gut-related microbes. J. Agric. Food Chem. **2010**, 58, 5334–40.
50. Boehm, G.; Moro, G. Structural and functional aspects of prebiotics used in infant nutrition. J. Nutr. **2008**, 138, 1818–28.

51. LoCascio, R.; Ninonuevo, M.; Freeman, S.; Sela, D.; Grimm, R.; Lebrilla, C.; Mills, D.; German, B. Glycoprofiling of Bifidobacterial consumption of human milk oligosaccharides demonstrates strain specific, preferential consumption of small chain glycans secreted in early human lactation. J. Agric. Food Chem. **2007**, 55, 8914–9.
52. Davis, J.; Totten, S.; Smilowitz, J.; German, B.; Mills, D.; Lebrilla, C. Identification of oligosaccharides in faeces of breastfed infants and their correlation with the gut microbial community. Mol. Cell Proteomics, **2016**, 15, 2987–3002.
53. Bych, B.; Miks, M.; Johanson, T.; Jondelius, M.; Vigsnaes, K.; Becker, P. Production of HMOs using microbial hosts – from cell engineering to large scale production. Curr. Opin. Biotechnol. **2019**, 56, 130–137.
54. Christensen, A.; Skov, S.; Lendal, S.; Hornshoj, B. Quantifying the human milk oligosaccharides 2'-fucosyllactose and 3-fucosyllactose in different food applications by high-performance liquid chromatography with refractive index detection. J. Food Sci. **2020**, 85, 332–9.
55. Reverri, E.; Devitt, A.; Kajzer, J.; Baggs, G.; Borschel, M. Review of the clinical experiences of feeding infants formula containing the human milk oligosaccharide 2'-fucosyllactose. Nutrients, **2018**, 10, 1161–73.
56. Inglingstad, A.; Devold, T.; Eriksen, E.; Holm, H.; Jacobsen, M.; Liland, K.; Rukke, E.; Vegarud, G. Comparison of the digestion of caseins and whey proteins in equine, bovine, caprine and human milks by human gastrointestinal enzymes. Dairy Sci. Technol. **2010**, 549–63.
57. Lönnerdal, B. Bioactive proteins in human milk: Health, nutrition, and implications for infant formulas. J. Pediatr. **2016**, 173, 4–9.

Summary

The work presented in this PhD thesis focused on the variability in type and levels of human milk oligosaccharides (HMOs) and serum proteins in milk of Chinese and Dutch mothers over lactation. In addition, it provides information on the serum protein *N*-glycan composition in milk of Chinese mothers over time, and elaborates on the enzymatic digestion of proteins from colostrum and mature milk of Chinese mothers in an *in vitro* infant (0–3 months) digestion model.

Chapter 1 described the current state of knowledge on human milk, including the benefits of human milk, the structures and characteristics of proteins and HMOs in human milk, and the fate of these constituents during digestion in the infant's gastrointestinal tract.

Chapter 2 provided information on the variability of serum proteins in human milk from Chinese mothers over a 20-week lactation period. Human milk of 7 mothers was collected at 6 time points over a 20-week lactation period, including colostrum (week 1 and 2) and mature milk (week 4, 8, 12, and 20). Human milk serum proteins were isolated, digested by trypsin, and labeled by stable isotopes. These samples were analysed with liquid chromatography–tandem mass spectrometry (LC-MS/MS), after which the proteins were identified and quantified. A similar dataset earlier collected from Dutch mothers was used as a comparison. The results showed that the total serum protein concentrations differed in colostrum among Chinese mothers. The protein concentrations in colostrum decreased significantly over a 20-week lactation period, although with variation between mothers in rate of decrease. The group of immune-active proteins, enzymes, and transport proteins were the most abundant for all mothers. Variation was also found in the composition of serum proteins in both colostrum and mature milk of the Chinese mothers. The group protease inhibitors, transport, enzyme, immune-active proteins encompass the 15 most abundant proteins, covering >95% of total protein present for both the Chinese and Dutch milk serum proteome. The levels of the protease inhibitors were twice as high in colostrum than in mature milk. In addition, the levels of 22 of the 166 common milk serum proteins varied in human milk from Chinese and Dutch mothers, and these differences found were more apparent in colostrum than in mature milk. The levels of protease inhibitors were 10-fold higher than proteolytic enzymes in colostrum of both Chinese and Dutch mothers. The levels of proteases in human milk remained relatively the same over lactation. Furthermore, a correlation was found between immune-active proteins and protease inhibitors in both Chinese and Dutch human milk during lactation. This might indicate that protease inhibitors protect immune-active proteins from degradation against native human milk proteases.

Chapter 3 provided insights on the degradation of human milk proteins in an *in vitro* infant digestion model, by comparing colostrum (week 1) and mature milk (week 4) of 7 Chinese mothers individually. An *in vitro* digestion model was successfully developed representing 0 to 3-month-old infant's digestion, which deviated from existing adult *in vitro* digestion

models by having a gastric pH of 5, lower pepsin and trypsin activities, and shorter incubations times, which mimic the infant's *in vivo* gastrointestinal tract better. Undigested proteins (> 10 kDa) were analysed after gel-electrophoretic separation and in-gel tryptic digestion by LC-MS/MS. LC-MS/MS provided a more detailed overview of the proteins present during *in vitro* infant digestion. Some of the highly abundant serum proteins (lactoferrin, bile salt-activated lipase, immunoglobulins, α_1-antichymotrypsin, serum albumin, fatty acid binding protein) were still partially intact after intestinal digestion of both colostrum and mature milk proteins. Other abundant serum proteins, α-lactalbumin, polymeric immunoglobulin receptor, clusterin, osteopontin, β_2-microglobulin, and the 3 caseins (β-, α_{S1}- and κ-casein) were almost completely digested after the *in vitro* digestion. This might be explained by the structural features of these proteins. The higher levels of both protease inhibitors and total protein in colostrum than in mature milk did not reduce overall proteolysis during digestion. Overall, small quantities of immune-active proteins were the most resistant against proteolytic enzymes during *in vitro* infant digestion.

Chapter 4 provided information on the Chinese human milk serum glycoproteome over lactation. Human milk of 7 mothers from 2 time points was studied, colostrum (week 1) and mature milk (week 4). The milk was assigned to the mother's Secretor (Se) Lewis (Le) histo-blood group. Milk serum was isolated by ultracentrifugation, after which the serum proteins were separated from the HMOs. Serum proteins were denatured, incubated with enzyme for 24h, and the released *N*-glycans after purification analysed by matrix assisted laser desorption ionization time of flight mass spectrometry (MALDI-TOF-MS). In total, 66 different serum protein *N*-glycans were found in colostrum and mature milk, with 15 structures covering >65% of the total *N*-glycan content. Variation was found in type and levels of individual *N*-glycans among mothers and over lactation. Based on the levels of the individual *N*-glycans in combination with principal component analysis, for the first time, a clear difference was observed between the milk of the 5 Se^+Le^+ and 2 Se^-Le^+ mothers. It was also observed that *N*-glycan profiles of the Se^+Le^+ mothers differed depending on lactation stage. A trend could also be observed for the milk of the Se^-Le^+ mothers. The top 15 serum protein serum protein *N*-glycans were identical in milk for both the 2 Se^-Le^+ mothers and the 5 Se^+Le^+ mothers. The type and level of the top 15 structures differed between these two genetic groups. The profiles of the total acidic and neutral (nonfucosylated and fucosylated) *N*-glycans over time varied between milk of Se^+Le^+ and Se^-Le^+ mothers. The relative amounts of neutral *N*-glycans covered >90% of the total *N*-glycan content, for all 7 mothers. Overall, this study investigated and compared for the first time single serum protein *N*-glycans in milk of individual mothers over time, and showed that fucosylation of *N*-glycans from serum proteins in human milk can be linked to the mother's secretor status.

Chapter 5 provided insights in the variability of the lactose and HMO concentrations in human milk of Chinese and Dutch mothers using capillary electrophoresis with laser-induced

fluorescence detection, by investigating the level and type of HMOs in Chinese human milk of 30 mothers over a 20-week lactation period. In order to study whether the observed clustering in HMO composition was typical for Chinese mothers only, HMO profiles of 28 Dutch mothers were determined 4 weeks after delivery, and used as comparison. The results showed that total HMO concentrations in Chinese human milk decreased significantly over a 20-week lactation period, although with variation between mothers in rate of decrease. The top 14 structures in human milk generally covered >90% of the HMO content. The total acidic and neutral HMO concentrations differed between Dutch and Chinese human milk. Based on the total neutral fucosylated HMO concentrations in both Chinese and Dutch human milk, Se^+Le^+ subgroups were identified for the first time. HMOs that differed in level between these Se^+Le^+ subgroups were the structures 2'fucosyllactose, difucosyllacto-*N*-hexaose, lacto-*N*-fucopentaose I, and fucosyllacto-*N*-octaose, having in common α1,2-fucosylated linkages to the core HMO structures. No subgroups could be found for the Se^-Le^+ mothers.

Chapter 6 summarized and discussed the results of the foregoing chapters, and provided supporting data on the digestion of individual proteins. The additional experiments were carried out to provide information on the peptide composition in milk and digesta using the same infant *in vitro* digestion model. β-casein was the largest supplier of potential bioactive peptides in colostrum, mature milk, and during *in vitro* infant digestion. Some of the major proteins (caseins, osteopontin, clusterin, PIGR, and β_2-microglobulin) were predigested by proteases in human milk, resulting in the presence of their peptides in human milk. Most serum proteins are less sensitive to digestion by these native milk proteases, which might be due to their compact folded globular structure. After summarizing and discussing the results of the foregoing chapters, several key components from human milk were highlighted. In addition, follow-up studies have been proposed, including among others an *in vitro* fermentation study with *N*-glycans and HMOs by infant microbiota. These latter milk constituents might both help in shaping a beneficial microbiota in the human infant, which is important for health maintenance. In conclusion, this thesis provided detailled information on the variability of serum proteins, *N*-glycans and HMOs in colostrum and mature milk of Chinese mothers, as well as on the digestion of human milk proteins from colostrum and mature milk using an newly developed infant *in vitro* model.

Acknowledgments

At this point, I would like to express my sincere gratitude to my highly learned promotors Henk Schols (Laboratory of Food Chemistry, FCH) and Kasper Hettinga (Food Quality and Design Group, FQD) for their guidance, support, and encouragement throughout these beautiful and lovely 4 years. It was a privilege and honor to work with you both and I really appreciated every advice. My thanks also goes out to my colleagues from Yili, especially Fang Fang, April Li, Dai Xiaoshuang, Wendan Wang, Carolien van Loo-Bouwman, Gerrit Smit, and Ignatius Szeto.

Jolanda van den Boomgaard you are truly the beating heart of FCH, and yes I enjoyed the candy! Edwin Bakx, René Kuipers, Margaret Bosveld, Wouter Lubbers, Peter de Gijsel, and Mark Sanders, thank you all for sharing your expertise on the lab equipment. Mirjam Kabel, Peter Wierenga, Willem van Berkel, Fangjie Gu, Jean-Paul Vincken, Melliana Jonathan, Matthias Frommhagen, Hugo Cardoso, Roelant Hilgers, Zhibin Liu, Wouter de Bruijn, Suzanne Broxterman, and many other FCH-family members, my thanks for the helpful discussions, comments, and profitable advices. Yuxi Deng, Bianca Martens, Claire Butre, Gijs Vreeke, Carla Araya Cloutier, Maud Meijers, and Dimitrios Kouzounis, it was a pleasure to have you all in the office over the years. Similar accounts for my lab roommates, Madelon Logtenberg, Eva Jermendi, Christiane Rosch, Gijs van Erven, and Thibaut Mouthier. Bake de Rink, I really liked our coffee breaks, talking about others things than work.

Lysanne Hoksbergen, Kimberley Boss and Corine van Huenen thanks for all the secretary support at FQD. Xandra Bakker-de Haan, thank you so much for ordering all the stuff I needed for my experiments. Charlotte van Twisk, Mike de Beijer, Erik Meulenbroeks, and Frans Lettink, thank you all for guiding my students around in the lab. Hein van Valenberg, Tiny van Boekel, Teresa Oliviero, Edoardo Capuano, Anita Linnemann, and many other FQD-family members, my gratitude for your help, advice, and support, when needed. Also thanks goes to the dairy group members, Thom Huppertz, Sine Yener, Etske Bijl, Pieter Dekker, Chunyue Zhang, Daylan Tzompa Sosa, Ling Xiong, Naomi Arita Merino, Huifang Cai, Elsa Antunes-Fernandes, Sara Pacheco, James Ledo, Eva Kontopodi, Hannah Zenker, and Ruben de Vries. I want also to thank all my students (Nynke, Kim, Joost, Yuhui, Yalei, Simone, Kunyuan, Thomas, and Dongyao). Thanks also goes to the people at Biochemistry, especially Sjef Boeren and Jacques Vervoort. Jos Hageman, my thanks for the fruitful discussions on statistics. Kelly Dingess, it was always nice to have lunch together and I will never forget that we missed our flight back from Oxford to Amsterdam. From the Dutch human milk bank, I want to thank Johannes van Goudoever and Marita de Waard.

If your name is not mentioned, but you worked at 1 of the 3 groups in Wageningen somewhere between 2015 to 2019, please forgive me, I give you all my thanks. I want to thank so many people, but due to time, I will stop here. Special thanks also goes to my wife Rewan, my parents, my brother, and all my other family members in the Netherlands and

Egypt for all the emotional support and encouragement throughout my research project. During my first year of my project, my beloved Dutch grandmother died, who was my role model in many aspects. I will never forget you. Hereby, I kept also my promise to my grandmother to never give up in pursue of my dreams. After 4 years of hard work, dedication, and determination, I completed my PhD thesis and I can now shout it from the rooftops. I want to dedicate my thesis to my parents. Without their support, I could not have managed. ''Als je een bloem vindt, dan pluk je haar. Maar als je van een bloem houdt, dan geef je haar elke dag water. Vrienden komen en gaan, maar ouders blijven onvoorwaardelijk achter je staan in goede en slechte tijden''. This is basically what they did for me every single day.

About the author

Mohèb Elwakiel (born 13 October 1985, Utrecht) studied at the Leiden University where he obtained a BSc degree in Bio-Pharmaceutical Sciences (2005–2010). During his BSc he also managed to receive a BSc and MSc degree in Islamic Jurisprudence (2006–2010). Furthermore, he was one of the student ambassadors of the Leiden University (2005–2006), practicum supervisor for several courses (2006–2007), co-founder and treasurer of a multicultural student association SABR in Leiden (2006–2008), and board member of the first Dutch mosque in the Netherlands located in Amsterdam (2008–2009). After his bachelor graduation in Bio-Pharmaceutical Sciences, he started to work as a junior researcher for two years, and then spend one year on teaching Chemistry. From 2010 until 2013, he was also chairman at an educational institute for children, and founded the organization. Mohèb Elwakiel obtained a MSc degree in Pharmaceutical Sciences (2013–2015) from Utrecht University. He finalized his MSc thesis at Medicinal Chemistry and Chemical Biology, and performed an internship at the Dutch Medicines Evaluation Board. After obtaining his MSc degree in Pharmaceutical Sciences, he started his PhD-project with Prof. dr. Henk Schols and dr. Kasper Hettinga at Wageningen University & Research. The results of his PhD research project (2015–2019) are presented in this thesis.

Email address: m.elwakiel@gmail.com

List of publications

Elwakiel, M.; Boeren, S.; Hageman, J.A.; Szeto, I.M.; Schols, H.A.; Hettinga, K.A. Variability of serum proteins in Chinese and Dutch human milk during lactation. Nutrients. **2019**, 11, 499–513.

Elwakiel, M.; Hageman, J.A.; Wang, W.; Szeto, I.M.; van Goudoever, J.B.; Hettinga, K.A.; Schols, H.A. Human milk oligosaccharides in colostrum and mature milk of Chinese mothers: Lewis positive secretor subgroups. J. Agric. Food Chem. **2018**, 66, 7036–43.

Werkhoven, P.; **Elwakiel, M**.; Meuleman, T.; Quarles van Ufford, H.; Kruijtzer, J.; Liskamp, R. Molecular construction of HIV gp120 discontinuous epitope mimics by assembly of cyclic peptides on an orthogonal alkyne functionalized TAC-scaffold. Org. Biomol. Chem. **2016**, 14, 701–10.

Overview of completed training activities

Discipline-related activities	Organization	Year
Courses		
Food and Bio-refinery Enzymology[a]	VLAG	2015
Reaction Kinetics in Food Sciences	VLAG	2016
Summer School Maxquant (Oxford, England)	MPS	2016
Glycosciences[a]	VLAG	2016
Advanced Food Analysis[a]	VLAG	2017
Healthy Food Design	VLAG	2018
Diary Protein Biochemistry	VLAG	2018
Intestinal Microbiome of Humans and Animals	VLAG	2019
Symposia & Conferences		
Immunity in Early Life	Friesland Campina	2015
Many Faces of Carbohydrates	CCC	2015
Immunomodulation	Friesland Campina	2016
Polyunsaturated Fatty Acids – Microbiota – Immune Health	WUR	2017
Carbohydrates for improving Health	CCC	2017
International Conference on Food Digestion (Rennes, France)	Infogest	2017
European Conference Personalised Nutrition and Health	TNO	2019
General Courses		
PhD Introduction Week	VLAG	2016
Essentials of Scientific Writing and Presenting	WGS	2016
Scientific Publishing	WGS	2016
Reviewing a Scientific Manuscript	WGS	2016
Mobilizing your Scientific Network	WGS	2016
Communication with Media and General Public	WGS	2016
Workshops PhD carousel, Publish for Impact, Teaching Tricks	WUR	2016
Lecturing[b]	WUR	2017
(Re)designing a Course[b]	WUR	2018
Brain Based Teaching	WUR	2018
Group Assignments: Design, Supervision, Feedback & Assessment[b]	WUR	2018

Optional		
PhD Research Proposal	VLAG	2015
FCH Protein Group Meetings[c]	WUR	2015–2017
FQD Dairy Science and Technology Group Presentations[c]	WUR	2015–2019
FCH Carbohydrate Group Meetings[c]	WUR	2015–2019
FCH PhD Lunch Presentations[c,d]	WUR	2017–2019
FCH PhD Trip Japan[a,c]	WUR	2016
FQD PhD Trip Australia[a,c,d]	WUR	2018
Teaching activities		
Supervision of BSc and MSc projects	WUR	2015–2019
BSc Course Nutritional Aspects of Food	WUR	2015–2019
BSc Course Quality Systems Operations	WUR	2015–2019
BSc Course Food Packaging and Design	WUR	2016
MSc Course Advanced Biochemical Analysis of Food	WUR	2018

Abbreviations: Dutch acronym VLAG (English: Advanced Studies in Food, Technology, Agrobiotechnology, Nutrition and Health Sciences); MPS, Max Planck Society; CCC, Carbohydrate Competence Center; WUR, Wageningen University & Research; WGS, Wageningen Graduate Schools; TNO, Netherlands Organization for Applied Scientific Research; FCH, Laboratory of Food Chemistry; FQD, Food Quality and Design. [a] Poster Presentations [b] Part of the University Teaching Qualification [c] Oral Presentations [d] Organization committee.

The research described in this thesis was financially supported by Inner Mongolia Yili Industrial Group Company. Financial support from Wageningen University for printing this thesis is gratefully acknowledged.

Cover Design: Fahui Lui
This PhD thesis was printed by Digiforce || ProefschriftMaken
Edition 250 copies