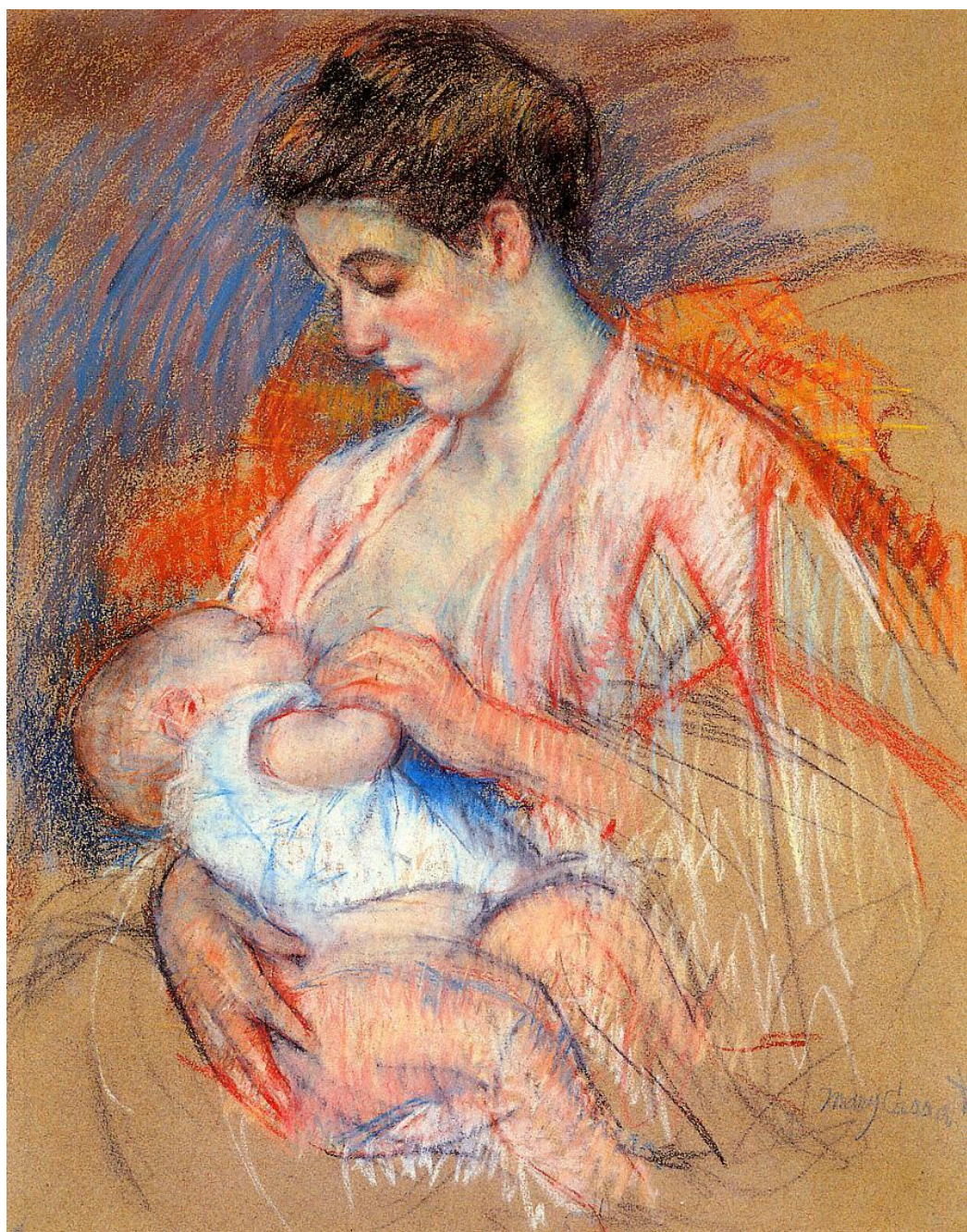




Mother Jeanne Nursing Her Baby
Mary Cassatt, 1908

**Characterizing immune-modulatory
components of human milk:
The fate and function of soluble CD14
and the human milk metagenome**

Tonya L. Ward

Thesis submitted to the
Faculty of Graduate and Postdoctoral Studies
in partial fulfillment of the requirements
for the degree Doctorate in Philosophy degree in Biochemistry

Department of Biochemistry, Microbiology and Immunology
Faculty of Medicine
University of Ottawa

© Tonya L. Ward, Ottawa, Canada, 2014

Abstract

Background

During the first stages of development human infants are either fed human milk or human milk substitutes (infant formulas). The composition of infant formulas and human milk differ drastically, including a difference in protein constituents and bacterial load. Due to the high global frequency of infant formula use, the humanization of infant formulas to better reflect the complex nature of human milk is warranted. To better understand the role of human milk components, the fate and function of a key bacterial sensor in human milk, soluble CD14, was determined. Additionally, the microbiome of human milk was analyzed from a metagenomic standpoint in an attempt to determine which types of bacteria are present in human milk and what their potential biological function might be.

Results

In rodent models, ingested sCD14 persisted in the gastrointestinal tract and was transferred intact into the blood stream. Once transferred to the blood, ingested sCD14 retained its ability to recognize lipopolysaccharide and initiate an immune response in pups. This transfer of sCD14 across the epithelial barrier was also observed in human cells in vitro, where it appears to be dependent on Toll-like receptor 4. Using Illumina sequencing and the MG-RAST pipeline, the human milk metagenome of ten mothers was sequenced. DNA from human milk aligned to over 360 prokaryotic genera, and contained 30,128 open reading frames assigned to various functional categories. The DNA from human milk was also found to harbor

immune-modulatory DNA motifs that may play a significant role in immune development of the infant.

Conclusions

Given the complex nature of human milk in comparison to its bovine or plant based substitutes, the results presented in this thesis warrant future modification of infant formulas to include non-nutritive bioactive components. Current human milk components not yet present in infant formulas include the diverse microbiome of human milk, the immune-modulatory DNAs which those microbes harbor, and bioactive human proteins such as sCD14.

Acknowledgements

I would like to express a special appreciation for my supervisor, Dr. Illimar Altosaar. Thank you for introducing me to the fascinating topic of human milk and allowing me the freedom to approach research questions using my own creativity. Your encouragement and mentoring has allowed me to grow as a scientist. I would also like to thank the members of my thesis advisory committee, Dr. Mack, Dr. Makrigiannis and Dr. Stintzi, for their advice and input on how to better my research.

Thank you, to my collaborators at The Ottawa Hospital and the Children's Hospital of Eastern Ontario, as my research would not be possible without your efforts on donor recruitment. Also, I would like to thank the staff of the Animal Care and Veterinary Services for their expertise and invaluable help, as well as the departmental staff for their technical assistance. A big thank you goes out to previous and current members of the Altosaar lab, your company made me feel at home in the lab and your assistance helped to shape my research to what it is today.

A special thank you goes to my family and friends. Words cannot express how thankful I am for your love, kindness, understanding and patience during the course of my studies. Despite the distance you were only ever a phone call away, providing a source of humour during times of frustration and always willing to provide a listening ear during experimental triumphs.

Table of Contents

Abstract	ii
Acknowledgements	iv
List of Abbreviations	vii
List of Figures	ix
List of Tables	xi
Chapter 1 – Introduction	1
1.1 Infant nutrition	2
1.2 Composition of milk	4
1.3 Bioactive protein in milk	8
1.4 CD14 in human milk	10
Hypotheses and objectives	14
1.5 Cells and DNA as milk bioactives	15
Hypotheses and objectives	19
Chapter 2 – Digestive fate of soluble CD14	20
2.1 Abstract	21
2.2 Introduction	22
2.3 Methods	25
Radiolabeling of proteins	25
Feeding studies and protein isolation	25
Quantifying intact and degraded proteins	26
Determining protein intactness	27
Detecting CD14 in rat milk	27
Data Analysis	28
2.4 Results	29
Ingested sCD14 persists in the GI tract	29
Ingested sCD14 is transferred to the blood	33
Rat milk contains sCD14	38
2.5 Discussion	43
Evasion of digestion by milk proteins	43
Activity of digested and intact sCD14	44
Potential mechanisms of CD14 uptake	45
Possible role of ingested CD14	46
Conclusions	48
Chapter 3 – Functionality of ingested soluble CD14	49
3.1 Abstract	50
3.2 Introduction	52
3.3 Methods	54
Animals, fostering and LPS injection	54
CD14 and cytokine concentrations	54
Caco-2 cells and transport assays	56
TLR4 siRNA knockdown	57
Immunocytochemistry and microscopy	58
Statistical analysis	58

3.4 Results	59
Milk sCD14 is transferred to the blood of pups	59
sCD14 transferred to the blood remains functional	60
Human cells transport sCD14 in vitro	60
sCD14 transfer is TLR4 dependent	67
3.5 Discussion	77
Fate of ingested sCD14	77
Functionality of ingested sCD14	79
sCD14 transfer by Caco-2 cells is TLR4 dependent	81
Implications of biologically active sCD14 post-ingestion	82
Conclusions	83
Chapter 4 – Human milk metagenome	84
4.1 Abstract	85
4.2 Introduction	87
4.3 Methods	90
Donors and sample collection	90
DNA isolation	90
DNA sequencing, filtering and contig assembly	91
Contigs, ORF prediction and characterization	91
Immune-modulatory motif identification	92
Availability of supporting data	93
4.4 Results	94
Phyla and genera within human milk	94
Open reading frames within human milk	100
Human milk compared to feces	101
Immune-modulatory DNA motifs	111
4.5 Discussion	115
Genera of bacteria within human milk	115
Phylogenetic differences between human milk and feces	117
Functionality of the human milk metagenome	118
Immune-modulatory landscape of human milk	120
Conclusions	121
Chapter 5 – General Discussion	123
5.1 sCD14 in human milk	124
5.2 Bacteria and DNA in human milk	126
5.3 Conclusions	128
References	131
Appendices	141
A. Supplemental figures	141
B. Supplemental tables	148
C. Curriculum vitae	161

List of Abbreviations

BF	Breast-fed
BSA	Bovine serum albumin
Caco-2	Human colonic intestinal epithelial cells (colorectal adenocarcinoma)
CD14	Cluster of differentiation 14
CI	Confidence Intervals
COG	Cluster of orthologous groups
CpG	Cytosine phosphate guanine
d	Day
DPM	Disintegrations per minute
DTT	Dithiothreitol
ECL	Enhanced chemiluminescent
EDTA	Ethylenediaminetetraacetate
ELISA	Enzyme linked immunosorbent assays
FF	Formula-fed
GAPDH	Glyceraldehyde 3-phosphate dehydrogenase
GI	Gastrointestinal
GPI	Glycosylphosphatidylinositol
HMO	Human milk oligosaccharide
hrCD14	Human recombinant CD14
HRP	Horseradish peroxidase
HT29	Human colonic intestinal epithelial cells (colorectal adenocarcinoma)
IL	Interleukin
LBP	Lipopolysaccharide binding protein
LP	Lipoproteins

LPS	Lipopolysaccharide
mCD14	Membrane bound CD14
mIC _{c12}	Murine intestinal crypt-derived cell-line
NEC	Necrotizing enterocolitis
ORF	Open reading frame
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PP	Post partum
rRNA	Ribosomal RNA
sCD14	Soluble CD14
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
sIgA	Secretory immunoglobulin A
siRNA	Small interfering RNA
TLR4	Toll-like receptor 4
TLR9	Toll-like receptor 9
TNF	Tumor necrosis factor
TIRAP	Toll-interleukin 1 receptor adapter protein
TBST	Tris buffered saline with Tween
WHO	World Health Organization
WT	Wild type

List of Figures

Figure 1.1	The gross composition of human milk.	Page 6
Figure 1.2	The involvement of CD14 in the innate immune response to lipopolysaccharide.	12
Figure 1.3	Recognition of DNA by Toll-like receptor 9.	18
Figure 2.1	Biodistribution of intact and degraded [^{14}C]-sCD14 and a digestible control protein, [^{14}C]-BSA, post-ingestion in 10-d-old rat pups.	32
Figure 2.2	Biodistribution of [^{125}I]-sCD14 and a digestible control protein, [^{125}I]-BSA, post-ingestion in 10-d-old rat pups.	35
Figure 2.3	Uptake of intact (>30 kDa) and degraded (<30 kDa) [^{125}I]-sCD14 into the blood of 10-d-old rat pups.	37
Figure 2.4	Intactness of ingested [^{125}I]-sCD14 in the blood of 10-d-old rat pups after size separation (< or >30 kDa), visualized by phosphor imaging and silver staining of SDS-PAGE gels.	40
Figure 2.5	The presence of rat-sCD14 in rat milk, 10 days PP, as detected by immunoblotting.	42
Figure 3.1	Quantity of sCD14 in the stomach contents and blood of mouse pups ingesting milk with or without sCD14.	62
Figure 3.2	Intactness of sCD14 in the stomach contents and blood of mouse pups ingesting milk with or without sCD14.	64
Figure 3.3	Immune response to LPS following sCD14 ingestion in mouse pups, assessed via sCD14, TNF-alpha, and IL-6 levels.	66
Figure 3.4	Quantity and intactness of sCD14 transferred across human intestinal monolayers in vitro.	70
Figure 3.5	Visualization of sCD14 within human intestinal cells treated with sCD14 in vitro.	72
Figure 3.6	TLR4 expression in Caco-2 and HT29 cells.	74

Figure 3.7	Effect of TLR4 knockdown on sCD14 transfer by human intestinal cells in vitro.	76
Figure 4.1	Best hit analysis of 51 bp DNA sequences from human milk.	96
Figure 4.2	Best hit analysis of open reading frames within human milk.	99
Figure 4.3	Functional categorization of open reading frames within human milk.	103
Figure 4.4	Best hit comparison of bacterial phyla in human milk, infants' feces and mothers' feces.	105
Figure 4.5	Functional category comparison of open reading frames within human milk versus infants' and mothers' feces.	108
Figure 4.6	Pair-wise comparison of categorized open reading frames from human milk versus infants' and mothers' feces.	110
Figure 5.1	The gross composition of human milk, modified to include cells of somatic and non-somatic origin as well as nucleic acids.	130
Supplemental		
Figure S2.1	Proteins isolated from the gastrointestinal tract and organs of rat pups fed [¹²⁵ I]CD14 or [¹²⁵ I]BSA.	142
Figure S2.2	Increased contrast of [¹²⁵ I]CD14 phosphor-images used in Figure 2.2A.	143
Figure S3.1	Timeline of pup fostering.	144
Figure S3.2	Passive transport of Lucifer Yellow across Caco-2 cell monolayers.	145
Figure S4.1	Pair-wise comparison of phyla abundance in human milk versus infants' and mothers' feces metagenomes.	146
Figure S4.2	Lowest common ancestor comparison of bacterial phyla in human milk, and in infants' and mothers' feces.	147

List of Tables

Table 3.1	Fostering scheme used to determine the quantity and functionality of ingested sCD14 transcytosed to the blood of mice.	Page 55
Table 4.1	Contig assembly and open reading frame (ORF) prediction of Illumina reads (51 bp) from human milk.	97
Table 4.2	Occurrence of immune suppressive motifs in various metagenomes.	113
Table 4.3	Occurrence of immune suppressive motifs TTAGGG and TCAAGCTTGA in contigs from human milk.	114
Supplemental		
Table S4.1	Abundance of DNA fragments in pooled human milk, sequenced seven times.	149
Table S4.2	Classification of 51 bp DNA sequences derived from human milk by best hit analysis.	150
Table S4.3	Predicted open reading frames from human milk DNA sequences aligning to rRNA genes of known organisms.	159
Table S4.4	Immune-modulatory DNA motifs sought in DNA sequences derived from human milk or feces.	160

Chapter 1

Introduction

1.1 Infant nutrition

One distinguishing feature that separates the organisms in the class Mammalia from other amniotes is the development of milk-secreting mammary glands. The production of milk post partum (PP) has evolved as a means to provide the first food needed by the offspring (1). Similar to other bodily fluids such as blood, milk is highly complex in composition, consisting of an emulsion of milk fat globules suspended in water containing carbohydrates, proteins, minerals, and whole cells. Ingestion of milk by the offspring is foremost for nutritive purposes and also provides immunological protection for the newborn through its bioactive components. In non-human mammals, milk is consumed as the first food, and lack of available milk to the newborn offspring can result in death without outside intervention.

Humans, however, do not always receive human milk as their first food, especially with the increasing availability of milk substitutes, which will be termed infant formulas herein. Due to the complex nature of human milk in comparison to infant formulas, the World Health Organization (WHO) recommends that exclusive human milk consumption should begin within one hour of birth and should continue for the first six months of life (2). Also recommended is continued breastfeeding complemented with supplementary foods for up to two years of age or beyond (2). Additionally, to promote breastfeeding and diminish outside influence to formula feed children, an international code was developed in 1981 to regulate the marketing of infant formulas. Specifically, this code calls for all formula labels to state the benefits of breastfeeding and risks of formula feeding, and also calls for a ban on the promotion of infant formulas through advertisements and free samples (2). Despite

agreement by health organizations on the benefits of human milk versus infant formulas, currently less than half of all infants exclusively receive human milk for the first six months of life (30% of infants, 95% CI 26-35), with the lowest rates of breastfeeding reported in Europe (14% of infants, 95% CI 7-22; 3).

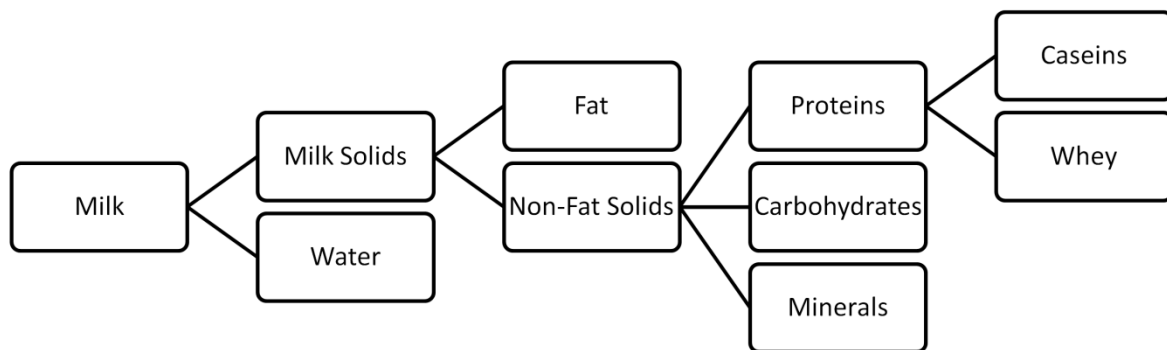
In developing countries, the benefits of human milk for the infant are even greater than in the more developed world, given the limited access of some communities to safe drinking water (4). In less developed countries one of the largest threats to infant health is the occurrence of diarrhea, which can lead to stunted growth and even death (3). By breastfeeding infants as opposed to formula feeding, the risk of diarrhea mortality is significantly diminished (5). Suboptimal breastfeeding has recently been reported as responsible for 804,000 deaths in 2011, which is 11.6% of total deaths of children under five years old (3). Despite efforts in recent years to improve breastfeeding practices through the WHO and UNICEF Baby Friendly Hospital Initiative (2), even countries with successful programs still have breastfeeding rates far below the levels recommended by the WHO (3). Therefore, while infant formulas continue to be widely used, the study of human milk components, their impact on infant health, and the potential addition of these components to infant formulas provide complementary avenues or strategies to improve the current status of infant nutrition.

1.2 Composition of milk

Similar to the milk of other mammals, human milk can be divided into major groups of components. The current concept of major milk components is shown in Figure 1.1, which includes fats, proteins, carbohydrates and minerals. The proportion of each milk component in Figure 1.1 varies from mother to mother and also changes over the duration of lactation. Throughout lactation there are three types of milk produced: colostrum (first 72 hours), transitional milk (72 hours – 2 weeks PP) and mature milk (2 weeks PP and onwards), each differing in composition (6). Colostrum, for example, contains significantly more protein than transitional milk (2.5 ± 0.2 g/100 mL versus 1.7 ± 0.1 g/100 mL), whereas the lipid content of human milk increases during lactation (2.2 ± 0.2 g/100 mL, 3.0 ± 0.1 g/100 mL and 3.8 ± 0.1 g/100mL for colostrum, transitional and mature milk, respectively (7). Also, when milk is monitored over longer periods of time (12 months), concentrations of other proteins are shown to fluctuate, some increasing in concentration while others decrease (8, 9). Therefore, unlike the stationary composition of infant formulas manufactured to a factory recipe, the composition of human milk is variable and fluctuates during lactation to ensure the full development of the growing infant.

The two major subdivisions of milk proteins are caseins and whey proteins, which differ significantly in their structure and solubility. Caseins lack disulfide bridges, contain highly hydrophobic regions and exist in milk as micelles, which contain calcium phosphate and casein complexes (10). Due to the lack of disulfide bridges, the micelles easily precipitate in acidic conditions (pH 4.6) causing a milk clot to form in the acidic stomach of the infant, aiding in more efficient digestion (10).

Figure 1.1. The gross composition of human milk. (Modified from 11, 12)



Conversely, whey proteins have a more ordered structure and are not easily precipitated by acidic conditions. Whey proteins include nutritive proteins, enzymes and bioactive proteins. One major difference between human and bovine milk, from which most infant formulas are derived, is the types of whey proteins present. For example, human milk does not contain the major bovine whey protein β -lactoglobulin (2-4 g/L in bovine milk), and human milk contains more total whey protein in comparison to bovine milk (whey to casein ratio is 60:40 in human milk, but 20:80 in bovine milk; 9, 13, 14). Also, bovine milk lacks human immunoglobulins and has lower levels of bioactive proteins, such as the antibacterial protein lactoferrin, than those in human milk (0.01-0.1 g/L versus 1.0-2.0 g/L; 15-17).

1.3 Bioactive proteins in milk

Although milk has classically been thought to be primarily nutritive in function, it also contains numerous bioactive components. The proteome of human milk consists of an array of immunologically relevant proteins including growth factors, cytokines, and antimicrobial peptides. For proteins to retain immunological function and provide benefit for the infant, bioactive milk proteins must remain intact following ingestion. Two examples of proteolytically resistant milk proteins include secretory immunoglobulin A (sIgA; 18, 19) and lactoferrin (20, 21), which have been shown to survive passage through the infant digestive system. Immunoglobulins and lactoferrin are also examples of proteins that are translocated intact across the mucosal barrier post-ingestion by the infant. Uptake of these two proteins along the gastrointestinal (GI) tract is receptor mediated by the neonatal Fc receptor and the lactoferrin receptor, respectively (22, 23). Once ingested, immunoglobulins from human milk provide passive humoral immunity donated by the mother to her infant (24). Similarly, lactoferrin has been shown to reduce the risk of sepsis and infection in preterm infants (25–27). Although the digestive fate and biological function of some milk proteins, such as those described above, have been characterized, the significance of the majority of milk proteins remains unknown. For example, although it is found in high concentrations in human milk, the fate and function of the pattern recognition receptor Cluster of Differentiation 14 (CD14) remains unknown. The lack of knowledge on the role of many human milk proteins causes difficulty in setting standards regarding which human proteins should be used to supplement infant

formulas. Therefore, further studies on the digestive fate and function of human milk proteins, such as CD14, are needed.

1.4 CD14 in human milk

CD14 is a pattern recognition receptor for several bacterial components including lipopolysaccharide (LPS) on the surface of Gram-negative bacteria. CD14 exists either as a glycosylphosphatidylinositol (GPI)-anchored membrane protein (mCD14) or as a soluble protein (sCD14). Soluble CD14 exists in human milk at $20.10 \pm 8.74 \mu\text{g/mL}$ (five days PP) and in serum at $3.71 \pm 0.59 \mu\text{g/mL}$, and is notably absent from commercially produced infant formulas (28, 29). Both forms of CD14 appear functionally interchangeable as they both enhance proinflammatory signalling in response to LPS through the Toll-like receptor 4 (TLR4)/MD-2 pathway (Figure 1.2A, B). In blood, sCD14 decreases LPS-related mortality and septic shock by sequestering LPS from mCD14/TLR4-expressing immune cells in coordination with accessory lipoproteins (Figure 1.2C; 30). Within the first year of life, the infant GI tract becomes colonized with bacteria, including those expressing LPS (31–34). The introduction of LPS, a molecule that in trace amounts ($15 \mu\text{g/kg}$ bodyweight) is able to induce a life threatening immune response (35), could be detrimental to the infant should the barrier function of its developing GI tract fail. Therefore, given the known role sCD14 plays in LPS recognition in conjunction with the considerable concentration of sCD14 in human milk warrants investigation into the role that ingested sCD14 may serve in LPS recognition by infants.

Previously, sCD14 was shown to partially survive a mock GI digestion, as sCD14 remained intact following pepsin digestion and was partially digested by pancreatin in vitro (36). Immunoprecipitation of sCD14 from milk also led

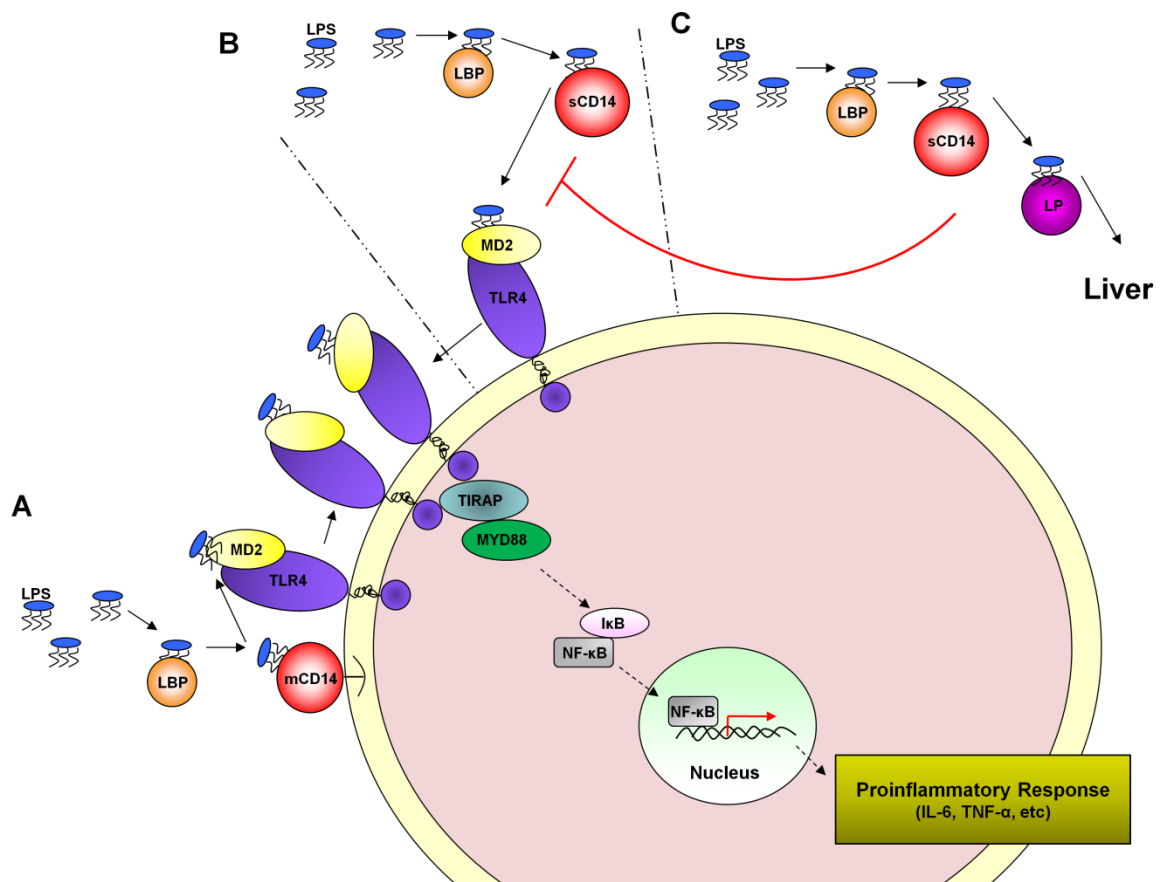
Figure 1.2. The involvement of CD14 in the innate immune response to lipopolysaccharide (LPS).

(A) LPS is delivered to mCD14 by LPS binding protein (LBP), which transfers LPS to MD2. The LPS:MD2 complex binds and activates TLR4. TLR4 signals internally through the Toll-interleukin 1 receptor adapter protein (TIRAP) and MYD88, leading to the migration of the NF- κ B transcription factor to the nucleus. NF- κ B then activates the transcription of proinflammatory genes such as tumor necrosis factor (TNF α) and interleukin 6 (IL-6).

(B) LPS is delivered to sCD14 by LBP. Soluble CD14 can then transfer LPS to MD2 and initiate the same TLR4 activation and proinflammatory response described in panel A.

(C) LPS is delivered to sCD14 by LBP, which can sequester LPS away from mCD14 and transfer LPS to serum lipoproteins (LP). Lipoproteins then carry LPS to the liver for clearance, thereby down regulating the immune response.

(Modified from 37–39).



to the discovery of sCD14-interacting proteins, such as α -lactalbumin that protect sCD14 from digestion (40). A lack of sCD14 in the feces of breast-fed infants (BF), in conjunction with the proteolytic protection of sCD14 in vitro suggests sCD14 may be taken up whole into the blood stream of BF infants. Additionally, the interaction between CD14 and TLR4 to signal an immune response to LPS has been shown to result in CD14 internalization in macrophages (41). Studies have shown TLR4 to be expressed by intestinal epithelial cells (42), which may represent a route of sCD14 transport from the GI lumen into the blood stream of BF infants.

This present thesis incorporates three major hypotheses derived from the above questions, and the research was conducted during the course of my training for the degree of Doctor of Philosophy. These three hypotheses have been submitted as two manuscripts (chapters 2 and 3) and are described below.

Hypotheses and Objectives

Manuscript 1 - Ingested soluble CD14 from milk is transferred intact into the blood of newborn rats

Hypothesis:

sCD14 remains intact along the digestive tract, and is absorbed intact into the circulatory system post-ingestion.

Objectives:

- Track the digestive fate of ingested sCD14 using ^{14}C - and ^{125}I -labeled sCD14 in a rat model of breastfeeding infants.
- Compare the digestive fate of ^{14}C - and ^{125}I -labelled sCD14 to a similarly labeled, fully digestible control protein, Bovine Serum Albumin (BSA).

Manuscript 2 - Ingested soluble CD14 contributes to the functional pool of circulating sCD14 in mice

Hypotheses:

Ingested sCD14 is a major contributor to circulating sCD14 in infants, and once transferred to the blood sCD14 remains functional in its ability to detect LPS.

Transfer of sCD14 across intestinal epithelial cells occurs in both rodent and human cells, and that sCD14 transfer across human intestinal cells is TLR4 dependent.

Objectives:

- Use CD14^{-/-} mouse pups fostered to wild type (WT) mothers to quantify the amount of sCD14 transferred from ingested milk to the circulatory system of infants.
- Compare the amount of sCD14 transferred to the circulatory system of CD14^{-/-} mice to that of WT mice in similar foster scenarios.
- Determine the functionality of ingested sCD14 in CD14^{-/-} pups following LPS injection using the immune response as a measurement of sCD14 functionality.
- Use human intestinal epithelial cells grown in Transwell assays to determine if sCD14 transfer is conserved across species.
- Use human intestinal epithelial cells in conjunction with small interfering RNA (siRNA) to determine if TLR4 is responsible for transfer of sCD14 across the epithelium.

1.5 Cells and DNA as milk bioactives

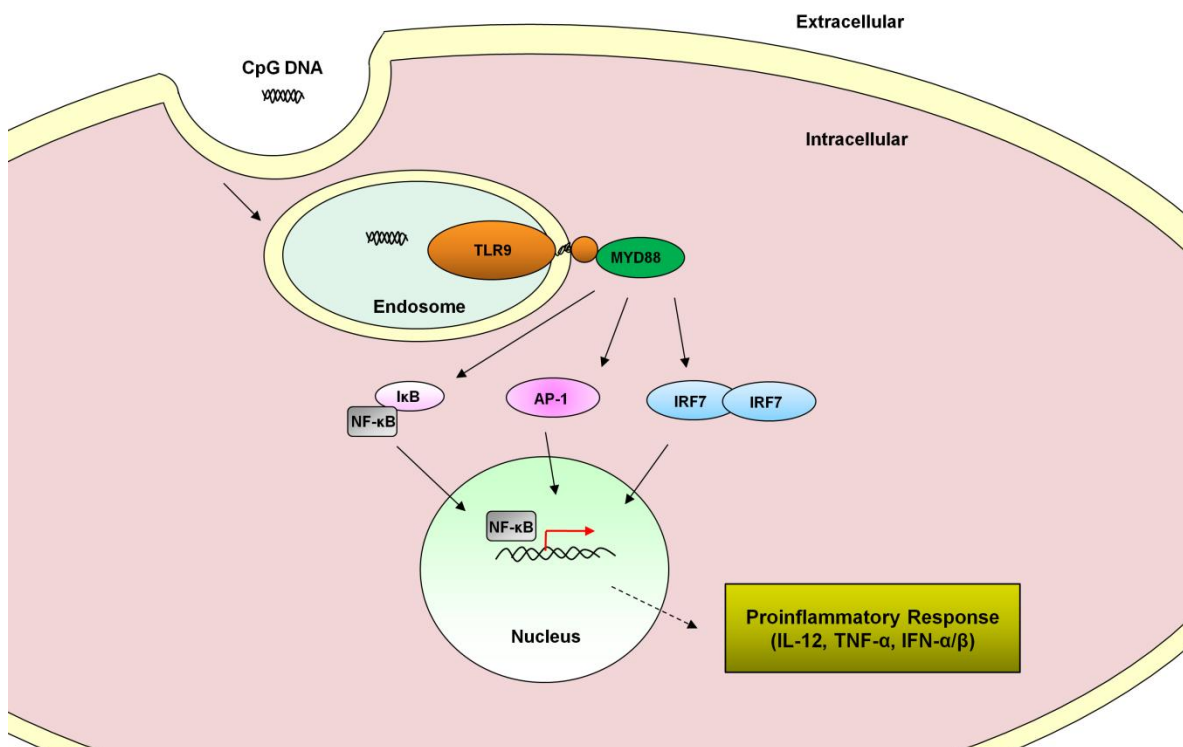
In human milk, bioactives can be proteinaceous in nature or can occur as live cells, either somatic or of other origin (43). For example, colostrum contains leukocytes (5×10^8 to 4×10^9 leukocytes/L) that provide a surveillance system for the naïve immune system of the infant (44). As the number of human microbiome studies increases, the biological importance of microbes and the influence of one's microbiome on one's health have come to light. Numerous studies have documented live bacteria within human milk, a bodily fluid once believed to be sterile. The bacteria within human milk include *Streptococcus* and *Bifidobacterium*, which can be cultured from human milk expressed from a sterilized breast (45, 46). Sequencing of the 16S ribosomal RNA (rRNA) gene has allowed for the analysis of the microbiome of milk, including those that are not easily cultured. One such analysis, which followed 16 different mothers throughout lactation, found *Staphylococcus* and *Streptococcus* to be the predominant milk genera, and *Lactobacillus* and *Bifidobacterium* as minor milk microbiota members (47). Another study found the microbiome of milk to be more variable and dependent on factors such as maternal weight (48).

Further advances and improvements in next generation sequencing technology and analysis techniques, such as whole genome sequencing using the Illumina platform, have allowed for more information to be extracted or 'mined' from microbiome studies. For example, metagenome sequencing, in which the entire genomes of the microbiome are sequenced, allows for the determination of which genes are present within each member of the community and therefore which proteins and pathways may also be present. Thus, metagenomic studies shed light on

the functional capacity of microbiomes. Sequencing just the 16S rRNA portion of the microbiome as a means to classify which bacteria are present is extremely useful and is thus far the standard in taxonomical analysis of microbiomes (49). Metagenomic sequencing, however, allows for the comparison of microbiomes at a functional level as well as a taxonomical level, placing it at the forefront of microbiology (50). Prior to the experiments conducted for the third manuscript incorporated within this thesis, no metagenomic studies on the human milk microbiome, to the best of the author's knowledge, had yet been published.

In addition to proteins and whole cells, DNA itself can act as a bioactive molecule that is able to stimulate and/or suppress the immune system through recognition by Toll-like receptor 9 (TLR9, Figure 1.3; 51, 52). Due to cells present in the fluid, human milk contains DNA of both somatic and non-somatic origin. If human milk DNA contains immune stimulatory motifs, such as unmethylated cytosine phosphate guanine (CpG) dinucleotides, exposure to the immune-modulatory motifs upon ingestion would impact the mucosal immune system of the infant as demonstrated in rodent models (53). In infants, ingested DNA motifs may also affect the immune system in a similar fashion to the DNA motifs used as adjuvants in vaccinations (54). Therefore it is important to understand which types of DNA motifs are found in human milk, and whether these motifs play a role in the development of the infant immune system. Given the recent supplementation of commercially-available infant formulas with both nucleic acids and bacteria, it is of utmost importance to better understand which bacteria and DNA sequences should be added to formulas to ensure positive health outcomes for the consuming infant.

Figure 1.3. Recognition of DNA by Toll-like receptor 9. Cells are able to endocytose circulating DNA. Once in the endosome, TLR9 is able to recognize DNA motifs, such as cytosine phosphate guanine (CpG) dinucleotides and initiate a signalling cascade involving NF- κ B, AP-1 and/or IRF7. Recognition of immune-stimulatory DNA results in a proinflammatory response of interleukin 12 (IL-12), tumor necrosis factor (TNF- α), and/or interferon (IFN- α/β) production. (Modified from 51, 52).



This present thesis also incorporates three major hypotheses derived from the above questions on the metagenome of human milk. Research on the three hypotheses described below was conducted during the course of my training for the degree of Doctor of Philosophy and has been submitted as one manuscript (chapter 4).

Hypotheses and Objectives

Manuscript 3 - Human milk metagenome: a functional capacity analysis

Hypotheses:

Human milk contains a wide variety of bacteria, whose genomes house open reading frames (ORFs) for functions that permit the bacteria to survive within human milk.

The metagenome of human milk is similar to that of infant's feces both at the taxonomical and functional level.

Human milk contains DNA motifs that are immune-modulatory in nature.

Objectives:

- Sequence the metagenome of human milk using Illumina Sequencing and the MG-RAST pipeline.
- Determine the genera of bacteria within human milk and the types of ORFs in human milk that may influence bacterial presence and stability.
- Determine the similarities and differences amongst the microbiome of human milk and mother's and infant's feces at taxonomical and functional levels.
- Search for immune-modulatory DNA motifs within human milk-derived DNA.

Chapter 2

Ingested soluble CD14 from milk is transferred intact into the blood of newborn rats

**Tonya L Ward, William J Spencer, Laura DR Davis,
JoAnn Harrold, David R Mack, Illimar Altosaar.**

Published in Pediatric Research
Volume 75, pages 252-258, 2014
doi:10.1038/pr.2013.225

Contribution of authors

Tonya Ward: Contributed to study design, conducted all experiments, analyzed all data and wrote and edited the manuscript.

William Spencer: Contributed to study design and aided in animal feedings.

Laura Davis: Aided in animal feedings.

JoAnn Harrold: Contributed to study design and edited the manuscript.

David Mack: Contributed to study design and edited the manuscript.

Illimar Altosaar: Contributed to study design and edited the manuscript.

2.1 Abstract

Background

Milk acts as an edible immune system that is transferred from mother to newborn. Soluble CD14 is a protein found in significant quantities in human milk ($\sim 8\text{--}29\text{ }\mu\text{g/mL}$). At a 10-fold lower concentration in the blood ($\sim 3\text{ }\mu\text{g/mL}$), the most notable role of sCD14 is to sequester LPS of Gram-negative bacteria from immune cells.

Methods

To explore the pharmacodynamics of this milk protein and its biological fate, the biodistribution of radiolabeled sCD14 (^{14}C , ^{125}I) was monitored in 10-d-old rat pups.

Results

Up to $3.4 \pm 2.2\%$ of the radiolabeled sCD14 administered was observed, intact, in the pup blood for up to 8 h post-ingestion. Additionally, $30.3 \pm 13.0\%$ of the radiolabeled sCD14 administered was observed degraded in the stomach at 8 h post-ingestion. A reservoir of intact, administered sCD14 ($3.2 \pm 0.3\%$), however, remained in the stomach at 8 h post-ingestion. Intact sCD14 was observed in the small intestine at $5.5 \pm 1.6\%$ of the dose fed at 8 h post-ingestion.

Conclusion

The presence of intact sCD14 in the blood and the GI tract of newborns post-ingestion has implications in the development of allergies, obesity, and other inflammation-related pathogenesis later in life.

2.2 Introduction

Cluster of Differentiation 14 is a 48-kDa pattern recognition receptor first discovered as a sensor for LPS of Gram-negative bacteria. CD14 exists either as a GPI-anchored membrane protein (mCD14) on the cell surface or as a soluble protein (sCD14) found in bodily fluids. Soluble CD14 is observed in adult blood at a concentration of $3.71 \pm 0.59 \mu\text{g/mL}$ and at up to ten-fold higher concentration in human milk, $20.10 \pm 8.74 \mu\text{g/mL}$ (at 5 d PP) to $12.16 \pm 3.75 \mu\text{g/mL}$ (at 3 mo PP; 28, 29). The two forms of CD14 (membrane-bound or soluble) appear functionally interchangeable because they both can enhance proinflammatory signaling in response to LPS through TLR4, alerting the immune system of potential infections (55).

In blood, circulating sCD14 decreases LPS-related mortality and septic shock, presumably by sequestering LPS from mCD14/TLR4-expressing immune cells (30). This allows clearance of LPS from the body before activation of the immune system. Recent studies have also implicated sCD14 in inflammation-related diseases. For example, both circulating and milk-derived sCD14 levels have been correlated with fat mass in humans, and the genomic elimination of the CD14 gene in mice attenuated symptoms of obesity, such as hypertension (56–58). Furthermore, CD14 is thought to influence the type of bacteria colonizing the GI tract of infants (59). Therefore, similar to many other immunologically relevant agents present in human milk, such as serum proteins, cytokines, and immunoglobulins, milk-derived sCD14 may play a role in inflammation, development, and overall infant health, as discussed in a recent review (60).

The high concentration of sCD14 in human milk exposes a breastfeeding infant to milligram quantities of the protein per day. In an initial study, however, neither intact nor degraded portions of sCD14 were found in the feces of BF infants (36). Immunoprecipitation of sCD14 from milk and in vitro digests demonstrated that sCD14 is able to complex with other milk proteins, namely, α -lactalbumin, which protect it from degradation (40). Taken together, the combined proteolytic protection of sCD14 by milk components and lack of sCD14 in infant feces raise the possibility that sCD14 may be absorbed intact along the GI tract of the infant, as earlier suggested (36).

Whole-protein uptake across the epithelium and into the blood stream has been previously described for other milk proteins such as immunoglobulins (22). Once translocated to the blood, these milk proteins contribute to the infant's endogenous serum pool of proteins, stimulating the immune system and offering passive immunity (for review, see ref. 61). Because sCD14 levels continue to increase during the first 18 months of life, sCD14 provided by the mother via her milk may afford additional surveillance against bacteria in the GI tract or blood of the infant (59).

In healthy full-term infants, "gut closure" (a decrease in intestinal permeability with age) occurs within a few days PP, which can be altered depending on the nutrient source (such as human milk vs. formula; 62). In rodents, gut closure is further delayed and correlates with the weaning age of 17–21 days PP (63). In this study, 10-d-old rat pups were used as a model for newborn human infants in whom gut closure has not yet occurred (term infants: 1–3 d old; or preterm infants: 1–10 d old). This age was chosen because it correlates with the greatest expression of sCD14 in human milk,

which can reach concentrations as high as $67.09 \pm 27.61 \mu\text{g/mL}$ in colostrum (28, 29).

Using radiolabeled proteins as a means to track digestive fate, the hypothesis that sCD14 remains intact along the digestive tract and is absorbed intact into the circulatory system, post-ingestion was addressed.

2.3 Methods

Radiolabeling of proteins

Human recombinant sCD14 (hrCD14, R&D Systems, Minneapolis, MN) and BSA (Sigma-Aldrich, Oakville, Ontario, Canada) were labeled with [^{14}C]-methyl iodide using 52.9 mCi/mmol (sCD14) or 10 mCi/mmol (BSA; Perkin Elmer, Waltham, MA; and Sigma-Aldrich). Proteins (100 μg of sCD14 or 500 μg of BSA) were lyophilized and labeled using a previously established protocol (64).

Separately, sCD14 and BSA were labeled with [^{125}I]-NaI (MP Biomedical, Solon, OH) using Iodination Beads (Thermo Scientific, Rockford, IL) following the manufacturer's protocol with 100 μg of protein and 1 mCi of [^{125}I]-NaI in a total volume of 1.1 mL phosphate buffered saline (PBS). Proteins were recovered by centrifugation at 1,000 $\times g$ for 2 min using a 2 mL Zeba Desalt spin column (Thermo Scientific). The resultant specific radioactivity levels were as follows: [^{14}C]-sCD14, 1.63×10^5 disintegrations per minute (dpm)/ μg ; [^{14}C]-BSA, 3.44×10^4 dpm/ μg ; [^{125}I]-sCD14, 12.16×10^6 dpm/ μg ; and [^{125}I]-BSA, 13.7×10^6 dpm/ μg . Radiolabeled proteins were stored at -70°C .

Feeding studies and protein isolation

Feeding studies were conducted in accordance with the University of Ottawa's Animal Care and Veterinary Service under approval permit ID number "BMI 77" and approved by the University of Ottawa's Animal Care Committee. Sprague Dawley rat pups aged 10 d were used as a model for preweaning infants as at this age, the animals only ingest their mother's milk. In total, 26 pups (derived from three litters of

13 pups each) weighing between 11 and 26 g were fasted for 2 h at 37°C to partially empty the stomach and accommodate for the gavage protein solution (65). Pups were gavage fed 25 µg of [¹⁴C]-sCD14, [¹⁴C]-BSA, [¹²⁵I]-sCD14, or [¹²⁵I]-BSA in 250 µL of PBS using a previous technique (65). As a control, one pup per experiment was fed PBS alone. BSA was chosen as a digestible control because it is known to be easily degraded and absorbed in the GI tract (66). Animals were anesthetized with isoflurane at 0.3, 4, and 8 h post-gavage and sacrificed by cardiac puncture. Organs were harvested, flash frozen in liquid nitrogen, and stored at -70°C. The GI tract was separated into the duodenum, jejunum, ileum, and large intestine by length (7, 80, 2, and 11%, respectively; 67).

To determine the amount of sCD14 transferred to the blood, five 10-d-old Sprague Dawley pups were fasted as described above and gavage fed 25 µg of [¹²⁵I]-sCD14 or PBS alone and returned to their mother. Blood samples (10 µL) were collected from the hind leg by needle prick for up to 8 h post-gavage and stored at -70°C.

Quantifying intact and degraded proteins

Organs and luminal contents were weighed and homogenized on ice in 150 µL buffer (containing 50 mmol/L tris(hydroxymethyl)aminomethane (pH 7.4), 2 mmol/L ethylenediaminetetraacetate (EDTA), 150 mmol/L NaCl, 0.5 mmol/L dithiothreitol (DTT), and protease inhibitor cocktail; Sigma-Aldrich) using a micropestle (Eppendorf, Westbury, NY). Samples were sonicated on ice three times using a 30-s on/off cycle at 5 W, and cell debris was removed by centrifugation. Samples were

prefiltered using a 0.22- μ m spin filter (Amicon, Millipore, Bedford, MA), followed by separation through a 30,000 Da molecular weight cutoff spin filter (Amicon). Flow-through solutions (<30 kDa) and retentates (>30 kDa) were added to 5 mL of liquid scintillation cocktail (ScintiSafe-Econo1, Fisher Scientific, Ottawa, Ontario, Canada) and radioactivity levels were quantified using a Tri-Carb liquid scintillation counter (Perkin Elmer).

Determining protein intactness

Isolated proteins were subjected to size separation by sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS-PAGE) using AnyKD SDS-polyacrylamide gradient gels (Bio-Rad, Mississauga, Ontario, Canada) and electrophoresis. Gels were silver stained, exposed to a storage phosphor screen (GE Healthcare, Piscataway, NJ), and imaged using a Typhoon Trio Imager (GE Healthcare).

Detecting CD14 in rat milk

From three pups fed PBS alone, stomach contents were resuspended in PBS and pooled. Proteins, including 100 ng of hrCD14 (R&D Systems) as a positive control and a biotinylated molecular weight marker (7727, Cell Signaling Technology, Danvers, MA), were size separated using SDS-PAGE and transferred to nitrocellulose membrane. A mouse anti-human CD14 antibody (cross-reactive to rat CD14, MAB 3831, R&D Systems), a goat anti-mouse IgG horseradish peroxidase (HRP)–linked antibody (HAF007, R&D Systems), and a goat anti-biotin HRP–linked antibody to recognize the molecular weight markers (7727, Cell Signaling Technology) were used

in a western blot of the milk proteins. Proteins were visualized using exposure to enhanced chemiluminescent (ECL) substrate (GE Healthcare).

Data analysis

Percentage dose was calculated using the formula $[(\text{dpm/g organ}) \times (\text{total organ weight})]/[\text{dpm dose fed}]$. Total organ weight of blood was calculated using the approximation that blood represents 7% of the body weight. Data were analyzed by two-tailed paired or unpaired *t*-tests (Sigma Plot 12.1, Systat Software, Inc., San Jose, CA). An alpha level of <0.05 was considered significant.

2.4 Results

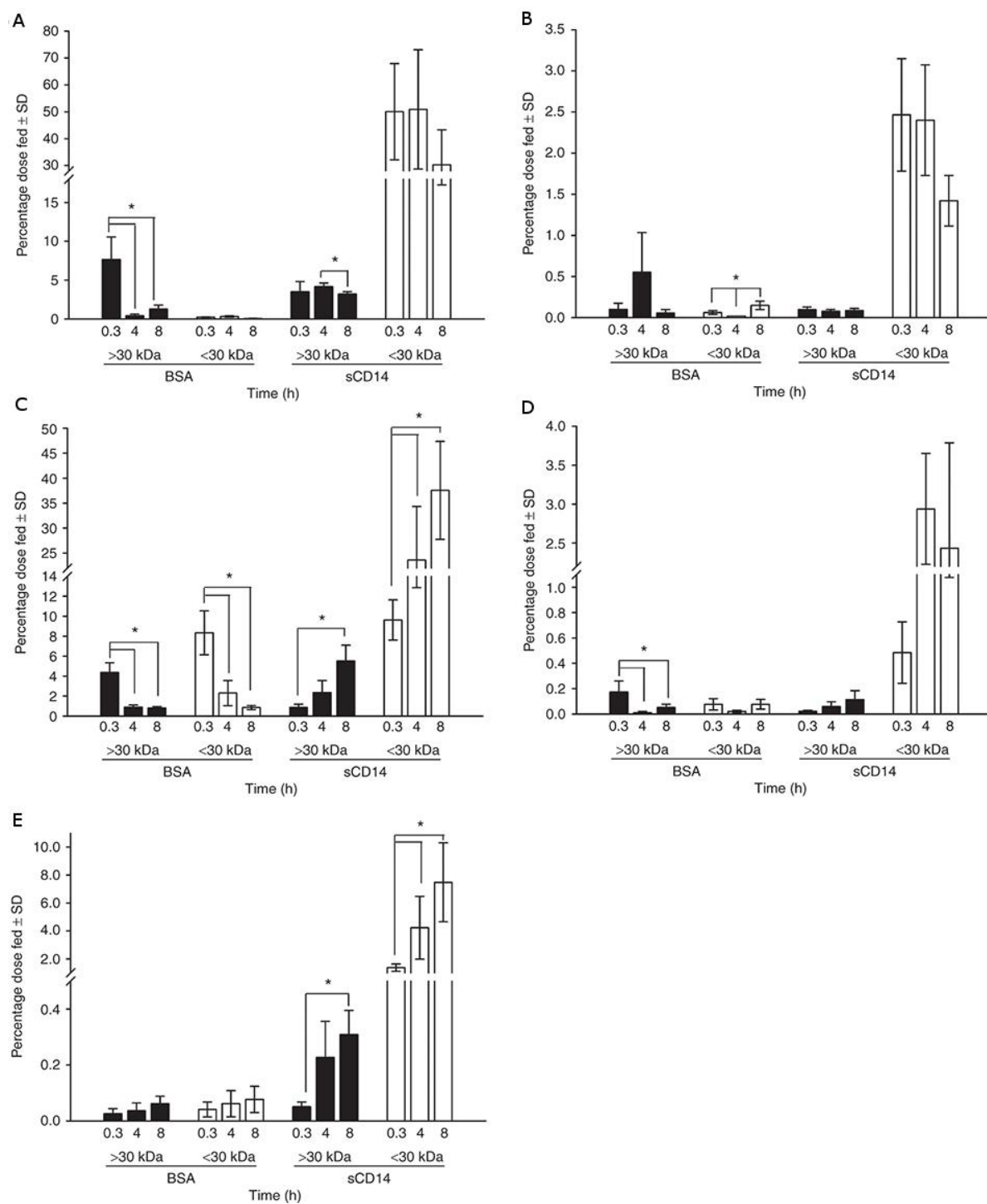
Ingested sCD14 persists in the GI tract

Within the stomach, $3.2 \pm 0.3\%$ of the [^{14}C]-CD14 dose fed remained intact at 8 h post-ingestion, which did not significantly differ from the amount at 0.3 h post-ingestion ($3.5 \pm 1.3\%$, $P = 0.838$; Figure 2.1A). The amount of degraded [^{14}C]-CD14 at 0.3 h post-ingestion ($50.0 \pm 17.9\%$) also did not significantly decrease over the duration of the experiment ($30.3 \pm 13.0\%$ at 8 h, $P = 0.423$). This differs from the profile of the digestible control protein, BSA. In the stomach, $7.7 \pm 2.9\%$ of the [^{14}C]-BSA dose fed was present intact at 0.3 h, which significantly decreased to $1.3 \pm 0.5\%$ by 8 h post-ingestion ($P = 0.027$; Figure 2.1A). The stomach also contained no more than $0.3 \pm 0.1\%$ of the dose fed of degraded [^{14}C]-BSA at all time points.

In the duodenum, no more than $2.5 \pm 0.7\%$ of the [^{14}C]-CD14 dose fed, degraded or intact, was observed at any time point, probably due to the small size of the tissue (Figure 2.1B). A similar trend was observed for ingested [^{14}C]-BSA, with no more than $0.5 \pm 0.5\%$ present in the duodenum, degraded or intact, at any time point (Figure 2.1B). In the jejunum, however, the amount of intact [^{14}C]-CD14 increased from $0.9 \pm 0.3\%$ of the dose fed at 0.3 h to $5.5 \pm 1.6\%$ of the dose fed at 8 h ($P = 0.045$; Figure 2.1C). Similarly, the amount of degraded [^{14}C]-CD14 in the jejunum increased from $9.6 \pm 2.0\%$ at 0.3 h to $37.5 \pm 9.8\%$ of the dose fed at 8 h post-ingestion ($P = 0.049$). The control protein, BSA, showed an opposite trend (Figure 2.1C). In the jejunum, the amount at 0.3 h for both intact and degraded BSA (4.4 ± 1.0 and $8.3 \pm 2.2\%$, respectively) significantly decreased by 8 h ($0.8 \pm 0.1\%$ ($P = 0.024$) and $0.8 \pm 0.2\%$ ($P = 0.027$), respectively).

Similar to the jejunum, the ileum and the large intestine demonstrated an increase in both intact and degraded [^{14}C]-CD14 from 0.3 to 8 h post-ingestion. For example, in the large intestine, the $0.05 \pm 0.01\%$ of the dose fed of intact [^{14}C]-CD14 significantly increased to $0.3 \pm 0.1\%$ of the dose fed at 8 h post-ingestion ($P = 0.042$; Figure 2.1E). The amount of intact BSA, however, decreased in the ileum over time (0.2 ± 0.09 to $0.05 \pm 0.03\%$ of the dose fed, from 0.3 to 8 h, $P = 0.045$), and the amount of BSA in the large intestine showed no significant change across time points (Figure 2.1E).

Figure 2.1. Biodistribution of intact (>30 kDa, black) and degraded (<30 kDa, white) [¹⁴C]-sCD14 and a digestible control protein, [¹⁴C]-BSA, post-ingestion in 10-d-old rat pups. (A) Stomach; (B) duodenum; (C) jejunum; (D) ileum; and (E) large intestine. Rat pups were gavage fed 25 µg of [¹⁴C]-sCD14 (1.63×10^5 dpm/µg) or [¹⁴C]-BSA (3.44×10^4 dpm/µg). Post-sacrifice (0.3, 4, or 8 h), proteins were extracted from organs, size separated using a 30-kDa cutoff spin column, and quantified using liquid scintillation counting. Data are presented as mean $\pm$ SD, and n=3 for each time point. * $P < 0.05$, using a Student's *t*-test.



Ingested sCD14 is transferred to the blood

Intact (48 kDa) and degraded [125 I]-sCD14 (30 and 25 kDa) were observed in the stomach for up to 8 h (Figure 2.2A). By exaggerating the image contrast, intact [125 I]-sCD14 was also observed in the jejunum for up to 8 h post-gavage (Supplemental Figure S2.1). No [125 I]-sCD14 was detected in the duodenum or ileum, probably due to the small size of tissue and relative protein content (Supplemental Figure S2.2). Most interestingly, intact [125 I]-sCD14 (48 kDa) was detected in the blood at 4 h post-ingestion and persisted up to 8 h post-ingestion (Figure 2.2A). In comparison, both intact (66 kDa) and degraded [125 I]-BSA (50, 40, and 30 kDa) were observed in the stomach for up to 8 h (Figure 2.2B). [125 I]-BSA was observed in both intact and in multiple degraded forms (50, 40, and 30 kDa) in the duodenum, jejunum, and ileum samples. Neither intact nor degraded [125 I]-BSA was detected in the blood at any of the sampling times (Figure 2.2B).

In a separate experiment, intact [125 I]-sCD14 was detected in the blood as early as 0.3 h post-gavage ($0.5 \pm 0.4\%$ of the dose fed), which significantly increased to $2.5 \pm 1.2\%$ of the dose fed by 1 h post-gavage ($P = 0.021$; Figure 2.3). The percentage of intact [125 I]-sCD14 in the blood did not significantly change after 1 h ($3.4 \pm 2.2\%$ of the dose fed at 8 h, Figure 2.3). Degraded [125 I]-sCD14 was also observed in the blood by 0.3 h post-ingestion, at a concentration of $1.08 \pm 0.5\%$ of the dose fed. The concentration of degraded [125 I]-sCD14 in the blood at 0.3 h significantly increased at 4 h post-ingestion ($P = 0.024$), at which time it reached $4.9 \pm 2.0\%$ of the dose fed (Figure 2.3).

Figure 2.2. Biodistribution of (A) [¹²⁵I]-sCD14 and (B) a digestible control protein, [¹²⁵I]-BSA, post-ingestion in 10-d-old rat pups. Rat pups were gavage fed 25 µg of [¹²⁵I]-sCD14 (12.16×10^6 dpm/µg) or [¹²⁵I]-BSA (13.7×10^6 dpm/µg). Post-sacrifice (0.3, 4, or 8 h), proteins were extracted from harvested organs, size separated by SDS-PAGE, and phosphor imaged. Stomach (St) and contents (St Con), duodenum (Duo) and contents (Duo Con), jejunum (Je) and contents (Je Con), ileum and contents (Ileum Con), large intestine (L In) and contents (L In Con), liver, and blood are shown. Figure is representative of n=3.

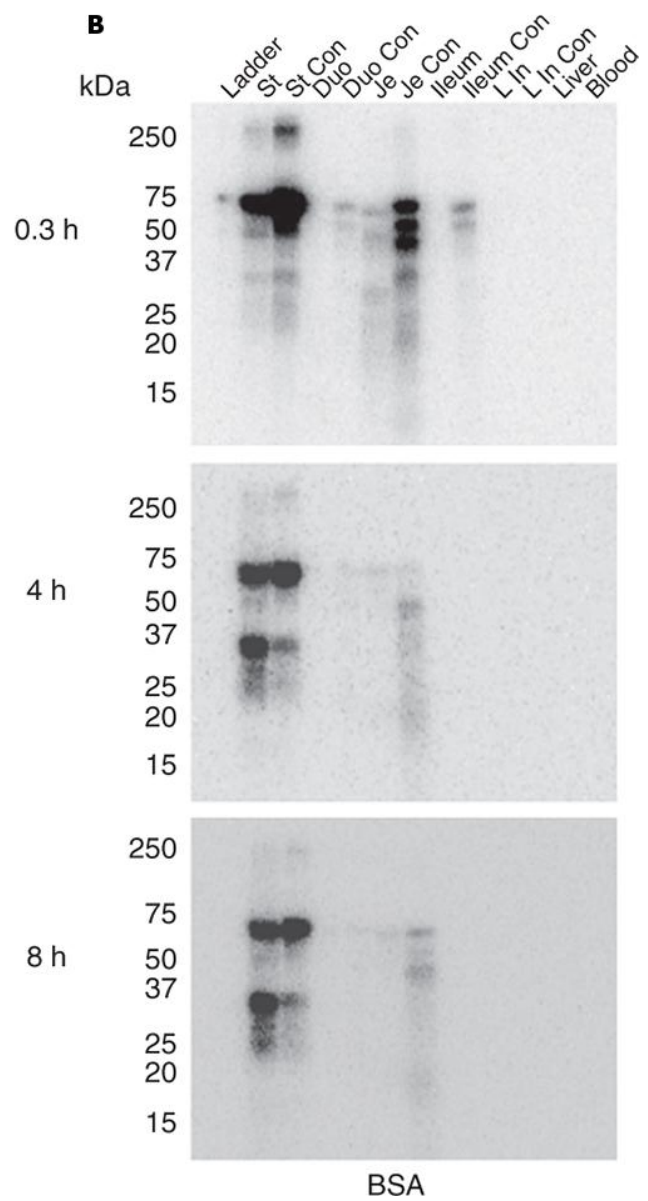
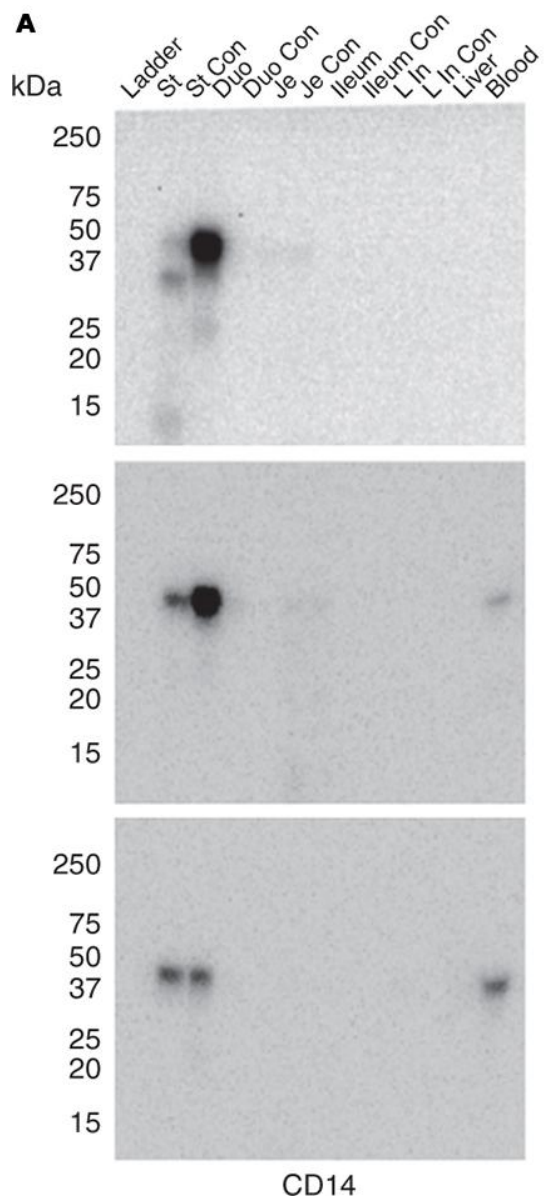
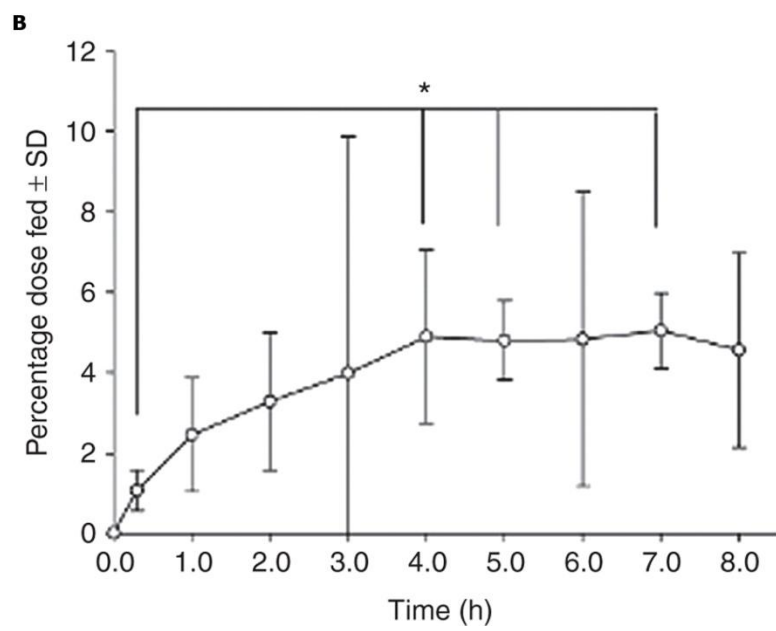
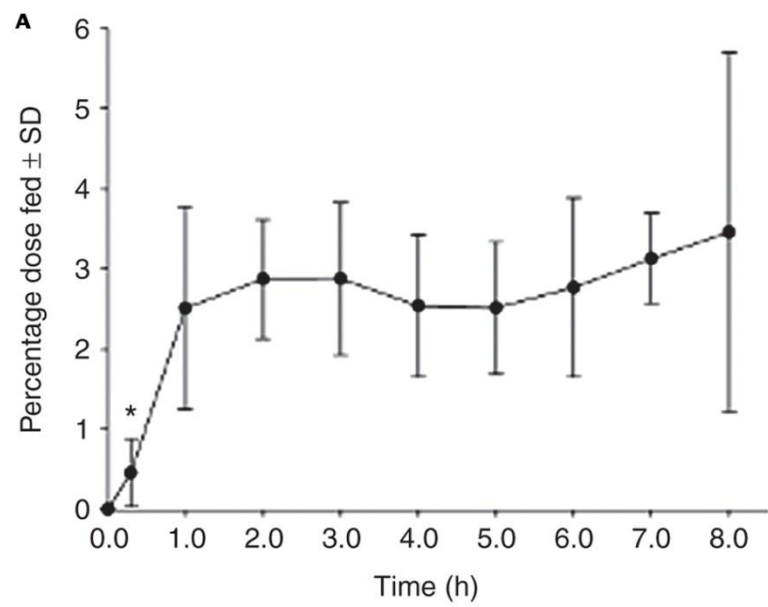


Figure 2.3. Uptake of (A) intact (>30 kDa) and (B) degraded (<30 kDa) [¹²⁵I]-sCD14 into the blood of 10-d-old rat pups. Data are presented as mean ± SD, and n=4. **(A)** Percentage dose fed of sCD14 >30 kDa at 0.3 h was significantly less than that at all subsequent time points (**P* < 0.05). **(B)** Percentage dose fed of sCD14 <30 kDa at 0.3 h was significantly less than that at 4, 5, and 7 h (**P* < 0.05, Student's *t*-test).



Using SDS-PAGE and phosphor imaging, the circulating [¹²⁵I]-sCD14 was confirmed to be fully intact (48 kDa; Figure 2.4A). Degraded sCD14 was not detected by these methods, probably due to its small size (<10 kDa). This was confirmed by silver staining the SDS-PAGE gels of the >30 or <30 kDa size-fractionated samples, where no protein was observed in the <30 kDa fraction (Figure 2.4B).

Rat milk contains sCD14

Because the pups were ingesting rat milk from their mother pre- and post-gavage, any unlabeled rat sCD14 within the milk may have hindered uptake of labeled CD14 across the GI tract. A western blot using milk clots from pups fed only PBS showed the presence of sCD14 in rat milk (Figure 2.5).

Figure 2.4. Intactness of ingested [¹²⁵I]-sCD14 in the blood of 10-d-old rat pups after size separation (< or >30 kDa) visualized by (A) phosphor imaging and (B) silver staining of SDS-PAGE gels. Intact and non-ingested [¹²⁵I]-sCD14 was used as a positive control. Figure is representative of n=4.

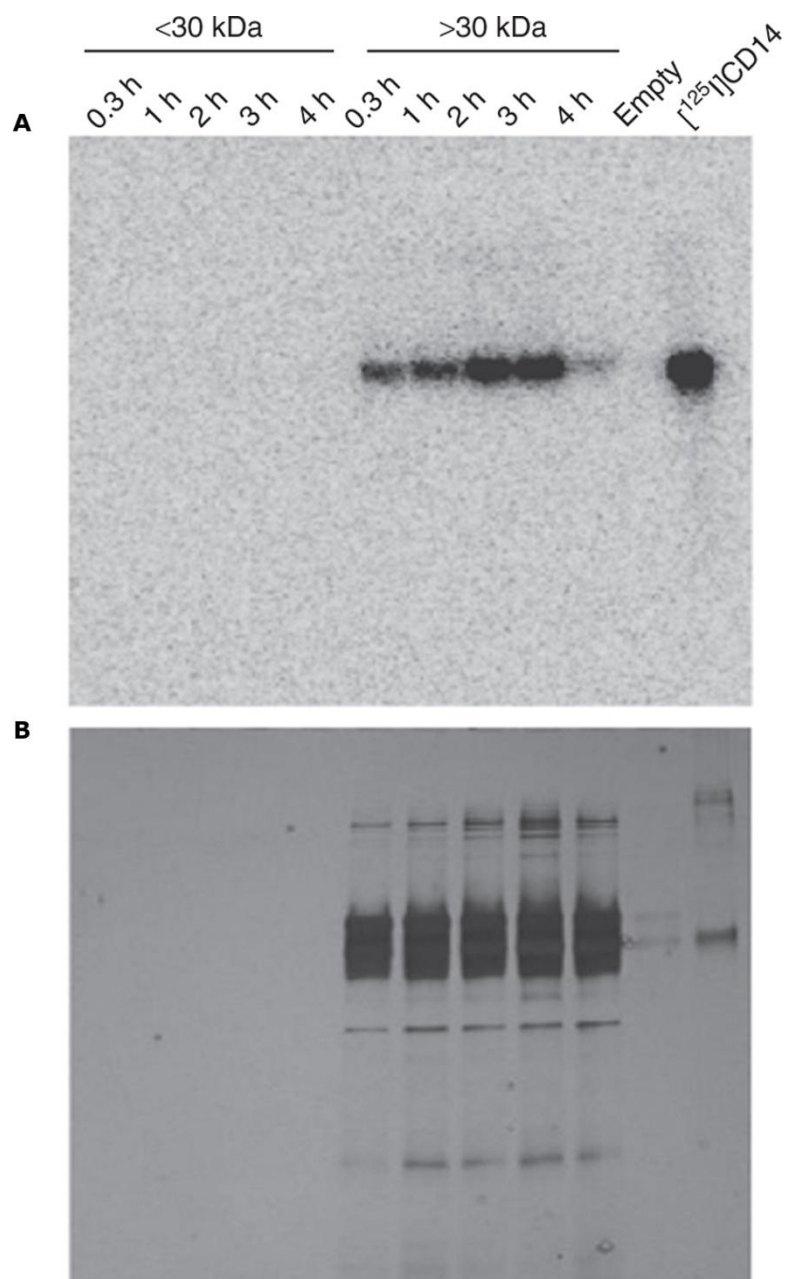
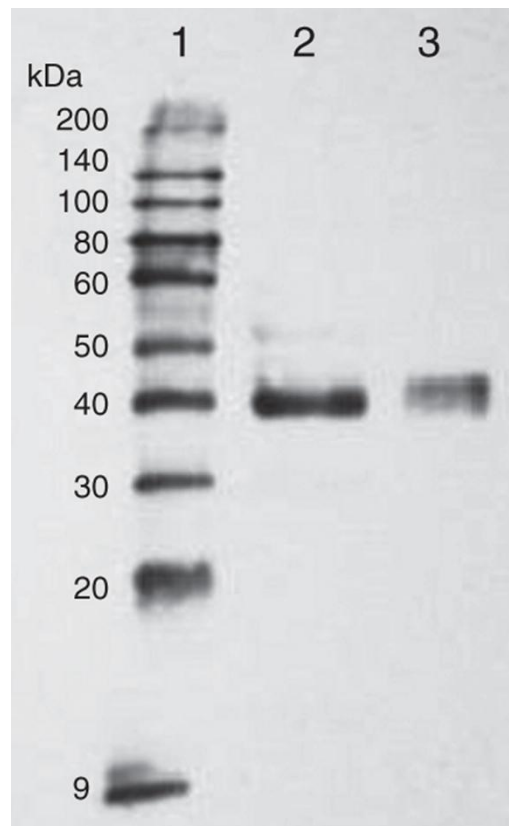


Figure 2.5. The presence of rat sCD14 in rat milk, at 10 d PP, as detected by immunoblotting. Lanes: 1, molecular weight marker; 2, rat milk; 3, human recombinant sCD14 (100 ng, positive control).



2.5 Discussion

Evasion of digestion by milk proteins

The bioactive components of human milk that promote development and provide protection to the infant, including the bacterial sensor sCD14, must be proteolytically resistant to digestion or must form complexes with other proteins to avoid degradation and retain functionality in the infant's GI tract or circulatory system. Secretory IgA and lactoferrin are well-known examples of bioactive proteins that survive passage through the infant's digestive system (21, 36). Similarly, in this study, sCD14 was able to evade digestion in the GI tract because $3.2 \pm 0.3\%$ of ingested [^{14}C]-CD14 remained undigested in the stomach for up to 8 h. In comparison, the digestible control, BSA, decreased from $7.7 \pm 2.9\%$ at 0.3 h to $1.3 \pm 0.5\%$ of the dose fed at 8 h post-ingestion (Figure 2.1A). Soluble CD14 was also able to enter and persist in the small intestine because $5.5 \pm 1.6\%$ of the dose fed of sCD14 was observed intact in the jejunum at 8 h post-ingestion (Figure 2.1C). This persistence may, in part, be due to the interaction of sCD14 with other milk proteins. For example, α -lactalbumin was previously shown to interact with sCD14 in milk and protect it from digestion in vitro (40). Rat milk, which was ingested pre- and post-gavage by the pups, contains α -lactalbumin at a concentration of 1.5 g/L at 8 d PP, which increases to 3 g/L at 15 d PP (68). In vivo, α -lactalbumin in rat milk probably confers the same proteolytic protection of sCD14 observed in vitro.

The high percentage of degraded sCD14 observed in all organs using liquid scintillation counting (Figure 2.1, white bars) was not observed by SDS-PAGE and phosphor imaging (Figure 2.2A). This is probably due to the small size of the peptides

produced during digestion (<10 kDa), which would migrate off the gel, limiting detection. Previous in vitro work showed human milk sCD14 to be partly resistant to pepsin digestion, which was confirmed herein through the observation of intact sCD14 in the stomach for up to 8 h post-ingestion (Figure 2.2A; 36). Additionally, in vitro pepsin and pancreatin digestion—simulating the small intestine—showed human milk sCD14 to be proteolytically sensitive, and the digestion produced fragments of ~20 kDa (36). In this report, [¹⁴C]-sCD14 of <30 kDa was detected in the small intestine at all time points (Figure 2.1b-d), but these fragments were not detected using phosphor imaging (Figure 2.2). This may be due to the location of the radiolabel along the peptide chain, which may have prevented detection of some fragments depending on how the protein was digested.

Activity of digested and intact sCD14

Whether it is intact or degraded, ingested sCD14 may be acting as an important regulator of development in the infant GI tract. Intact sCD14 in the GI tract may aid in dampening the immune response to the influx of colonizing bacteria as the child develops just as sCD14 is able to neutralize bacterial LPS in the blood (30, 69). Similarly, *degraded* sCD14 may provide protection to the infant because sCD14 peptides as small as 20 amino acids (residues 81–100 especially) have been shown to contain effective LPS-neutralizing capabilities (70). Furthermore, a fragment of CD14 (residues 139–160) has been shown to protect human lymphocytes from gliotoxin-induced death, and sCD14 has been shown to be a survival factor for leukemic cells (71, 72). If sCD14 behaves similarly along the GI tract of infants, it may be promoting

gut closure and intestinal epithelial cell homeostasis. In support of this, it has been shown that digested infant formulas, which notably lack sCD14, increase death of intestinal cells, whereas digested human milk preserves their function (73).

Strikingly, sCD14 was detected intact in the blood at 4 h post-gavage, where it remained for up to 8 h (Figure 2.2a). This finding was confirmed by a second experiment in which the blood of gavage-fed pups was sampled each hour following sCD14 ingestion (Figures 2.3a and 2.4a). Approximately 3% of the ingested sCD14 was absorbed intact into the blood, which did not significantly change up to 8 h post-ingestion (Figure 2.3a). When extrapolated, this amount may represent up to 30% of a newborn's circulating sCD14. As more studies describing the milk constituents that are able to cross the infant epithelial barrier are published, the notion of "you are what you eat" gains validity. The transfer of beneficial immune components, such as the sCD14 reported here, immunoglobulins, and noninherited maternal antigens, can positively affect the infant's health by conferring additional immunity (74, 75). For example, once ingested, noninherited maternal antigens present in the mother's milk convey tolerance to the infant during bone marrow transplantation later in life (75). This type of donated protection via milk into the infant's blood is probably conserved across other immune proteins, such as sCD14.

Potential mechanisms of CD14 uptake

Whole-protein uptake of milk components has been well described for immunoglobulins and lactoferrin, which are receptor mediated and conserved across species, including humans and rodents (22, 23). One potential route of sCD14 uptake

may involve the bacterial recognition receptor TLR4, which is present on multiple gut-associated cell types, including intestinal epithelial cells of adults and newborns (42). TLR4 interacts with both mCD14 and sCD14 to signal the presence of bacteria and alert the immune system. In vitro studies have shown that TLR4 internalization during LPS exposure is CD14 dependent in B cells, and translocation of LPS and whole bacteria by intestinal epithelial cells has been demonstrated in vitro and in vivo (76, 77). This internalization of LPS and whole bacteria may represent a potential mechanism for sCD14 uptake in vivo along the GI tract. Soluble CD14 may “piggyback” on luminal LPS while the latter are being translocated. If the uptake of sCD14 is specific and receptor mediated, the proportion ($3.4 \pm 2.2\%$) of the intact [^{125}I]-sCD14 fed that is translocated to the blood at 8 h may be an underrepresentation of the total amount of sCD14 absorbed in the GI tract (Figure 2.3A). Because rat milk contains endogenous rat sCD14 (Figure 2.5), unlabeled endogenous protein may be competing for the uptake mechanisms used by radiolabeled sCD14, thereby hindering uptake of the labeled protein. Consequently, further experiments to better quantify the total amount of sCD14 translocated to the blood are warranted.

Possible role of ingested CD14

In the blood, the classic role of circulating sCD14 is to attenuate bacteria-induced immune responses by sequestering bacterial components, including LPS, from immune cells and protecting against exaggerated immune responses such as septic shock (30). The exact role of *ingested* sCD14 in the infant is yet to be determined, but the molecule’s newly found associations with inflammation-related diseases suggest

that its transfer from mother to infant holds significance. For example, there are numerous situations where removing both mCD14 and sCD14 expression results in attenuation of inflammation-related disease phenotypes. In obesity, there is a chronic low-grade state of inflammation, which is affected by the presence of CD14. When the CD14 coding region was removed from the genome, mice still possessed the ability to become obese, but they gained less fat compared with WT mice and suffered from less obesity-related phenotypes, such as hypertension (57). Another closely related example is the effect of CD14 on insulin resistance. Specifically, when WT mice were given a bone marrow transplant with CD14-null cells, they no longer became glucose intolerant or insulin resistant when fed a high-fat diet (58). Therefore, sCD14 levels may affect inflammatory responses in the infant.

In the context of LPS response, elevated levels of sCD14 were able to attenuate exaggerated immune responses in mice in otherwise-normal physiological contexts (69). Perhaps the presence of additional sCD14, obtained from the mother's milk, provides similar protection for the infant both during the LPS response and in other systems. For example, delivering additional hrCD14 via injection to adult WT mice was able to increase glucose tolerance and increased the action of insulin in the same manner as a total CD14 knockout (58). Furthermore, the ability of additional sCD14 to ameliorate negative inflammatory symptoms was also seen in allergy development. It has been shown that children with allergies often have decreased levels of circulating sCD14 and that exogenous sCD14 was able to suppress allergen-induced T helper cell 2 differentiation in vitro (59, 78). The link between the level of sCD14 in mother's

milk and the development of atopy and eczema in the infant, however, remains controversial (79–81).

Conclusions

The significance of the transfer of ingested sCD14 from human milk into the circulation of the infant should not be underestimated. As new roles of sCD14 in disease pathogenesis are uncovered, the protein's impact on infant development may become clearer.

Chapter 3

Ingested soluble CD14 contributes to the functional pool of circulating sCD14 in mice

Tonya L Ward, Kagami Goto, Illimar Altosaar.

In print at Immunobiology
doi: 10.1016/j.imbio.2014.03.008

Contribution of authors

Tonya Ward: Conceived study design, conducted all experiments, analyzed all data and wrote and edited the manuscript.

Kagami Goto: Aided in microscopy.

Illimar Altosaar: Contributed to study design and edited the manuscript.

3.1 Abstract

Background

Soluble CD14 is a pattern recognition receptor and Toll-like co-receptor observed in human milk (5-26 $\mu\text{g/ml}$) and other bodily fluids such as blood (3 $\mu\text{g/ml}$). The most well defined role of sCD14 is to recognize LPS of Gram-negative bacteria and signal an immune response through TLR4. Previous research has shown ingested sCD14 to transfer from the GI tract and into the blood stream in neonatal rats. The contribution of human milk sCD14 to circulating levels in the infant and the functionality of the protein, however, remain unknown.

Results

Using CD14^{-/-} mouse pups fostered to WT mothers expressing sCD14 in their milk, ingestion of sCD14 resulted in blood sCD14 levels up to $0.16 \pm 0.09 \mu\text{g/ml}$. This represents almost one third (26.7%) of the circulating sCD14 observed in WT pups fostered to WT mothers ($0.60 \pm 0.14 \mu\text{g/ml}$). Ingested-sCD14 transferred to the blood also remained functional in its ability to recognize LPS as demonstrated by a significant increase in immune response (IL-6 and TNF- α) in CD14^{-/-} pups fostered to WT mothers in comparison to control animals ($P=0.002$ and $P=0.007$, respectively). Using differentiated human epithelial colorectal adenocarcinoma cells (Caco-2), a significant decrease in sCD14 transcytosis was observed when TLR4 was knocked down ($P<0.001$), suggesting sCD14 transfer is TLR4-dependent.

Conclusions

The bioavailability of ingested sCD14 established in this study confirms the importance of human milk proteins for the infant and demonstrates the need to improve infant formulas which are lacking in immune proteins such as sCD14.

3.2 Introduction

Human milk consumption in comparison to infant formula has been linked to beneficial health outcomes for the infant such as decreased incidence of necrotizing enterocolitis (NEC; 82, 83). The bioactive components of human milk linked to health benefits for infants include prebiotics, probiotics and innate immune components. One such immune component is CD14. CD14 is a pattern recognition receptor found in two forms, mCD14 on the surface of cells, or as a sCD14 in bodily fluids such as tears (0.5 µg/mL), blood (3 µg/ml) and milk (5-26 µg/ml), depending on time of lactation (56, 84). Both forms of CD14 are responsible for the detection of several bacterial components including LPS of Gram-negative bacteria (85). In combination with TLR4, detection of LPS by CD14 results in a proinflammatory immune response through both the TIRAP-MyD88 dependent pathway and the TRAM-TRIF dependent pathway (86, 87). sCD14, in comparison to mCD14, can also decrease the immune response to LPS by sequestering LPS from mCD14-expressing cells and providing clearance of LPS through the liver (88, 89).

To provide benefit for the infant, bioactive milk components, including sCD14, must remain intact and functional post-ingestion. In a previous study, a significant portion ($3.4 \pm 2.2\%$) of radiolabeled-hrCD14 survived GI transit post-ingestion and was transferred, intact, into the blood of rat pups (90). The total amount of endogenous milk CD14 transferred into the blood and its functionality, however, remain unknown. In this study, soluble CD14 was hypothesized to be a major contributor to circulating sCD14 in infants, and once transferred to the blood, sCD14 was hypothesized to remain functional.

Additionally, this study attempted to validate sCD14 transfer across the GI tract using human intestinal epithelial cells, and to elucidate the mechanism by which sCD14 is transferred from the GI tract into the blood. In macrophages and dendritic cells, internalization of TLR4 has been demonstrated through a CD14 interaction (91). In addition to immune cells, intestinal epithelial cells also express TLR4 (42), and therefore transfer of sCD14 across human intestinal cells was hypothesized to occur in a TLR4 dependent mechanism.

3.3 Methods

Animals, fostering and LPS injection

Animal studies were conducted in accordance with the University of Ottawa's Animal Care and Veterinary Service under approval permit ID number 'BMI-129', and were approved by the University of Ottawa's Animal Care Committee. CD14 null mice (B6.129S-*Cd14^{tm1Frm}*/J) and WT controls of the same background (C57BL/6J) were purchased from Jackson Laboratory (Bar Harbor, Maine, USA). Using an experimental design outlined in Table 3.1, mice were bred and pups were fostered at day 6 PP to different mothers. Starting at day 0 (pre-foster) until day 4 post-foster, pups were removed from the litter, euthanized by CO₂ inhalation, and organs were harvested (Supplemental Figure S3.1 for the fostering timeline). Separately, pups were fostered as described above and injected intraperitoneally on day 2 post-foster (day 8 PP) with LPS (200 ng/g body weight, Sigma-Aldrich). Pups were monitored for signs of discomfort and at 2 hours post injection pups were euthanized by CO₂ inhalation and organs were harvested (90).

CD14 and cytokine concentrations

Milk clots from the stomach of pups (15 mg) were resuspended in 150 µL of protein extraction buffer (50 mM Tris-HCl pH 7.4, 2 mM EDTA, 150 mM NaCl, 0.5 mM DTT and protease inhibitor cocktail, Sigma-Aldrich). Stomach contents and blood samples were analyzed using mouse-specific enzyme linked immunosorbent assays (ELISAs) for CD14, TNF-α and IL-6 (R&D Systems). Stomach contents and blood samples were also analyzed by western blot to determine CD14 intactness. Samples, including 30 ng of

Table 3.1. Fostering scheme used to determine the quantity and functionality of ingested sCD14 transcytosed to the blood of mice.

Pup	Foster Mother	Endpoint
WT	WT	- Total circulating sCD14 - Positive control for LPS detection by sCD14
WT	CD14 ^{-/-}	- Contribution of pup-CD14 to circulating sCD14 - Impact of lack of ingested sCD14 on LPS detection
CD14 ^{-/-}	WT	- Contribution of ingested sCD14 to circulating sCD14 - Functionality of transferred sCD14 to detect LPS
CD14 ^{-/-}	CD14 ^{-/-}	- Negative control

mouse recombinant CD14 as a positive control (R&D Systems), were electrophoresed using SDS-PAGE and proteins were transferred to nitrocellulose membranes which were blocked with 5% BSA-Tris buffered saline with Tween (TBST) followed by 0.5 µg/mL rabbit anti-CD14 antibody (M-305), Santa Cruz Biotechnology, Dallas, TX, USA) for one hour. Membranes were washed three times and incubated with a 1:30,000 dilution of a chicken anti-rabbit IgG HRP-linked antibody (SC-2955, Santa Cruz Biotechnology). Following an additional three washes, proteins were visualized using exposure to ECL substrate (GE Healthcare).

Caco-2 cells and transport assays

Human intestine-derived Caco-2 and HT29 cells were grown in Eagle's minimum essential media (Life Technologies, Burlington, ON, Canada) with a final concentration of 2 mM L-glutamine, 1 mM sodium pyruvate, sodium bicarbonate, 1x Antibiotic-Antimycotic (Life Technologies) and 20% FBS (Life Technologies) at 37°C and 5% CO₂. Biocoat HTS Caco-2 transport assays were completed as per the manufacturer's protocol (BD Biosciences, Franklin Lakes, NJ, USA). Briefly, cells were seeded on transport membranes and differentiated for three days. Human recombinant CD14 (R&D Systems) was added to the apical well. Media from both the apical and basal well was collected 2 hours post-addition for analysis by human CD14-specific ELISA (R&D Systems). Passive transport of Lucifer Yellow (444.25 mol mass) was used to determine monolayer intactness as per the manufacturer's protocol. Lucifer Yellow (70 µM, Sigma-Aldrich) was added in addition to the soluble CD14 and was detected in the media using a Typhoon Trio Imager (GE Healthcare). Intactness of hrCD14 was

determined by western blot, as described above with the exception of using an anti-human CD14 primary antibody (1:1000 dilution, R&D Systems) and a rabbit anti-mouse IgG HRP-linked secondary antibody (1:10,000 dilution, HAF007, R&D Systems).

TLR4 siRNA knockdown

Caco-2 cells and HT29 cells were grown as described above. RNA was isolated from cells using an RNeasy mini kit (Qiagen, Valencia, CA, USA) and subsequently used to create cDNA using a Maxima cDNA kit (Thermo Scientific). TLR4 and glyceraldehyde 3-phosphate dehydrogenase (GAPDH) cDNA was amplified using PCR conditions previously described and the following primers

TLR4: 5'-CTGCGTGAGACCAGAAAGC-3', 5'-TTCAGCTCCATGCATTGATAA-3' (75 bp amplicon); GAPDH: 5'-AGCCACATCGCTCAGACAC-3',

5'-GCCCAATACGACCAAATCC-3' (66 bp amplicon; 92). Caco-2 cells were seeded and differentiated on a Biocoat HTS transport assay as described above (BD Biosciences).

Following the manufacturer's instructions, cells were transfected with siRNA for TLR4 (SC-40260) or scrambled siRNA as a control (Santa Cruz Biotechnology). After 48 hours, hrCD14 (24 µg/ml, R&D Systems) was added to the apical well and transfer to the basal well was measured via ELISA, as described above. RNA was isolated from Caco-2 monolayers and analyzed by RT-PCR as described above.

Immunocytochemistry and microscopy

Following transport assays, monolayers were washed twice with PBS and fixed with neutral buffered formalin. Fixed monolayers on Transwell membranes were excised from the plastic trays, hydrated for one hour with buffer (0.1% tween, 0.15M NaCl, 5mM EDTA 20mM HEPES pH 7.5, 0.02% NaN₃) and blocked with 1% BSA in buffer for 30 min. Membranes were incubated overnight with mouse anti-CD14 antibodies (1µg/mL, UCH-M1, Santa Cruz Biotechnology) at 4°C, washed three times with buffer, and incubated with FITC-labeled goat anti-mouse IgG antibodies (10 ng/mL, F0257, Sigma-Aldrich) for 1.5 hours at room temperature. Membranes were washed three times with buffer and stained with Hoechst (10µg/mL, Sigma-Aldrich) for 1 hour at 4°C, followed by three more washes with buffer. Membranes were mounted on slides using ProLong Gold Anti-fade Reagent (Life Technologies) with a coverslip.

Microscopy was performed using an Axio Observer D1 microscope (Carl Zeiss, Göttingen, Germany) and images were prepared using Axio Vision 4.8.2 software (Carl Zeiss). Relative signal intensities were determined using ImageJ (NIH, Bethesda, MD).

Statistical analysis

Data were analyzed by 2-tailed paired or unpaired *t*-tests (Sigma Plot 12.1, Systat Software, Inc.). An alpha level of <0.05 was considered significant. In the cases where the measurands were below the level of detection (see Figures, “Not detected”) values were assumed to be zero.

3.4 Results

Milk sCD14 is transferred to the blood of pups

In CD14^{-/-} pups, CD14 was not detected in either the stomach contents or blood prior to fostering the pups to different mothers (Figure 3.1E, F). Once fostered to WT mothers, sCD14 was detected in the blood of CD14^{-/-} pups at day eight PP (two days post-foster) at a concentration of 0.16 ± 0.09 µg/ml, which decreased to non-detectable levels by day ten PP (4 days post-foster, $P < 0.001$, Figure 3.1F). This decrease in sCD14 coincided with a decrease in sCD14 concentrations in the milk of WT mice, which significantly decreased from 2.23 ± 0.37 µg/ml at day seven PP to 0.02 ± 0.01 µg/ml at day ten PP ($P < 0.001$, Figure 3.1A). The endogenous level of circulating sCD14 observed in mouse pups changed over time. In WT pups fostered to WT mothers, sCD14 was observed at 0.69 ± 0.14 µg/ml in the blood which significantly decreased to 0.31 ± 0.04 µg/ml by day nine PP ($P = 0.011$, Figure 3.1B). In WT pups fostered to CD14^{-/-} mothers, sCD14 concentrations in the blood were similar to that of WT pups fostered to WT mothers. At day six PP, sCD14 was detected at 0.63 ± 0.15 µg/ml in the blood of WT pups (pre-foster, Figure 3.1D). The concentration of sCD14 in the blood decreased to 0.16 ± 0.01 µg/ml on day seven PP when fostered to CD14^{-/-} mothers ($P = 0.006$), and was similar to the amount in WT pups fostered to WT mothers by day nine PP (0.20 ± 0.10 µg/ml $P > 0.05$). Soluble CD14 was not detected in CD14^{-/-} pups fostered to CD14^{-/-} mothers at any time point (Figure 3.1H).

Soluble CD14 in the stomach contents and blood of pups from all treatment groups was confirmed to be intact (48 kDa, Figure 3.2). The amount of sCD14 in the

stomach contents and blood of pups visualized by western blot also corresponded to the sCD14 concentrations obtained by ELISA (Figure 3.1).

sCD14 transferred to the blood remains functional

CD14^{-/-} pups fostered to WT mothers had significantly higher IL-6 and TNF- α concentration following LPS-injection ($3,225 \pm 610$ and 15.9 ± 2.7 fold increase, respectively) in comparison to CD14^{-/-} pups fostered to CD14^{-/-} mothers (693 ± 184 and 6.1 ± 1.5 fold increase for IL-6 and TNF- α , respectively, ($P=0.002$, $P=0.005$) Figure 3.3). The levels of sCD14 did not increase post-LPS injection in CD14^{-/-} pups irrespective of the foster mothers' genotype (Figure 3.3A). WT pups fostered to WT mothers showed an increase in circulating sCD14 levels (1.6 ± 0.1 fold increase), IL-6 levels (5333 ± 853 fold increase) and TNF- α levels (40.3 ± 8.2) post-LPS injection (Figure 3.3). WT pups fostered to CD14^{-/-} mothers showed a significantly higher increase in IL-6 concentration post-LPS injection (7764 ± 946 fold increase) in comparison to all other genotype/foster groups ($P<0.031$, Figure 3.3C). Conversely, sCD14 and TNF- α levels post LPS-injection in WT pups fostered to CD14^{-/-} mothers (2.0 ± 0.3 fold increase and 41.7 ± 12.5 fold increase, respectively) were not significantly different from that of WT pups fostered to WT mothers ($P>0.05$, Figure 3.3).

Human cells transport sCD14 in vitro

When added to the apical media of Caco-2 cell Transwell assays, hrCD14 was transcytosed through cell monolayers and into the basal media (Figure 3.4).

Figure 3.1. Quantity of sCD14 in the stomach contents (left panel) and blood (right panel) of mouse pups ingesting milk with or without sCD14. On day six post partum, pups were fostered to different mothers (dotted line). (A) and (B), WT pups fostered to WT mothers; (C) and (D), WT pups fostered to CD14^{-/-} mothers; (E) and (F), CD14^{-/-} pups fostered to WT mothers; (G) and (H), CD14^{-/-} pups fostered to CD14^{-/-} mothers. *, $P < 0.05$ by t -test, $n=3$, bars represent mean $\pm$ standard deviations.

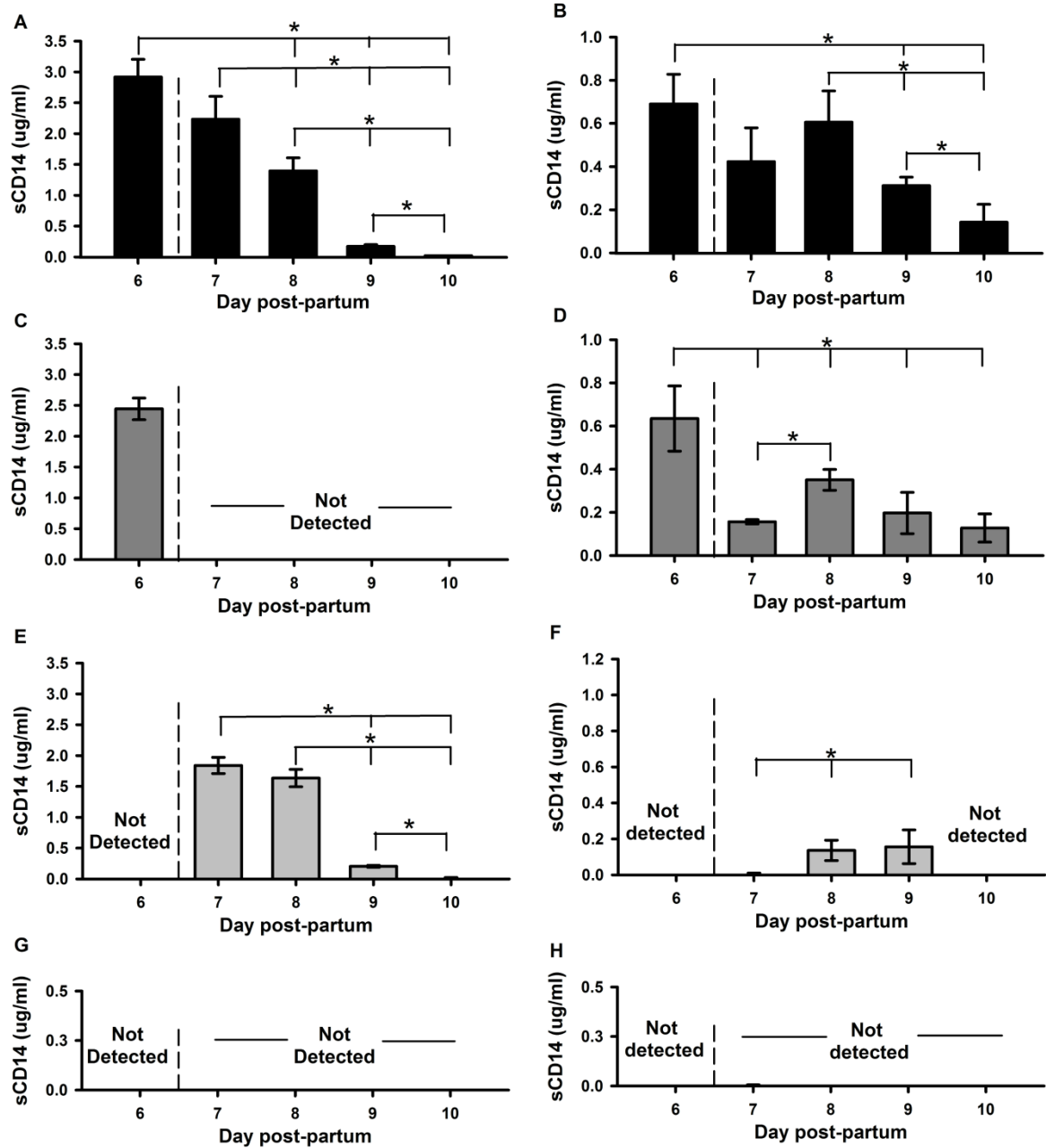


Figure 3.2. Intactness of sCD14 in the stomach contents and blood of mouse pups ingesting milk with or without sCD14. (A) WT pups fostered to WT mothers; (B) WT pups fostered to CD14^{-/-} mothers; (C) CD14^{-/-} pups fostered to WT mothers; (D) CD14^{-/-} pups fostered to CD14^{-/-} mothers. Dotted line indicates when the pups were fostered to different mothers.

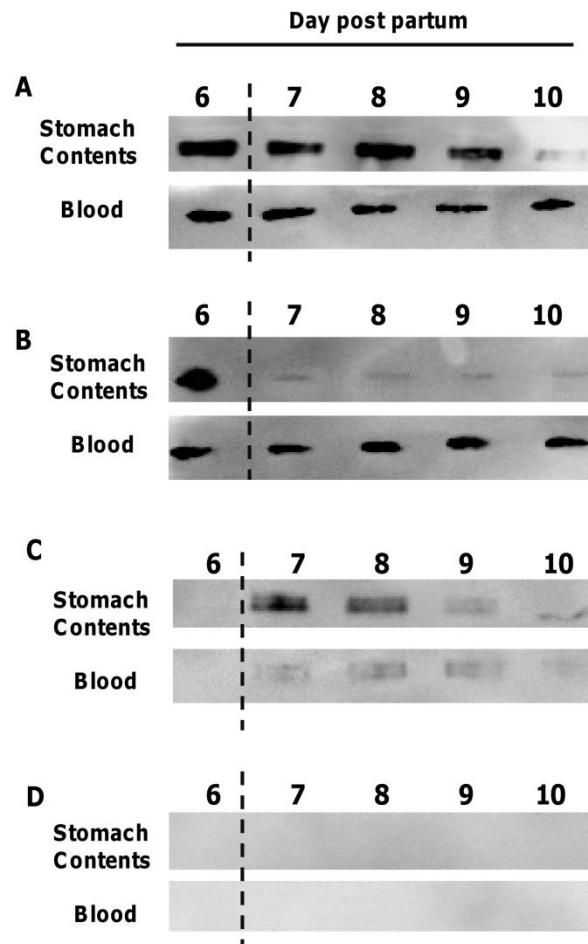
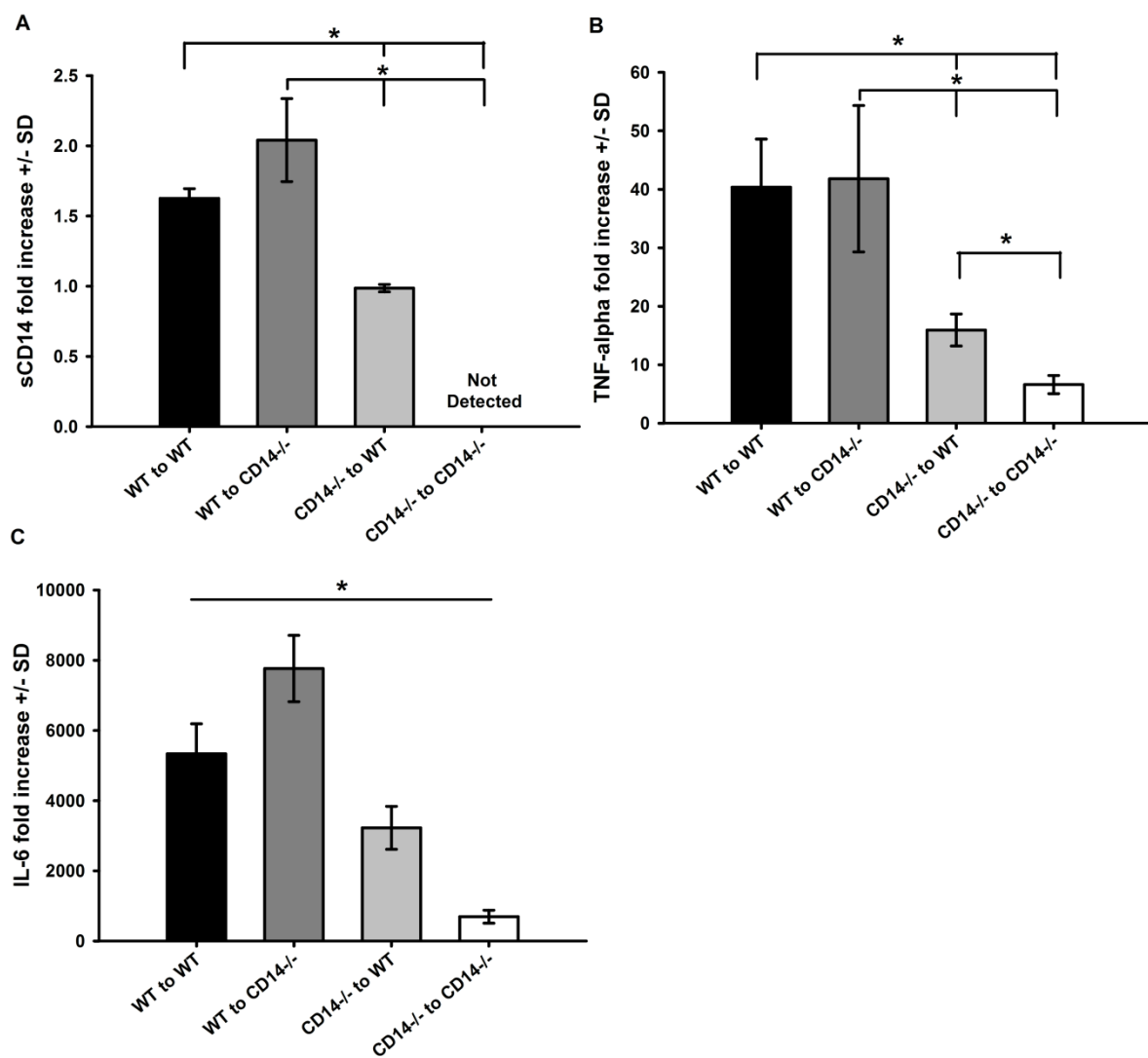


Figure 3.3. Immune response to LPS following sCD14 ingestion in mouse pups, assessed via sCD14, TNF-alpha, and IL-6 levels. Pups were fostered to mothers expressing (WT) or not expressing sCD14 (CD14^{-/-}) in their milk and blood was sampled pre and post LPS injection (200 ng/g body weight) on day two post-foster. Immune response was measured by quantifying sCD14 (**A**), TNF-alpha (**B**) and IL-6 (**C**) by ELISA and is expressed as a fold increase over pups fostered similarly but injected with PBS alone. *, $P < 0.05$ by t -test, $n = 3$.



The amount of hrCD14 transcytosed was dependent on the initial dose in the apical well; higher initial concentrations resulted in significantly more hrCD14 observed in the basal well (Figure 3.4). After two hours of incubation, up to 3.41 ± 0.06 % of the initial hrCD14 added to the apical well was transcytosed to the basal well (Figure 3.4). The integrity of the membrane was monitored by passive diffusion of Lucifer Yellow, which was minimally detected in the basal media (Supplemental Figure S3.2). The intactness of pre- and post- transcytosed hrCD14 was monitored by western blot, which showed hrCD14 from the apical and basal well to be intact (48 kDa, Figure 3.4). To confirm the cellular uptake of hrCD14 by Caco-2 monolayers, hrCD14 was monitored by immunocytochemistry and microscopy. The amount of hrCD14 within the Caco-2 cells was dependent on the initial apical hrCD14 concentration (Figure 3.5), further confirming the ELISA results shown in Figure 3.4. The production of endogenous CD14 (mCD14 and sCD14) by Caco-2 cells was also observed by microscopy (Figure 3.5A).

sCD14 transfer by Caco-2 cells is TLR4 dependent

TLR4 expression was observed in Caco-2 cells but not with another intestinal epithelial cell line, HT29, via RT-PCR (Figure 3.6). Following TLR4-specific siRNA treatment, TLR4 expression was significantly decreased in Caco-2 cells in comparison to control siRNA (scrambled) or non-treated Caco-2 cells (Figure 3.7A). Knockdown of TLR4 expression resulted in a significant decrease in sCD14 transfer across Caco-2 monolayers. Specifically, in control Caco-2 cells, 3.45 ± 0.13 % of hrCD14 was

transferred to the basal media, whereas $0.005 \pm 0.001\%$ of hrCD14 was transferred from the apical to basal media in TLR-siRNA treated cells ($P < 0.001$, Figure 3.7B).

Figure 3.4. Quantity (A) and intactness (B) of sCD14 transferred across human intestinal monolayers in vitro. Caco-2 cells were grown and differentiated as monolayers on Transwell plates. Human recombinant CD14 (hrCD14) was added to apical media and transfer of hrCD14 to basal media was measured by ELISA. Intactness of hrCD14 in the apical (a) and basal (b) media was detected by western blot. *, $P < 0.05$ by t -test, $n = 3$.

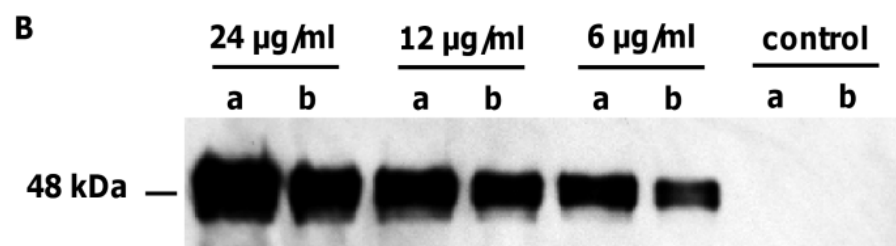
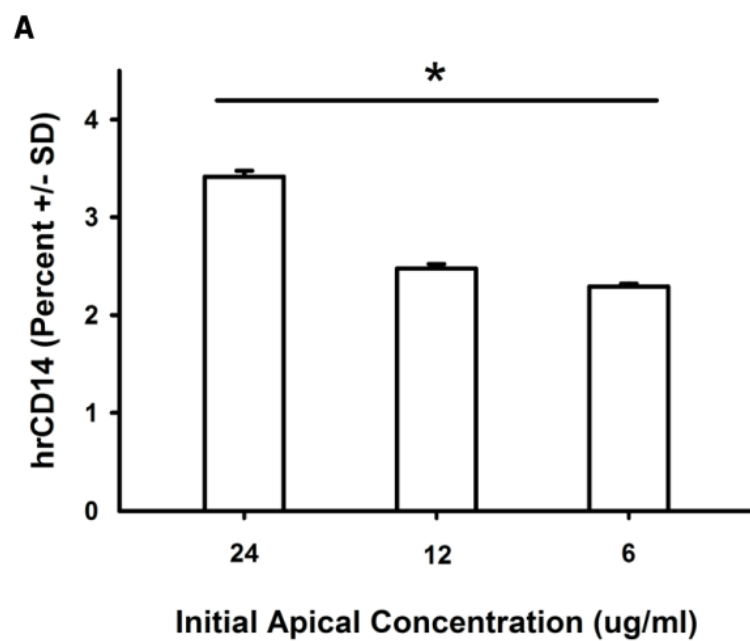
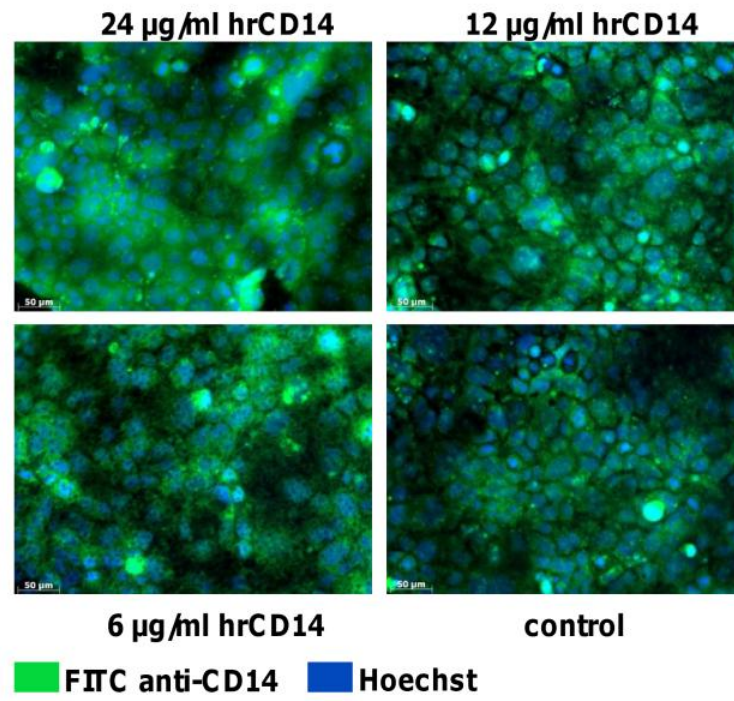


Figure 3.5. Visualization of sCD14 within human intestinal cells treated with sCD14 in vitro. Caco-2 cells were grown and differentiated as monolayers on Transwell plates. Human recombinant CD14 (hrCD14) was added to the apical media and presence of hrCD14 within the monolayers was detected by microscopy (**A**) after incubation with or without hrCD14. Signal intensities for hrCD14 (**B**) were calculated using ImageJ. *, $P < 0.05$ by t -test, $n = 3$.

A



B

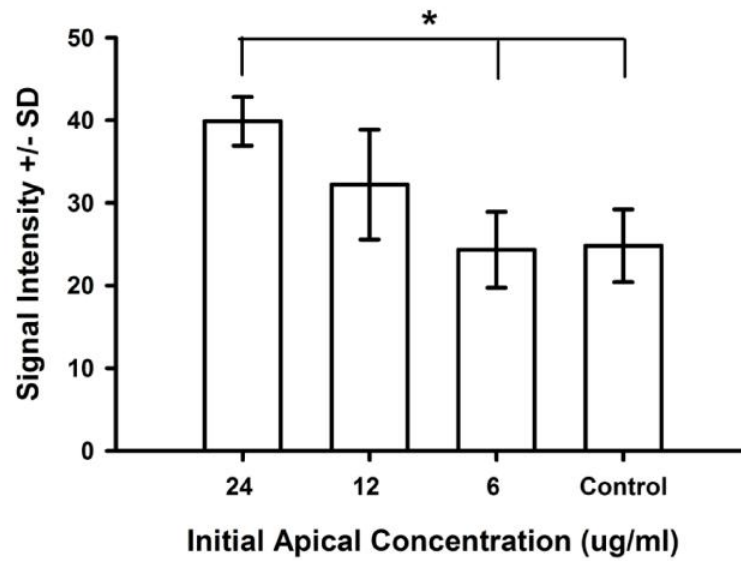


Figure 3.6. TLR4 expression in Caco-2 and HT29 cells. Intestinal epithelial cell lines were grown in monolayers and TLR4 expression was determined by reverse transcription PCR. TLR4 (75 bp) and GAPDH (66 bp) amplicons were electrophoresed following PCR on polyacrylamide gels.

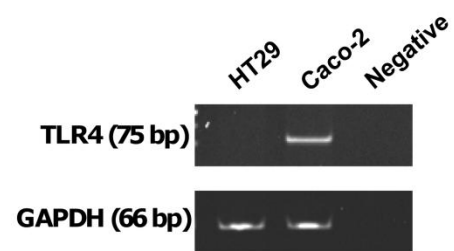
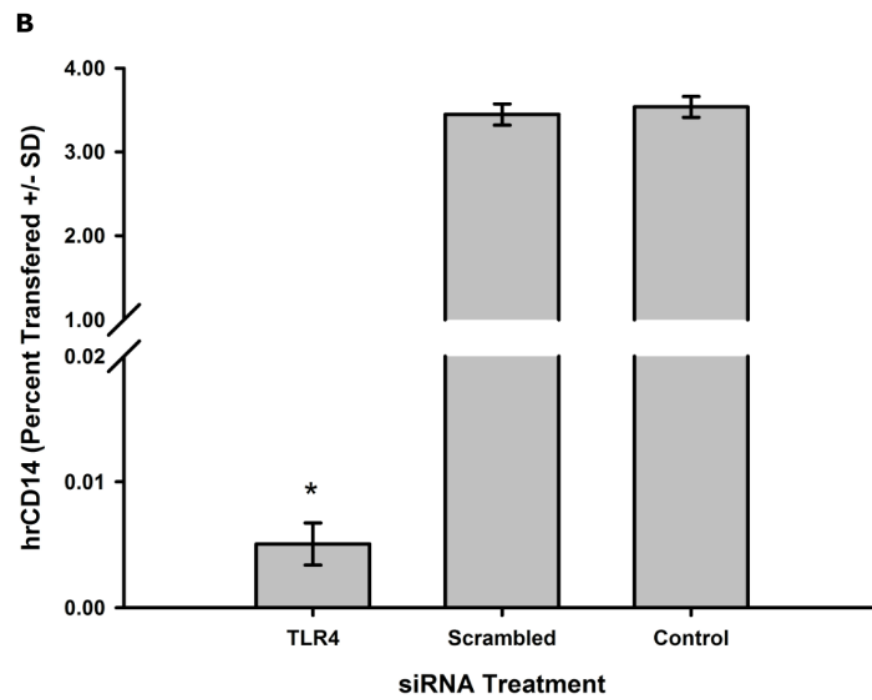
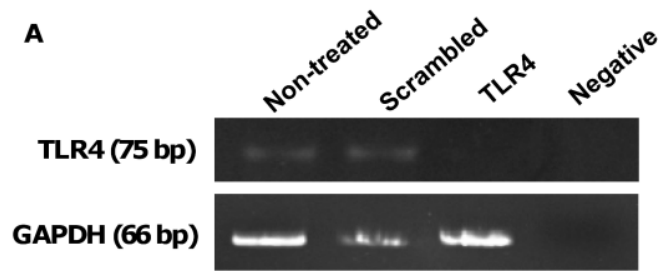


Figure 3.7. Effect of TLR4 knockdown on sCD14 transfer by human intestinal cells in vitro. Caco-2 cells were grown and differentiated as monolayers on Transwell plates, treated with TLR4 specific siRNA, scrambled siRNA or no siRNA, and TLR4 expression was measured by reverse transcription PCR. Following PCR, TLR4 (75 bp) and GAPDH (66 bp) amplicons were electrophoresed following PCR on polyacrylamide gels (**A**). Human recombinant CD14 (hrCD14) was added to apical media and transfer of hrCD14 to basal media was measured by ELISA (**B**). *, $P < 0.05$ by t -test, $n = 4$.



3.5 Discussion

Fate of ingested sCD14

Ingested sCD14 was transferred to the blood of mice as shown by detection of CD14 in the blood of CD14^{-/-} pups fostered to WT mothers (0.16 ± 0.09 $\mu\text{g/ml}$, day eight PP, Figure 3.1F). Ingested sCD14 contributed approximately 26.7% of the circulating sCD14 on day eight PP, as assessed by comparing the concentration of sCD14 in the blood of WT pups fostered to WT mothers to CD14^{-/-} pups fostered to WT mothers (0.60 ± 0.14 $\mu\text{g/ml}$ and 0.16 ± 0.09 $\mu\text{g/ml}$, respectively, Figure 3.1). The contribution of ingested sCD14 to circulating levels of sCD14 pups was significant earlier in lactation, when sCD14 levels in milk were higher as (Figure 3.1A). Specifically, when comparing WT pups fostered to CD14^{-/-} mothers to WT pups fostered to WT mothers, the amount of circulating sCD14 significantly decreased from 0.60 ± 0.14 $\mu\text{g/ml}$ at day eight PP (WT to WT) to 0.35 ± 0.05 $\mu\text{g/ml}$ at day eight PP (WT to CD14^{-/-}) due to the absence of sCD14 from their diet ($P=0.016$, Figure 3.1B and D). Over time, the amount of sCD14 in the mouse milk decreased, but the level of circulating sCD14 in WT pups did not show a similar decrease (Figure 3.1). For example, when comparing sCD14 levels in the blood of WT pups fostered to WT mothers to WT pups fostered to CD14^{-/-} mothers, the amount of circulating sCD14 on days nine and ten PP do not significantly differ between the groups regardless of the absence of sCD14 from the diet of the WT pups fostered to CD14^{-/-} mothers ($P>0.05$, Figure 3.1B and D). The lack of correlation between milk-sCD14 levels and circulating sCD14 levels after eight days of suckling suggests pups begin producing more of their own circulating sCD14 at this time and rely less on contributions from ingested sCD14.

A decrease in sCD14 levels during lactation is also seen in humans, where sCD14 levels are reported as high as 26 µg/ml in colostrum, which decreases to 5µg/ml by 30 days PP (56). Additionally, circulating sCD14 levels in infants do not reach levels similar to those in adults until four months PP (79). Therefore, uptake of ingested sCD14 by infants as a means to compensate for a lack of endogenously produced sCD14 is likely a conserved mechanism in both rodents and humans. This is supported by the observation of human intestinal cells transferring hrCD14 across monolayers in vitro (Figure 3.4). When exposed to hrCD14, Caco-2 cells were able to transfer hrCD14 from the apical to basal well of a Transwell assay in a dose-dependent manner (Figure 3.4).

Caco-2 monolayers have also been used as a model to demonstrate the transcytosis of other milk components including proteins, bacteria and viruses. For example β-lactoglobulin, a major bovine milk allergen, was recently shown to evade digestion similarly to sCD14, and was also transferred across Caco-2 monolayers in vitro (93). Viruses found in human milk, such as human T-cell leukemia virus type 1 (HTLV-1) which resides within infected milk-lymphocytes in HTLV-1 positive mothers, has also been shown to transfer across Caco-2 monolayers. This transfer occurs via transcytosis despite the resistance of enterocytes to HTLV-1 infection (94). Similarly, Caco-2 monolayers and brain microvascular endothelial cell monolayers have been shown to transcytose *Cronobacter*, a type of bacteria linked to neonatal meningitis (95). *Cronobacter* can be found in human milk (96) and can also reside within powdered infant formulas, the latter of which has been reported as a source of neonatal meningitis outbreaks and discussed in a recent review (97).

In vivo uptake of other milk proteins, such as immunoglobulin, has also been extensively studied. Recently, transcytosis of IgG by intestinal cells of rats was imaged by electron tomography, which showed the antibody to be taken up by epithelial cells along the intestine and transported to the blood in an endosome dependent pathway (22, 98). Further analysis revealed the intestinal cells of weaned animals had less multivesicular bodies containing early endosomal markers in comparison to neonatal animals (22). The change in vesicle phenotype throughout suckling suggests the capacity of the neonatal intestine to transport functional proteins from the lumen of the intestine and into the blood diminishes during maturation. Because the present study observed sCD14 transport across a window of five days (days six through ten PP), decreased intestinal epithelial transport during suckling likely did not contribute greatly to any changes in sCD14 transport during the study (Figure 3.1).

Functionality of ingested sCD14

One well studied function of CD14 (both mCD14 and sCD14) is its ability to recognize LPS from Gram-negative bacteria and initiate an innate immune response through TLR4 (85). Because of this, CD14^{-/-} mice are resistant to LPS mediated septic shock, as a minimal TNF- α and IL-6 response is observed following injection of LPS in concentrations normally lethal to WT mice (99). When CD14^{-/-} pups were fostered to WT mothers, their level of circulating sCD14 increased as ingested milk-sCD14 was transferred to the blood (Figure 3.1F). When injected with LPS, CD14^{-/-} pups fostered to WT mothers presented with significantly higher TNF- α and IL-6 responses than CD14^{-/-} pups fostered CD14^{-/-} mothers ($P < 0.05$, Figure 3.3B and C). This suggests that

ingested sCD14 transferred to the blood of pups remains functional in its ability to detect LPS.

CD14^{-/-} pups fostered to WT mothers had a significantly lower immune response (both IL-6 and TNF- α) following LPS injection in comparison to WT pups fostered to WT mothers ($P=0.002$ Figure 3.3B and $P=0.007$ Figure 3.3C). This is likely due to the presence of only sCD14 and no mCD14 in the CD14^{-/-} pups fostered to WT mothers. In WT mice, mCD14 can be found on monocytes, hepatocytes and intestinal epithelial cells where it is responsible for recognizing LPS and initiating the TLR4 mediated immune response (100, 101). Therefore, because mCD14 is found on numerous cells in WT pups, WT pups are more readily equipped to recognize LPS regardless of sCD14 ingestion. When comparing the immune response to LPS of WT pups fostered to WT mothers (WT to WT) to WT pups fostered to CD14^{-/-} mothers (WT to CD14^{-/-}), there was a significantly higher IL-6 response in WT pups fostered to CD14^{-/-} mothers ($P=0.030$, Figure 3.3C). On the second day post-foster, significantly less sCD14 was observed in the blood of WT pups fostered to CD14^{-/-} mothers in comparison to WT pups fostered to WT mothers ($P=0.016$, Figure 3.1B and D). As previously stated, WT pups express mCD14 on various cell types which are then able to detect LPS, regardless of sCD14 ingestion. Although milk sCD14 is capable of recognizing LPS and subsequently passing LPS to TLR4 (29), sCD14 can also sequester LPS away from mCD14 expressing cells and bring LPS to the liver for clearance by TLR4 expressing hepatocytes (30, 88). Therefore, the lower levels of sCD14 in WT pups fostered to CD14^{-/-} mothers may decrease the amount of LPS sent to the liver for clearance, resulting in a higher immune response to LPS exposure.

sCD14 transfer by Caco-2 cells is TLR4 dependent

Caco-2 cells in Transwell assays were able to transfer hrCD14 across the cell monolayer from the apical media to the basal media in a dose dependent manner (Figure 3.4). Previous research has shown TLR4 internalization to be CD14 dependent in macrophages and dendritic cells and thus CD14 internalization was hypothesized to be TLR4 dependent in intestinal cells (91). When TLR4 expression was knocked down using siRNA, the capacity of Caco-2 cells to transport hrCD14 across the monolayer was diminished ($3.45 \pm 0.13\%$ of hrCD14 transferred by control cells versus $0.005 \pm 0.001\%$ of hrCD14 transferred by TLR4-siRNA treated cells, $P < 0.001$, Figure 3.7B). In rats, LPS transport across the intestinal epithelial ex vivo was observed to be inhibited using anti-CD14 or anti-TLR4 antibodies, supporting a role of these two pattern recognition receptors in mucosal LPS transport (102). Similarly, a study using mouse intestinal cells in vitro (mIC_{c12}) showed an increase of CD14 expression following LPS exposure, and LPS was internalized and colocalized with TLR4 within mIC_{c12} cells (103). Recently, a study with Caco-2 cells showed that LPS increased TLR4 and CD14 expression in intestinal monolayers, and a similar trend was observed in vivo along the mouse GI tract (100). In healthy intestine, LPS does not permeate through the intestinal barrier to a great extent (intestinal lumen to blood transfer), likely due to a low presence of TLR4 and CD14 along the GI tract (100, 104–106). Once exposed to LPS, TLR4 and CD14 concentrations increase, resulting in greater intestinal permeability (100). This is likely how milk-derived sCD14 is able to cross the intestinal barrier, through involvement with LPS recognition.

Implications of biologically active sCD14 post-ingestion

Although high concentrations of sCD14 in early-lactation milk in conjunction with sCD14's ability to transfer to the blood post-ingestion suggest sCD14 is an important human milk protein, sCD14's role in infant health remains unknown. Similarly to following LPS exposure as shown by Guo *et al*, TLR4 is also upregulated along the GI tract of premature infants (100, 104–106). Once activated, TLR4 is able to induce a proinflammatory response, and this activation in the intestine has been implicated as a cause of NEC (107, 108 for review). Incidence of NEC is higher in premature infants and lower birth weight infants have the highest NEC-caused death rates of up to 30% (109). During NEC, activation of TLR4 leads to autophagy and subsequent impaired migration of intestinal epithelial cells from the crypts of the intestine resulting in decreased barrier integrity (110). Additionally, TLR4 activation within the endothelium has recently been shown to reduce mesenteric perfusion resulting in intestinal ischemia and increased NEC severity (111). Activation of TLR4 may be dependent on CD14, since blocking CD14 through inactivating antibodies decreased the severity of induced NEC by decreasing the TLR4 immune response as measured by IL-6 and TNF- α expression (112). WT pups not ingesting sCD14 were seen to have an increase in IL-6 response to LPS compared to those ingesting sCD14 (Figure 3.3), which confirms previous studies that show increased sCD14 levels prior to or at the time of LPS exposure dampens the LPS-induced immune response (113, 114). Therefore, high levels of sCD14, such as those in colostrum and early lactation milk (26 $\mu\text{g/mL}$) may be beneficial to the infant prior to exposure of LPS.

Conclusions

Ingestion of sCD14 within milk results in the transfer of intact sCD14 from the GI tract to the circulatory system of mouse pups. Transfer of sCD14 across the intestinal barrier is TLR4 dependent in vitro and newly circulating sCD14 derived from ingested sources remains functional in its ability to detect LPS in vivo. High quantities of sCD14 in early lactation milk implicate sCD14 as a factor influencing infant health, especially given the role of the LPS recognition pathway, which includes CD14 and TLR4, in the development of NEC.

Chapter 4

Human milk metagenome: a functional capacity analysis

Tonya L Ward, Sergey Hosid, Ilya Ioshikhes, Illimar Altosaar

Published in BMC Microbiology
Volume 13, page 116, 2013
doi:10.1186/1471-2180-13-116

Contribution of authors

Tonya Ward: Contributed to study design, conducted experiments, analyzed the data, interpreted results, wrote and edited the manuscript.

Sergey Hosid: Performed the data analysis for Figure 4.1, Tables 4.1, 4.2, 4.3 and Table S4.1, Table S4.2, Table S4.4.

Ilya Ioshikhes: Contributed to study design and edited the manuscript.

Illimar Altosaar: Conceived the study design and edited the manuscript.

4.1 Abstract

Background

Human milk contains a diverse population of bacteria that likely influences colonization of the infant GI tract. Recent studies, however, have been limited to characterization of this microbial community by 16S rRNA analysis. In the present study, a metagenomic approach using Illumina sequencing of a pooled milk sample (ten donors) was employed to determine the genera of bacteria and the types of bacterial ORFs in human milk that may influence bacterial establishment and stability in this primal food matrix. The human milk metagenome was also compared to that of BF and formula fed (FF) infants' feces (n = 5, each) and mothers' feces (n = 3) at the phylum level and at a functional level using ORF abundance. Additionally, immune-modulatory bacterial-DNA motifs were also searched for within human milk.

Results

The bacterial community in human milk contained over 360 prokaryotic genera, with sequences aligning predominantly to the phyla of Proteobacteria (65%) and Firmicutes (34%), and the genera of *Pseudomonas* (61.1%), *Staphylococcus* (33.4%) and *Streptococcus* (0.5%). From assembled human milk-derived contigs, 30,128 ORFs were annotated and assigned to functional categories. When compared to the metagenome of infants' and mothers' feces, the human milk metagenome was less diverse at the phylum level, and contained more ORFs associated with nitrogen metabolism, membrane transport and stress response ($P < 0.05$). The human milk

metagenome also contained a similar occurrence of immune-modulatory DNA motifs to that of infants' and mothers' fecal metagenomes.

Conclusions

The results presented here further expand the complexity of the human milk metagenome and enforce the benefits of human milk ingestion on the microbial colonization of the infant gut and immunity. Discovery of immune-modulatory motifs in the metagenome of human milk indicates more exhaustive analyses of the functionality of the human milk metagenome are warranted.

4.2 Introduction

The benefits of human milk compared to the use of commercial infant formulas are largely realized because of its bioactive components, including prebiotics, immune proteins and the microbiome of human milk itself. Breastfeeding is associated with a decreased incidence of GI tract infections (115, 116), which is corroborated by several studies that have correlated breastfeeding with a lower incidence of NEC in humans and animal models (117–119). Breastfeeding is also associated with an altered fecal microbiome; two studies showed at two weeks of age over 90% of the total fecal bacteria of a BF infant is *Bifidobacteria*, whereas in most formula-fed (FF) infants *Bifidobacteria* is non-detectable (120, 121). Because the community of gut-colonizing bacteria prevents adhesion and colonization of pathogenic bacteria whilst stimulating mucosal cell proliferation and enhancing immune development, the types of predominant bacteria in the fecal microbiome of the developing infant can affect the health outcomes of the individual, as has been discussed in a recent review article (122). Human milk, the infant's first food, is a primary source of ingested microbiota. Therefore, it is paramount to fully understand the human milk microbiome and how it might influence colonization of the infant GI tract.

Ingestion of viable bacteria in human milk may lead to effective colonization of the infant GI tract, but the presence of bacterial DNA alone may also hold responsibility for proper infant immune development. For example, unmethylated CpG dinucleotides within bacterial DNA are known as potent immune stimulators, acting through TLR9 (123). Conversely, immune suppressive motifs including poly-guanosine or guanosine cytosine-rich sequences, such as those on the telomere region

of mammalian DNA, can block immune activation induced by CpGs (124). Recently, immune suppressive motifs (TTAGGG and TCAAGCTTGA) that are able to counter the effects of CpGs have been discovered in *Lactobacillus* (125). If immune-modulatory motifs occur in human milk derived DNA, they could contribute to proper immune development by decreasing exaggerated inflammatory responses to colonizing bacteria, which are seen in infants with NEC (126).

Human milk bacteria have previously been analyzed by culture-dependent and -independent mechanisms, confirming the presence of a magnitude of bacterial phylotypes (45–48, 127–130). In one study, *Staphylococcus* and *Streptococcus* dominated the milk microbiome of most mothers, whereas commercially well known bovine milk-associated genera, *Lactobacillus* and *Bifidobacterium*, contributed as minor milk microbiota members (2–3% of genera; 47). Another study showed that the human milk microbiome changes over time, and may be dependent on the mother's weight and the baby's mode of delivery (48). Most recent methods for determining the milk microbiome have included amplification of 16S ribosomal RNA genes (rRNA) followed by pyrosequencing (47, 48). Although this technique is widely accepted as a means to determine microbial diversity, it does present limitations such as a lack of information on the functional capacity of the microbes within the milk matrix and also prevents data accumulation on the types of DNA motifs to which an infant is exposed.

In this study a metagenomic analysis of the bacteria in human milk using Illumina sequencing and the MG-RAST pipeline was performed (131). The aims were to determine the genera of bacteria in human milk, search for immune-modulatory DNA motifs, and determine the types of bacterial ORFs in human milk that may

influence bacterial presence and stability in this complex yet foundational food matrix.

4.3 Methods

Donors and sample collection

Breastfeeding women ($n = 10$) were recruited from the Children's Hospital of Eastern Ontario (CHEO, Ottawa, Canada) in accordance with the Research Ethics Board of CHEO and the University of Ottawa Research Ethics Board (2007303-01H). Informed consent was given by all participants, all donors were healthy, and milk was donated between 9 and 30 days PP. Milk samples were collected by either manual or electric breast pump expression into a sterile milk collection bag (Medela AG, Baar, Switzerland). To better represent a milk sample that would be received by the infant, breasts were not sterilized prior to collection. Samples were immediately frozen and then stored at -70°C .

DNA isolation

Milk samples (1 mL) were centrifuged at $5,000 \times g$ for 10 minutes to pellet eukaryotic cells. Prokaryotic cells were pelleted from milk serum by centrifugation at $13,000 \times g$ for 15 minutes. Pellets were resuspended in 2 mL PBS with 1% Triton X-100 and incubated for 2 hours at 37°C to lyse any remaining eukaryotic cells. Bacteria were pelleted by centrifugation at $13,000 \times g$ for 15 minutes and pellets were resuspended in 500 μL TE with 30 μL of 10% SDS and 5 μg proteinase K. Samples were incubated for 2 hours at 37°C , and DNA was isolated using phenol/chloroform as previously described (132). DNA pellets were resuspended in 50 μL Tris-EDTA buffer and pooled. A total of ~ 4 μg of double stranded DNA was isolated as quantified with Quant-iT PicoGreen (Invitrogen, Burlington, ON, Canada) using a Typhoon Trio Imager and

Image Quant TL software (GE Healthcare). DNA integrity was also determined by agarose gel electrophoresis prior to sequencing.

DNA sequencing, filtering and contig assembly

The pooled DNA sample was sequenced seven independent times by StemCore Laboratories (Ottawa, Ontario, Canada). DNA was prepared according to the DNA sample preparation protocol 1003806 Rev. B for Illumina sequencing (Illumina Inc., San Diego, CA, USA). Sequencing was performed using an Illumina GAIIx Genome Analyzer and Illumina CASAVA analysis pipeline (v 1.7.0). Sequences were aligned to the human genome (hg19/NCBI37) with a stringency of 2 bp mismatching using ELAND (Illumina Inc). Prokaryotic genomes (1,731 genomes) were imported from NCBI. Sequences were aligned to the genomes using BLAT (Kent Informatics, Inc.) and sorted via best hit analysis to genera according to “List of Prokaryotic Names with Standing in Nomenclature” (<http://www.bacterio.cict.fr/> website, accessed February 2012). Unidentified sequences were further filtered by using BLAT against the human genome with a stringency of ≤ 10 mismatches or gaps. Both prokaryotic and remaining unknown sequences were assembled into contigs using Ray v1.7 (133).

Contigs, ORF prediction and characterization

Assembled contigs were uploaded to the MG-RAST pipeline (131). Organism abundance was analyzed using a lowest common ancestor approach with a maximum e-value of 1×10^{-5} , a minimum identity of 60%, and a minimum alignment length of 15 measured in amino acids for protein and base pairs for RNA databases. A functional

abundance analysis of ORFs was performed using "Hierarchical Classification" by comparing to subsystems with a maximum e-value of 1×10^{-5} , a minimum identity of 60%, and a minimum alignment length of 15 measured in amino acids for protein and base pairs for RNA databases. Previously reported and publicly available metagenomes of feces from five unrelated BF-infants, five FF-infants (metagenome IDs: USinfTW4.1, 6.1, 10.1, 11.1, 12.1, 13.1, 15.1, 19.1, 20.1, and 21.11) and three unrelated mothers (metagenome IDs: USchp [1,3,4,18,33]mom) were compared at the phylum level to the human milk metagenome within MG-RAST using the same lowest common ancestor analysis described above (131). The mean percent alignments of the individuals were used in Figure 4.4 and Supplemental Figures S4.4 and S4.5. The normalized mean percent of ORFs in each functional category was used in Figures 4.5 and 4.6. Metagenome comparisons were statistically compared by Student's *t*-tests ($P < 0.05$) using SigmaPlot (Systat Software, Inc.).

Immune-modulatory motif identification

An identity of 100% was used to search for immune-modulatory motifs by alignment with assembled contigs from the human milk metagenome (56,712 contigs) or the fecal metagenomes described above (834,774, 64,662 and 553,391 contigs from BF-infants', FF-infants' and mothers' feces, respectively). The human genome (2,865,822,365 bp) was used as a comparative reference. Z-score was calculated using the formula $Z = (O-E)/\text{Stdev}$, where O was the observed number of hits and E was the expected number of hits using the formula $E = (L\text{cont})(N_h/L_h)$ where Lcont was length of sequences or assembled contigs, N_h was number of sites found in the human

genome (or compiled bacterial genomes); Stdev was the standard deviation of occurrence of each motif in 22 + X + Y human chromosomes.

Availability of supporting data

The data set supporting the results of this article is available in the MG-RAST repository, under the project name Human_milk_microbiome, <http://metagenomics.anl.gov/linkin.cgi?project=2959>

4.4 Results

Phyla and genera within human milk

Metagenome sequencing of a pooled human milk sample resulted in 261,532,204 sequenced reads of 51 bp, which were binned into those aligning to the human genome (186,010,988, 72.01 $\pm$ 3.06%), known prokaryotic genomes (1,331,996, 0.53 $\pm$ 0.16%) or those not aligning to either category (74,189,220, 27.46 $\pm$ 3.72%, Supplemental Table S4.1). Using a best hit analysis of the 1,331,996 51-bp sequences, 75% aligned to *Staphylococcus*, 15% to *Pseudomonas*, 2% to *Edwardsiella*, and 1% to *Pantoea*, *Treponema*, *Streptococcus* and *Campylobacter*, respectively (Figure 4.1). The remaining 3% of the known prokaryotic sequences mapped to 361 bacterial genera, demonstrating the diversity of the human milk metagenome while confirming the presence of key genera like *Akkermansia* (134; Supplemental Table S4.2).

Sequences not aligning to prokaryotic or human genomes with a ≤ 2 bp mismatch were re-aligned to the human genome with decreased stringency (≤ 10 bp mismatch), leaving 32,991,450 sequences for contig assembly (Table 4.1). Using Ray v1.7 (133), 56,712 contigs were assembled and submitted to the MG-RAST pipeline (131). Post quality control, 53,785 sequences (94.8%), with a mean length of 160 $\pm$ 55 bp, were used for further analysis (Table 4.1). When the contigs were analyzed using a best hit approach through MG-RAST, they aligned predominantly to the phyla of Proteobacteria (65.1%) and Firmicutes (34.6%, Figure 4.2). The contigs aligned to 194 known genomes at the genus level, predominantly *Pseudomonas* (61.1%), *Staphylococcus* (33.4%) and *Streptococcus* (0.5%), with the highest level of diversity at the genus level within the Proteobacteria phylum (125 different genera, Figure 4.2).

Figure 4.1 Best hit analysis of 51 bp DNA sequences from human milk. DNA from human milk was sequenced using Illumina sequencing followed by alignment to known prokaryotic genomes. Sequences (75,521,216) were BLATed against 1,731 known prokaryotic genomes imported from NCBI (min 95% identity), with 1,331,996 sequences aligning to 370 prokaryotic genera. Other refers to genera each representing <0.1% of all sequences.

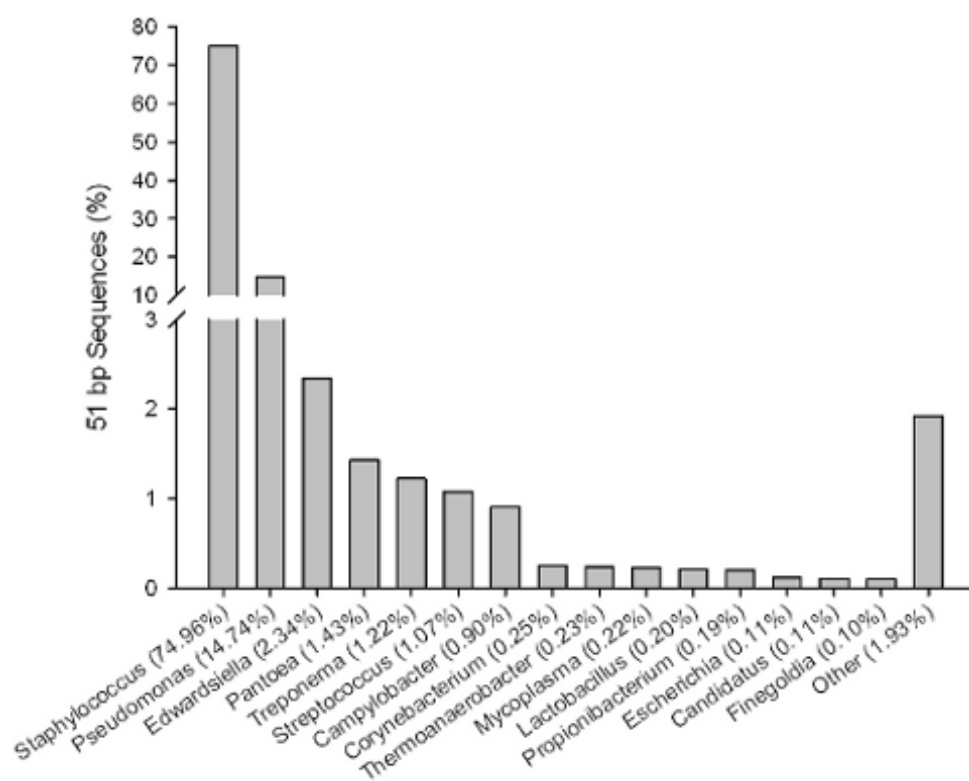


Table 4.1. Contig assembly and open reading frame (ORF) prediction of Illumina reads (51 bp) from human milk.

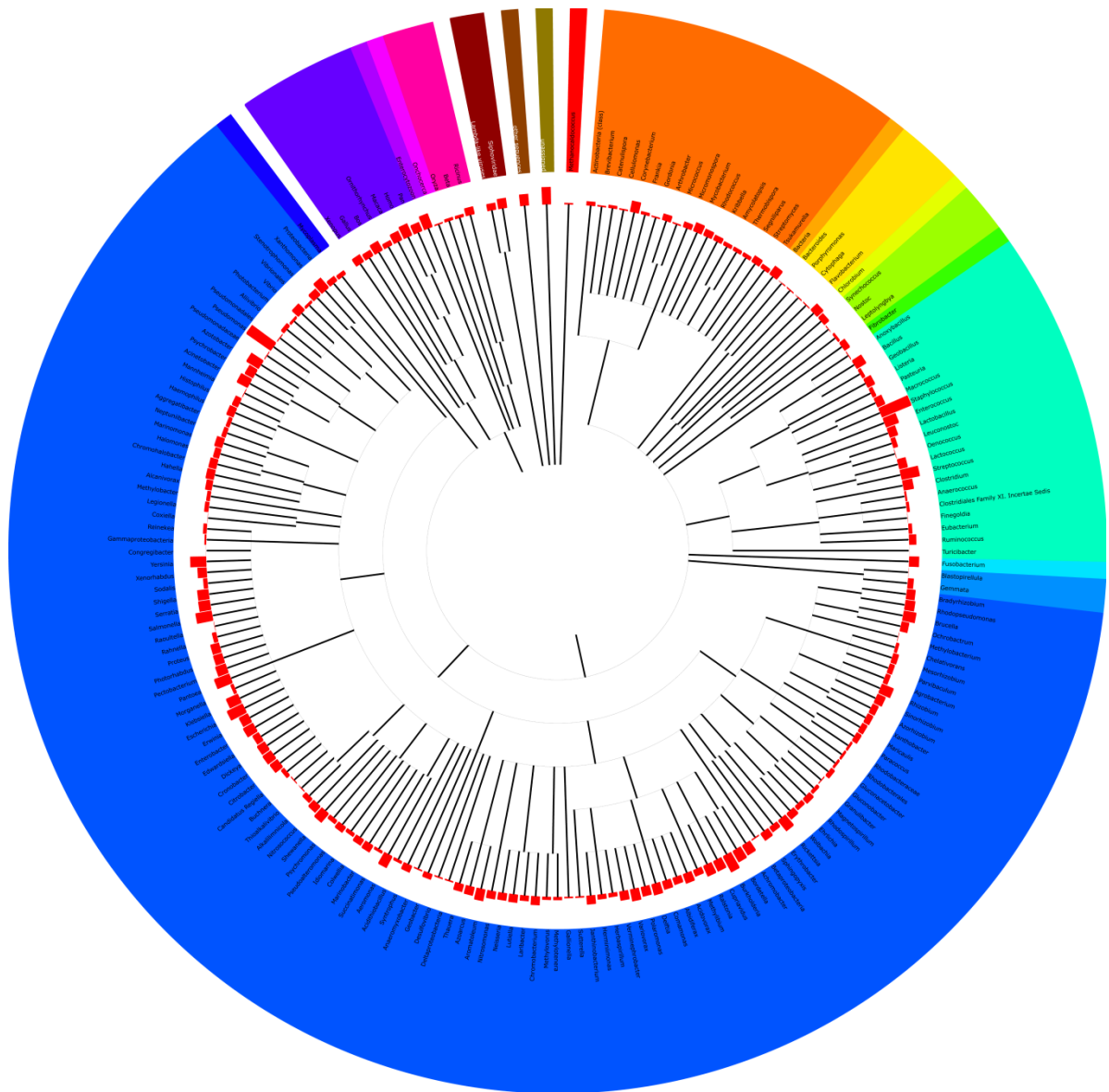
Sequenced reads (51 bp)	261,532,204
Matching human	186,010,988
Matching prokaryotic	1,331,996
Used in contig assembly ¹	32,991,450
Contigs	56,712
Post quality control	53,785
Average length (bp)	160 ± 55
Total length (bp)	8,630,997
Predicted ORFs	41,352
Annotated	33,793
rRNAs	103
Functional category	30,128
Unrecognized	7,559

¹ all sequences not matching the human genome (≤10 bp mismatch)

Figure 4.2. Best hit analysis of open reading frames within human milk.

Assembled contigs (56,712) were submitted to MG-RAST for analysis. Contigs aligned to 194 known genomes at the genus level (maximum e-value of 1×10^{-5} , minimum identity of 60%, and minimum alignment length of 45 bp). Color denotes phylum and red bars indicate the number of positive alignments.

Figures in this chapter can be seen in expanded view mode at [doi:10.1186/1471-2180-13-116](https://doi.org/10.1186/1471-2180-13-116).



These results are similar to the best hit analysis performed with the non-assembled sequences in that the majority of sequences are from *Staphylococcus* and *Pseudomonas*, but differ in their proportion (Figure 4.1). Contigs matching viral genomes were observed (<0.04%), including phages derived from *Pseudomonas* and *Staphylococcus* (Figure 4.2). Contigs also aligned to the genomes of humans, gorillas, chimps and orangutans, likely due to the 60% identity criteria used (Figure 4.2). The observation of some of the genera, including *Staphylococcus*, *Pseudomonas* and *Pantoea*, was further validated through the presence of their rRNA ORFs (Supplemental Table S4.3).

Open reading frames within human milk

A total of 41,352 ORFs were predicted using MG-RAST, of which 82% were annotated (33,793 ORFs), and 18% were unrecognized (7,559 sequences, Table 4.1). A total of 30,128 ORFs corresponded to a functional category (Figure 4.3). For example, many ORFs encoded proteins for basic cellular function, including those for respiration (4.2%), cell signaling (4.8%), RNA (7.0%), DNA (2.6%), and amino acid metabolism (5.3%, Figure 4.3). ORFs encoding proteins for carbohydrate metabolism (5.7% of all ORFs) included those for lactose metabolism (oligosaccharide, 6.7%), but none for human milk oligosaccharide (HMO) metabolism (Figure 4.3), likely due to the lack of sequences aligning to the genome of *Bifidobacteria* (Figure 4.2). Virulence-related ORFs (4.5% of all ORFs) included those for antibiotic resistance (60.2%), adhesion (17%), bacteriocins (2.7%), as well as others (Figure 4.3). Stress-related ORFs (4.0%

of all ORFs) included those for oxidative stress (40.3%), osmotic stress (20.2%), heat and cold shock (12.0% and 4.0%, respectively) and many others (Figure 4.3).

Human milk metagenome compared to mothers' and infants' feces

The metagenome of human milk was compared to that of feces from 10 unrelated infants (five BF and five FF) and three unrelated mothers (Figure 4.4). Using a best hit analysis at the phylum level, contigs from human milk were dissimilar from contigs from feces in regards to the lack of diversity within the human milk metagenome, as over 99% of the contigs were from just two phyla, Proteobacteria and Firmicutes (65.1% and 34.6%, respectively, Figure 4.4). BF-infants' feces had a high proportion of Actinobacteria (70.4%), followed by FF-infants' feces (27.3%), mothers' feces (12.6%), and human milk (0.15%). The proportion of Proteobacteria in the human milk metagenome (65.1%) was most similar to that of BF-infants' feces (10.8%), but was significantly different from FF-infants' feces and mothers' feces (7.5% and 4.3%, respectively, $P < 0.05$, Figure 4.2 and Supplemental Figure S4.1). The metagenomes of FF-infants' feces and mothers' feces were most similar in regards to their high proportion of Bacteroidetes (17.6% and 20.6%, respectively). Conversely, when using a lowest common ancestor approach at the phylum level in comparison to the best hit analysis, human milk appeared more similar to the fecal metagenomes in terms of an increase in diversity (Supplemental Figure S4.2), but was still dominated by Proteobacteria (38.5%). Also, using the lowest common ancestor analysis increased the proportion of contigs aligning to Actinobacteria in human milk (0.15% to 11.58%), as well as in mothers' feces (12.6% to 30.6%).

Figure 4.3. Functional categorization of open reading frames within human milk. The percent of ORFs assigned to each functional category is shown. Using the “Hierarchical Classification” tool within MG-RAST, 41,352 ORFs were submitted, 33,793 were annotated and assigned to a functional category (maximum e-value of 1×10^{-5} , minimum identity of 60%, and minimum alignment length of 15 amino acids). Three categories of genes (stress, virulence, carbohydrates) are expanded on the right to demonstrate the diverse capabilities of milk-derived DNA sequences.

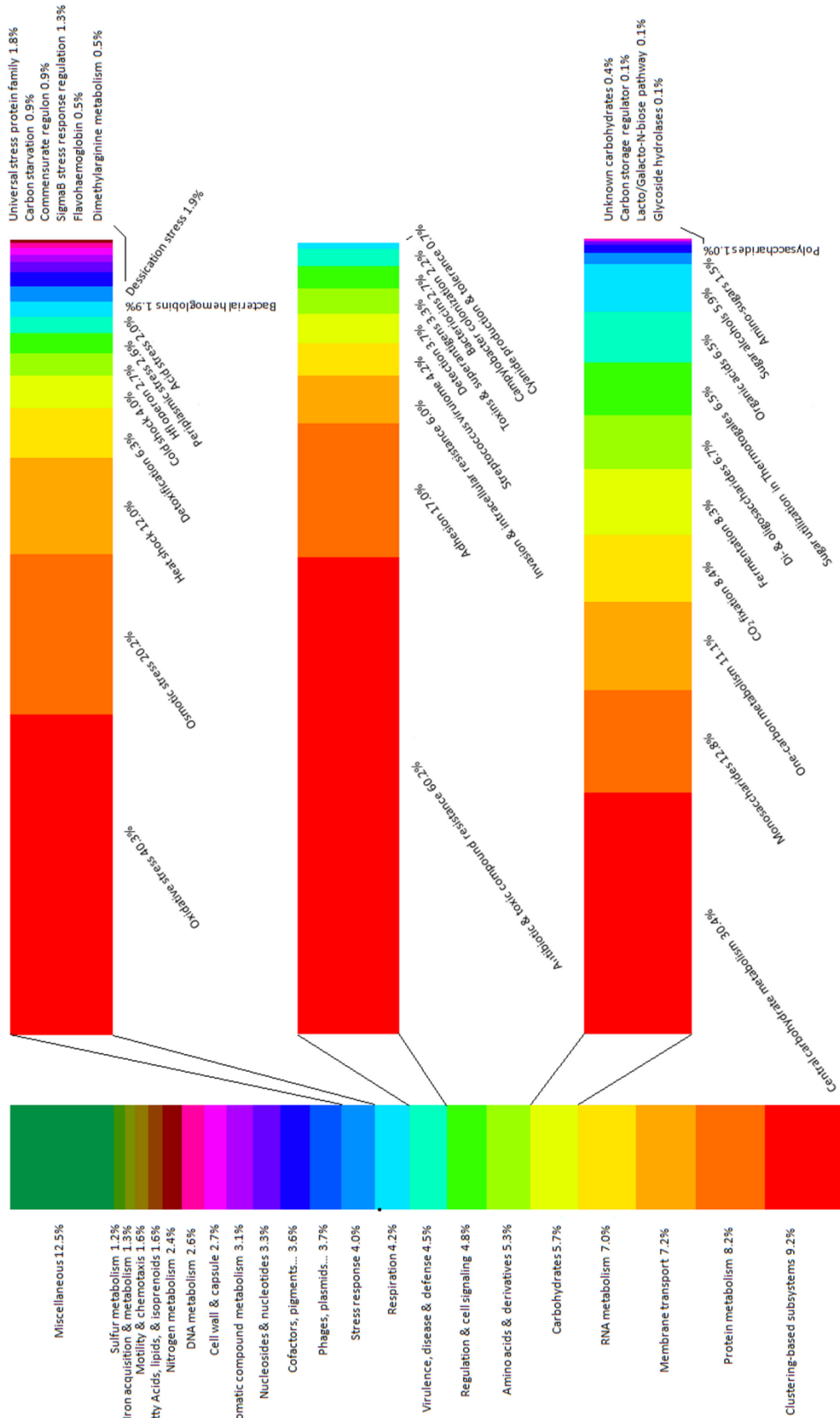
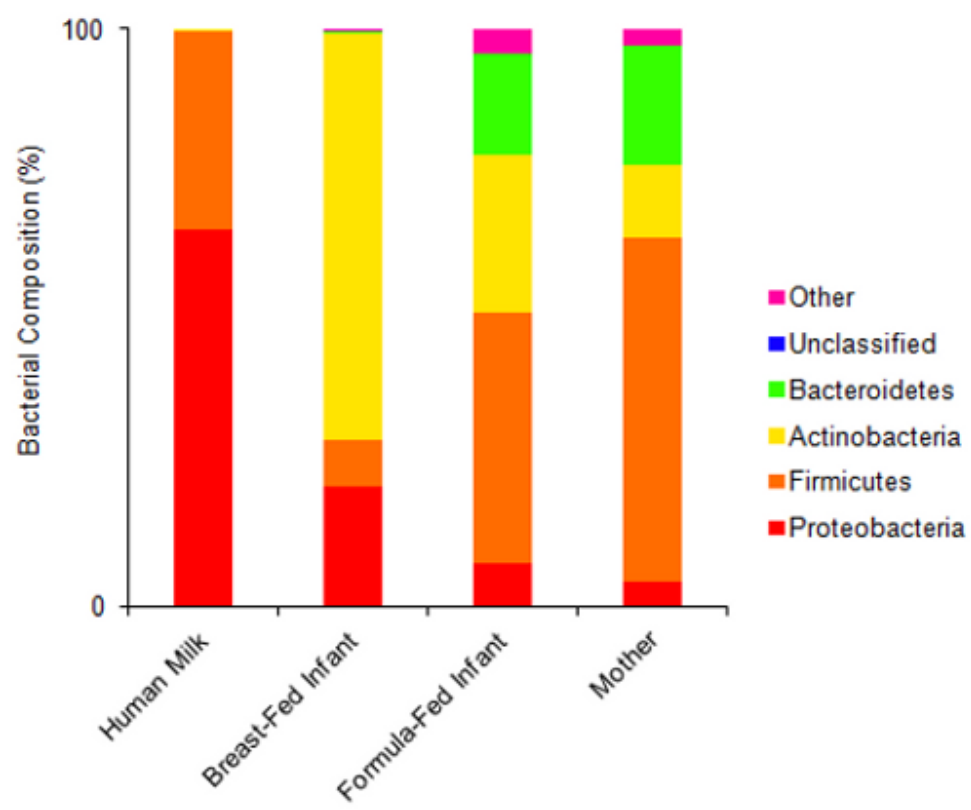


Figure 4.4. Best hit comparison of bacterial phyla in human milk, infants' feces and mothers' feces. The percent of sequences assigned to each phylum according to MG-RAST (maximum e-value of 1×10^{-5} , minimum identity of 60%, and minimum alignment length of 45 bp) is shown. Breast-fed and formula-fed infant feces values are an average of five individuals, and mothers' feces values are an average of three individuals. All subjects were unrelated. Other contains phyla each representing <1% of the contigs.



The metagenomes of human milk and feces were also compared at the functional level (Figure 4.5). The functional ORF profile of the human milk metagenome is similar to that of each fecal metagenome, but two fecal profiles were even more similar, for example BF- versus FF-infants' feces, as seen using pair-wise comparison plots (Figure 4.6). The human milk metagenome is most dissimilar from that of FF-infants' feces as 17 out of the 26 functional categories contain a significantly different proportion of the ORFs (Figure 4.6). The three fecal metagenomes had a significantly higher proportion of ORFs encoding genes for dormancy and sporulation (2.3%, 2.3% and 2.7%, for BF-infants', FF-infants' and mothers' feces, respectively) than did the human milk metagenome (no associated ORFs, Figures 4.5 and 4.6). Both BF- and FF-infants' fecal metagenomes had significantly higher proportions of cell division (3.5% each, respectively) and phosphorus metabolism related ORFs (3.1% and 3.0%, respectively) than did the human milk metagenome (2.3% and 2.1%, Figures 4.5 and 4.6). The human milk metagenome, in comparison to BF- and FF-infants' feces, did, however, have significantly higher proportions of membrane transport (5.0% compared to 4.0% and 4.0%), nitrogen (3.5% compared to 3.1% and 3.0%) and RNA metabolism (4.9% compared to 4.1% and 4.3%), cell regulation (4.4% compared to 3.5% and 3.3%), respiration (4.3% compared to 3.4% and 3.4%), stress response (4.2% compared to 3.7% and 3.5%) and virulence-related ORFs (4.4% compared to 3.7% and 3.7%, Figures 4.5 and 4.6).

Figure 4.5. Functional category comparison of open reading frames within human milk versus infants' and mothers' feces. The percent of ORFs assigned to each functional category of genes is shown. Using the "Hierarchical Classification" tool within MG-RAST, ORFs within each metagenome were assigned to a functional category (maximum e-value of 1×10^{-5} , minimum identity of 60%, and minimum alignment length of 15 amino acids). Asterisk denotes that the proportion of ORFs within the category is significantly different from that in human milk (Student's *t*-test, $P < 0.05$). Breast-fed and formula-fed infant feces values are an average of five individuals, and mothers' feces values are an average of three individuals. All subjects were unrelated.

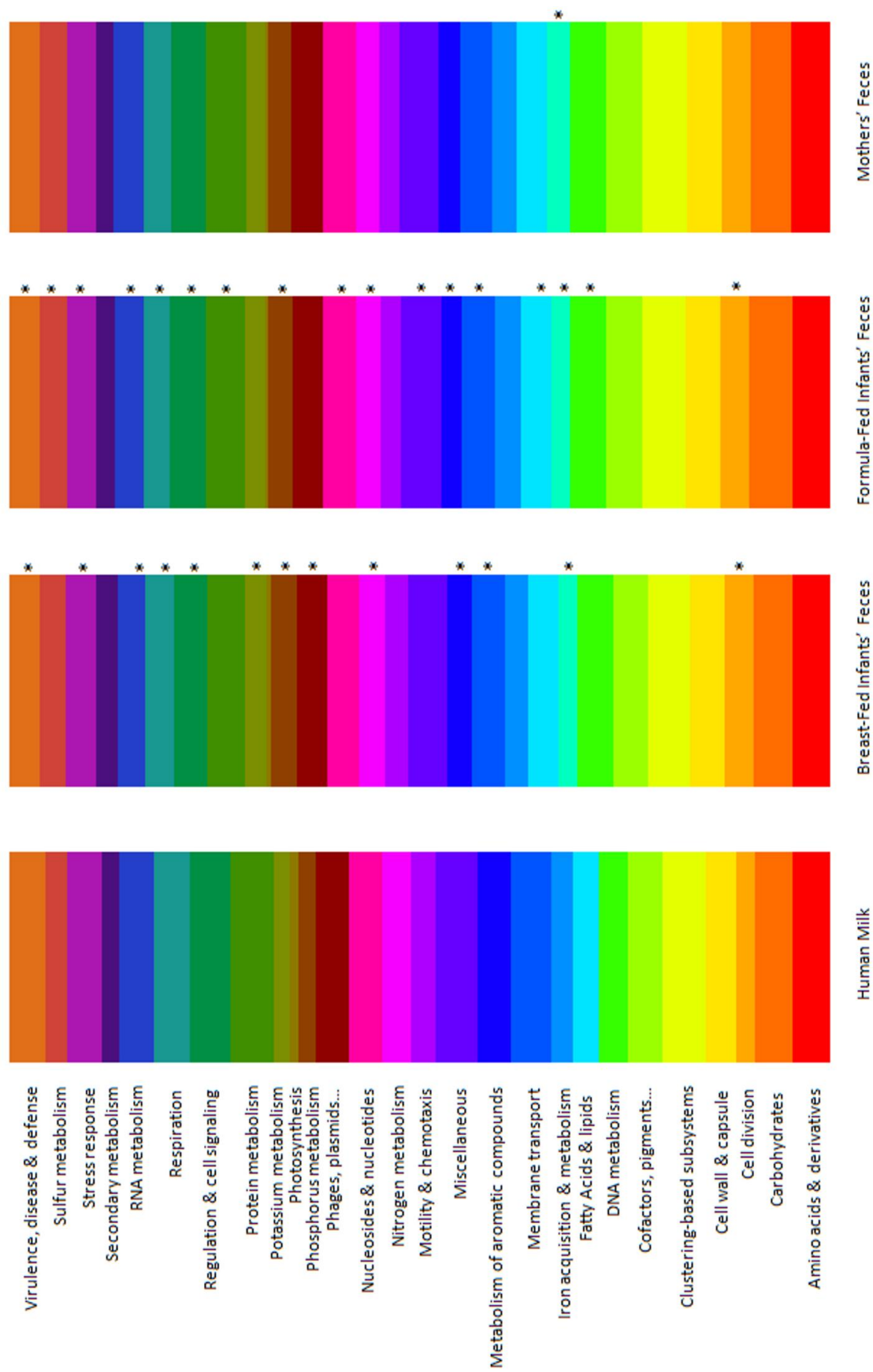
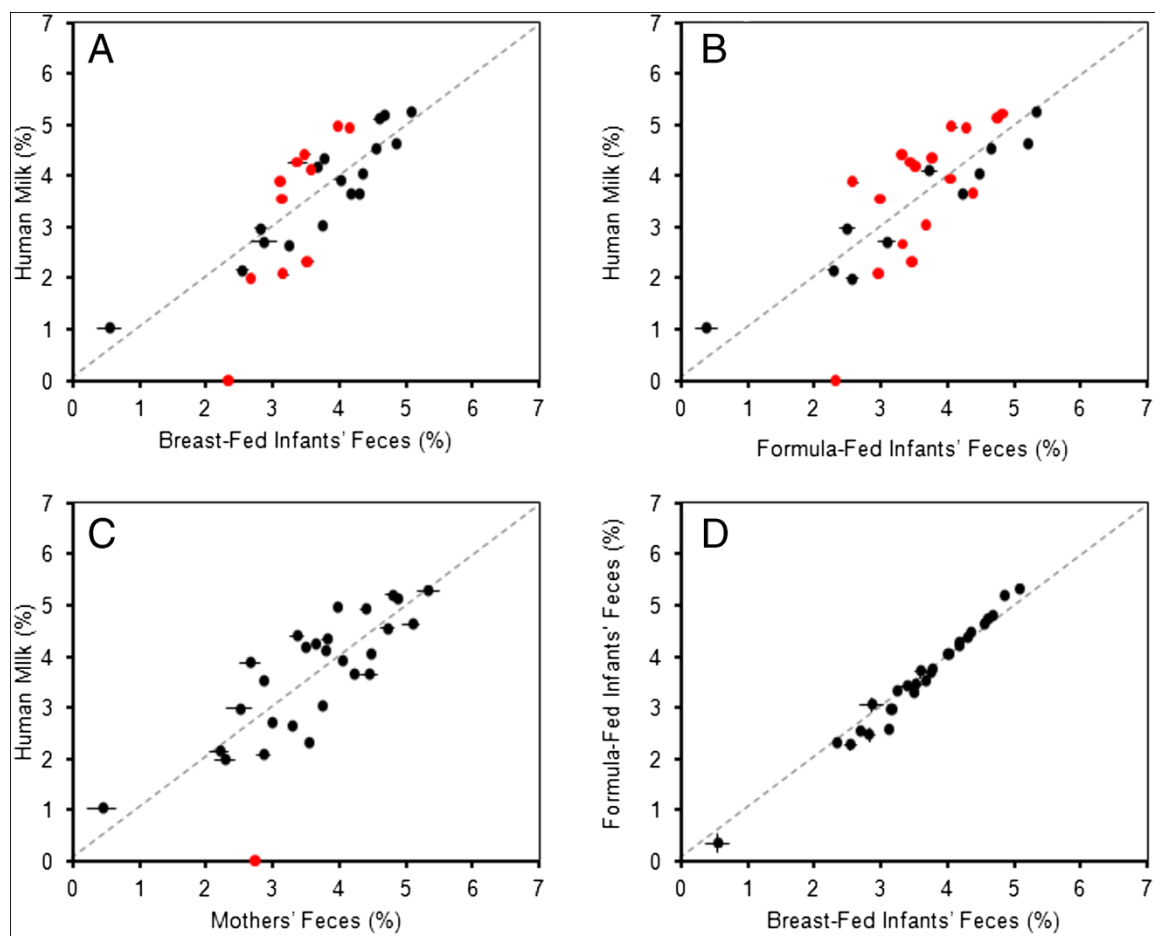


Figure 4.6. Pair-wise comparison of categorized open reading frames from human milk versus infants' and mothers' feces. Pair-wise comparisons for the human milk metagenome versus (A) breast-fed infants' feces, (B) formula-fed infants' feces and (C) mothers' feces are shown. For comparison, a plot of breast-fed infants' feces and formula-fed infants' feces (D) is also shown. Each point represents a different SEED subsystem and its relative abundance within the human milk metagenome compared to the fecal metagenomes. Points lying on or near the dotted line have equal or similar abundance in both metagenomes. Red dots signify those with significantly different proportions between the two metagenomes (Student's *t*-test, $P < 0.05$). Breast-fed and formula-fed infants' feces values are an average of five individuals, and mothers' feces values are an average of three individuals. All subjects are unrelated.



Immune-modulatory DNA motifs in human milk and feces

When contigs were searched for the presence of immune suppressive motifs, TCAAGCTTGA was found in 0.02% of the human-milk assembled contigs (11 sites, Table 4.2) with an occurrence 1.5 times that of the human genome alone (once per 844,000 bp compared to once per 1,276,500 bp in the human genome, Z-score -1.6). The contigs positive for TCAAGCTTGA aligned to the genomes of *Pseudomonas* (45%), *Nocardia* (9%), *Staphylococcus* (9%) and contigs of unknown origin (36%, Table 4.3). When the contigs from BF-infants' feces, FF-infants' feces and mothers' feces were scanned for TCAAGCTTGA, it was found at a relative occurrence of 1.19, 1.64, and 1.64 times that in the human genome, respectively (Table 4.2). Another immune suppressive site, TTAGGG was observed 1,684 times in the human milk metagenome (3.2% of contigs), and at a relative occurrence 0.48 times that of the human genome (once per 5,600 bp compared to once per 2,670 bp, Z-score 22.54, Table 4. 2). Contigs containing TTAGGG corresponded to genomes of *Staphylococcus* (59%), *Pseudomonas* (10%), *Lactobacillus* (0.5%), 21 other known prokaryotic genomes (2.7%), and contigs from unknown genomes (27%, Table 4.3). When the contigs from BF-infants' feces, FF-infants' feces and mothers' feces were scanned for TTAGGG, this sequence was observed at a relative occurrence of 0.33, 0.18 and 0.26 times that in the human genome, respectively (Table 4.2). Assembled contigs were also searched for the presence of synthetically-assembled immune suppressive or immune stimulatory DNA motifs (7 and 5 motifs, respectively), such as those used in vaccine production (Supplemental Table S4.4; 135–139). No synthetically-assembled sequences were observed in the human-milk contigs, whereas three motifs were found in less than 5 ×

10^{-4} % of contigs from the fecal metagenomes (maximum of 4 hits per 834,774 contigs, Supplemental Table S4.4).

Table 4.2. Occurrence of immune suppressive motifs in various metagenomes.
DNA motifs TTAGGG and TCAAGCTTGA were searched for in contigs derived from human milk, breast-fed infants' feces (BF infant), formula-fed infants' feces (FF infant) and mothers' feces. Relative occurrence is in comparison to the human genome.

Sequence	Number of Hits	Base Pairs per Hit		Relative Occurrence (Z-score)
TCAAGCTTGA	11	844,000	(Human Milk)	1.51 (-1.6)
	344	1,077,000	(BF Infant)	1.19 (-0.74)
	124	779,000	(FF Infant)	1.64 (-1.84)
	268	777,000	(Mother)	1.64 (-1.85)
	2,245	1,276,500	(Human Genome)	
TTAGGG	1,684	5,600	(Human Milk)	0.48 (22.54)
	18,118	8,200	(BF Infant)	0.33 (42.54)
	16,410	15,000	(FF Infant)	0.18 (94.85)
	20,612	10,200	(Mother)	0.26 (57.92)
	1,082,623	2,670	(Human Genome)	

Table 4.3. Occurrence of immune suppressive motifs TTAGGG and TCAAGCTTGA in contigs from human milk.

Sequence	Genus	Number of Hits
TCAAGCTTGA	<i>Pseudomonas</i>	5
	<i>Nocardia</i>	1
	<i>Staphylococcus</i>	1
	Unknown	4
TTAGGG	<i>Staphylococcus</i>	1000
	<i>Pseudomonas</i>	169
	<i>Lactobacillus</i>	8
	<i>Bacillus</i>	6
	<i>Streptococcus</i>	6
	<i>Streptomyces</i>	4
	<i>Tetragenococcus</i>	4
	Other	25
	Unknown	461

4.5 Discussion

Genera within human milk

Determining the human milk metagenome, a bodily fluid notably absent from the human microbiome project (140), is crucial for enabling better insight on the process of infant GI colonization and immune development. Pooling DNA from ten human milk samples and subjecting it to Illumina sequencing has demonstrated the large diversity of the human milk metagenome with over 56,000 contigs aligning to 177 bacterial genera (Figure 4.2). Previous studies investigating the microbiome of human milk have used both culture-dependent and -independent approaches. Using 16S rRNA sequencing, Hunt *et al.* have reported several predominant species in human milk including a core of genera found in 16 human milk samples across time: *Streptococcus*, *Staphylococcus*, *Serratia*, *Pseudomonas*, *Corynebacteria*, *Ralstonia*, *Propionibacteria*, *Sphingomonas*, and *Bradyrhizobiaceae* (47). Other studies showed colostrum was populated mostly by *Weisella* and *Leuconostoc*, followed by *Staphylococcus*, *Streptococcus*, and *Lactococcus*, and that *Akkermansia* were more prevalent in overweight mothers (48, 56). Using a best hit analysis of the 51 bp Illumina reads, alignments for *Akkermansia*, *Propionibacteria*, *Sphingomonas* and *Weisella* were observed (Supplemental Table S4.2), but because of the small number of base pairs used for the alignment (51 bp) and the lack of assembled contigs associated with these microbes, their presence in the milk samples used here is a tentative identification. Using PCR-denaturing gradient gel electrophoresis and quantitative PCR, two studies from Martin *et al.* reported the presence of *Bifidobacterium breve*, *B. adolescentis*, *B. bifidum* and *B. dentium* in human milk, which differs from the

current findings (Figure 4.2; 45, 129). This is likely due to the method of DNA extraction used in this study, as bead-beating as a means to extract DNA from the hard to rupture *Bifidobacterium* was not incorporated (141). The differences between the previously reported microbial communities and this analysis may also be due, in part, to the geographic location of the mothers, which has been shown to greatly impact the microbiome of individuals (142). Furthermore, other differences between this characterization of the milk microbiome and those previously reported may be attributed to the means of milk collection. In comparison to previous studies where human milk was expressed from an aseptic breast (45–48, 127–130), the current method determines the total microbiome (i.e. metagenome) ingested by the infant (from a non-sterilized breast), indicative of what an infant would receive from its mother during suckling.

Because the samples used here were collected from a non-sterilized breast, it could be hypothesized the human milk metagenome reported here would be similar to that of the skin microbiome. Although no reference database was freely available within MG-RAST for comparison, the metagenome of human milk is similar to previously reported skin profiles in that there is a large proportion of *Staphylococcus*, which is found in moist areas of skin. These moist areas, such as the antecubital fossa (inner fold of the elbow), also contain Betaproteobacteria, such as *Burkholderia* and *Bordetella*, which are present in the milk metagenome (Figure 4.2; 143, 144). The human milk metagenome is also similar to drier areas of the skin such as the plantar heel, which contains Gamaproteobacteria such as *Pseudomonas* (143). The human milk metagenome is, however, more similar to fecal microbiomes (as described in 16S

rRNA studies) due to the large proportion of Firmicutes bacteria within human milk, which is a very minor member of the skin microbiome (Figure 4.4; 143, 144). Also, the skin of adults tends to contain a high level of *Propionibacteria*, which notably tends to colonize the skin of cesarean-section birthed babies, whereas this genus is minimally represented in this human milk sample using a best hit analysis of the 51 bp Illumina reads (0.2%, Supplemental Table S4.2; 145, 146). This observation suggests that mother's milk may prove useful as a skin lotion, to re-balance the skin microbiome of C-section babies.

Phylogenetic differences between human milk and feces

Comparing the metagenome of human milk to that of publicly available infants' and mothers' fecal profiles provides insight as to how human milk may lead to proper colonization of the infant gut. When comparing the human milk metagenome to the infant fecal metagenome, there are numerous differences. For example, the metagenome of BF-infants' feces contains a high proportion of Actinobacteria (70.4%, Figure 4.4), which correlates with previous studies demonstrating a high abundance of *Bifidobacterium* in the feces of BF-infants whereas FF-infants had a more varied microbiota (120, 142, 147). Contigs from human milk, however, aligned mostly with Proteobacteria and Firmicutes (65.1% and 34.6%, respectively, Figure 4.4). At the phylum level, the present milk metagenome was less diverse than the fecal metagenomes as over 99% of the contigs were from just two phyla, Proteobacteria and Firmicutes (Figure 4.4). FF-infants' feces and mothers' feces were similar in that they both contained contigs aligning to the phylum Bacteroidetes (17.6% and 20.6%,

respectively), whereas Bacteroidetes was a very minor component of BF-infants' feces and human milk (0.3% and 0.02%, respectively, Figure 4.4). Also, the similar proportion of Firmicutes in human milk compared to mothers' feces (34.6% and 59.6%, respectively, Figure 4.4) correlates with the hypothesis that mothers' milk may be inoculated by immune cells carrying bacteria from the GI tract of the mother to her breast (148–150). This may be a mechanism by which the human milk microbiome is shaped by the general health of the mother, including her weight (48).

Functionality of the human milk metagenome

Using Illumina sequencing of all DNA within milk samples permits the prediction of ORFs within assembled contigs and allows for determination of the functional capability of the milk metagenome. A total of 41,352 ORFs were predicted, including those for basic cell function, as well as those that may enable the bacteria to remain in human milk, such as ORFs for carbohydrate metabolism (5.7% of ORFs, Figure 4.3). The predominant carbohydrate in human milk, lactose, is a potential carbon source for human milk bacteria, and therefore the presence of ORFs associated with its metabolism (6.7% of carbohydrate-associated metabolism, Figure 4.3) is expected. Another carbon source for bacteria in human milk is HMOs, which cannot be digested by the infant (151). These oligosaccharides, which are heavily fucosylated and readily digested by *Bifidobacteria*, are thought to be responsible for the colonization of BF-infants with high levels of *Bifidobacteria* (152). Due to a lack of contigs aligning to *Bifidobacteria* (Figure 4.2), no ORFs encoding genes for HMOs were observed (Figure 4.3). Recently, HMOs have also been correlated with increased abundance of

Staphylococcus within human milk, regardless of their inability to utilize the HMOs as a carbon source (153). The predominance of *Staphylococcus*-aligning contigs in these milk samples supports these findings (Figure 4.2). Furthermore, there was a significantly higher number of ORFs related to nitrogen metabolism within the human milk metagenome in comparison to BF- and FF-infants' feces (Figure 4.5, $P < 0.05$). Because human milk contains 1.48-2.47 g of nitrogen per 100 g of milk, the bacteria within human milk may use it as a nutrient source in addition to lactose and HMOs (154).

Human milk contains an abundance of immune cells, antibodies and antimicrobial proteins (such as lactoferrin, CD14, α -lactalbumin, and lysozyme), and therefore the bacteria residing within human milk must harbor mechanisms to combat the milk-endogenous immune system (36, 40, 155). For example, the metagenome of human milk includes ORFs for stress response and defense (4.0% and 4.5% of all ORFs, respectively) including those for oxidative stress (40.3% of stress-related ORFs) and toxic compound resistance (60.2% of defense ORFs, Figure 4.3). The human milk metagenome also contains ORFs for both heat and cold shock (12% and 4% of stress-related ORFs, Figure 4.3), which may enable the bacteria to persist in milk post-breast pumping, refrigeration and reheating. This may be of particular importance as human milk banks gain more popularity over time. For example, as described in a recent review by Urbaniak *et al.*, some milk banks deem pasteurization of breast milk unnecessary, while others have an upper limit of 10^5 organisms per mL (156). In unpasteurized banked milk and in-home stored milk, if some organisms are able to survive the storage and re-heating process better than others, the bacterial

profile of human milk may change to favor better surviving (and not necessarily more beneficial) bacteria. Furthermore, ORFs encoding genes related to virulence and disease (4.5% of all ORFs, Figure 4.3), are also observed in the human milk metagenome. These ORFs could allow some of the human milk microbes, such as *Staphylococcus aureus*, to cause mastitis in humans when the balance of human milk-antimicrobials to microbes is tilted towards microbial growth (157). For example, some bacteria within human milk harbor antibiotic resistance genes (60.2% of virulence associated ORFs) allowing them to proliferate regardless of the mother's potential antibiotic use, and some bacteria are able to produce bacteriocins (2.7% of virulence associated ORFs, Figure 4.3), which could impact the growth of other, less virulent, microbes within the community.

Immune-modulatory landscape of the human milk metagenome

Because human milk contains a broad spectrum of microbes at the genus level (Figure 4.2), it likely contributes significantly towards effective colonization of the infant GI tract. In the case of banked human milk, which is Holder pasteurized (65°C for 5–30 min), most bacteria are destroyed, but their proteins and DNA remain (158). The presence of non-viable bacteria and bacterial DNA in human milk, which are indistinguishable from live bacteria using the current approach of DNA isolation and sequencing, may be a way to prime the infant immune system and lead to tolerance of the trillions of bacteria that will inhabit the gut following birth. For example, the immune suppressive motifs, TTAGGG and TCAAGCTTGA (125), are present in 3.0% and 0.02% of the 56,950 human milk-contigs, respectively (1,684 sites, and 11 sites,

Table 4. 2). The occurrence of the immune suppressive motifs is similar to that in the metagenomes of BF- and FF- infants' feces, as well as mothers' feces. This suggests that having a diverse community of microbes may lead to a similar abundance of immune suppressive motifs, regardless of the genera present in the sample.

Interestingly, the immune suppressive motif TTAGGG was found in higher abundance in the human genome than in bacterial contigs (one per 2,670 bp in the human genome compared to one per 5,600 bp in the bacterial contigs, Table 4.2). Colostrum and mature human milk contain between 5×10^8 to 4×10^9 leukocytes/L and between 1×10^8 to 4×10^9 leukocytes/L, respectively, which are mostly macrophages (55%-60%) and neutrophils (30%-40%), with natural killer cells representing up to 12% of the population (44, 159). This suggests that ingestion of the mothers' DNA, through ingestion of her immune cells and any free circulating DNA may also lead to proper immune development through a balance of concomitant exposure to immune stimulatory bacterial CpGs and immune suppressive DNA in the mothers' genome and bacterial genomes.

Conclusions

Current microbiome studies characterizing the microbial communities of various anatomical niches have revealed vast differences between healthy individuals (140). These differences can often be attributed to the host's environment and diet. As demonstrated previously by preliminary 16S rRNA sequencing, the human milk microbiome is similar to other areas of the body in that its composition is unique to each individual (47). Milk has evolved as the first nutrient source for mammals ex

utero, with a high level of inter-mother diversity as to the proportions of bacterial genera, immune proteins and nutrients within it (56). Perhaps, it is the diversity and/or sequences of DNA within the milk metagenome that is beneficial for infants, as opposed to any one specific bacterial genus or species. Recent reviews on human milk outline the phylotypes of bacteria within human milk, but only speculate on the function of the human milk microbiome due to a lack of data on the functional capacity of the microbes within human milk (156, 160). Because of this, a better understanding of the human milk metagenome on a functional level rather than a solely phylogenetic level was sought. The discovery of the abundance of immune suppressive DNA motifs observed within bacterial and human DNA from human milk, as well as ORFs within the human milk metagenome that allow bacteria to persist in the biological fluid provides a first glance into the functionality of the milk metagenome. Further studies should include those determining the efficacy of milk DNA to modulate the immune system in the GI tract, and a more exhaustive look at the metagenome of human milk and how it relates to infant health outcomes.

Chapter 5

General Discussion

5.1 sCD14 in human milk

The research findings presented in the above thesis demonstrate the complex nature of human milk and specifically attempts to clarify the digestive fate and function of one bioactive human milk protein, sCD14. The bacterial sensor, sCD14, was able to partially evade digestion in the GI tract of rat pups (Figures 2.1 and 2.2), and enter intact into the blood stream of those animals studied (Figures 2.3 and 2.4; 90).

Additional and complementary studies in mice showed ingested sCD14 to contribute almost one third (26.7%) of all circulating sCD14 in the bloodstream of pups (51.6%, Figure 3.1). The uptake of ingested sCD14 along the digestive tract of rodents was also demonstrated in human cells in vitro (Figure 3.4; 161). When Caco-2 monolayers were grown in Transwell plates they were observed to transfer hrCD14 from the apical wells to basal wells (Figure 3.4), and this transfer appeared to be dependent on the membrane-bound receptor, TLR4 (Figure 3.7; 161). The transfer of ingested sCD14 to the circulatory system in human infants, however, has yet to be studied. In mice, sCD14 post-ingestion and post-transfer into the blood remained functional, as measured by sCD14's ability to detect injected LPS and elicit an immune response (IL-6 and TNF- α production, Figure 3.3; 161). Additionally, WT mouse pups not receiving enteral sCD14 because they were nursed by CD14^{-/-} mothers, were shown to have a significantly higher IL-6 response in comparison to those ingesting sCD14 (Figure 3.3; 161). Therefore, if ingested sCD14 behaves similarly in human infants as it does in the rodent models and in vitro work presented here, human milk derived sCD14 likely impacts the innate immune response of breastfeeding infants.

Given CD14's recent implication in inflammatory diseases such as obesity and insulin resistance, as well as NEC, the lack of sCD14 in infant formulas may be a contributing factor in the negative health outcomes that are associated with infant formula feeding (57, 58, 112). Future experiments could include the supplementation of infant formulas with a physiologically similar level of human sCD14 to determine if infant health is altered following ingestion. Infants normally receiving formula would either ingest sCD14-supplemented or non-supplemented formula and health outcomes such as NEC development, insulin resistance and obesity would be tracked until later in life. Such a study would require a source of hrCD14, which is costly to produce by today's standard protein production methods (Chinese Hamster Ovary cells, \$319 for 50 µg, R&D Systems). In the past, tobacco has been genetically modified to produce human sCD14 that is bioactive (162). Tobacco, however, is not an ideal plant for protein production due downstream purification costs to obtain a protein of ingestible or injectable grade (163, 164). Future experiments could include the production of hrCD14 in a readily edible host, such as within the hypoallergenic seeds of the rice plant (165). Seeds harbouring hrCD14 could easily be milled into rice flour, and added directly to commercially available infant formulas (166). If proven a successful avenue for production of edible sCD14, the production of other human milk proteins in rice (167), such as milk lactoferrin, could also be achieved thus presenting a way to further affordably humanize current infant formulas.

5.2 Bacteria and DNA in human milk

In addition to exploring the digestive fate and function of sCD14, this thesis also asked the ‘systems biology’ questions of who is there and what might they be doing?

Therefore, the data presented in this thesis also elucidates the metagenome of human milk and is the first to this author’s knowledge to describe the immune modulatory potential of human milk DNA. DNA isolated and sequenced from human milk aligned predominantly to the genomes of *Pseudomonas* (61.1%), *Staphylococcus* (33.4%) and *Streptococcus* (0.5%), which correlates with previous studies that analyzed the bacteria in human milk by 16S rRNA sequencing (Figure 4.2; 47, 48, 96, 129). The metagenome of human milk was found to differ from that of infants’ and mothers’ fecal profiles in many ways, including a low level of diversity in the human milk metagenome (over 99% of the contigs were from just two phyla, Proteobacteria and Firmicutes) compared to that of infants’ and mothers’ feces (Figure 4.4; 96). A total of 41,352 ORFs were predicted within the human milk metagenome, including those for protein metabolism (8.2% of all ORFs), defense (4.5% of all ORFs), and stress response (4.0% of all ORFs, Figure 4.3; 96). The human milk metagenome was also shown to harbour the immune suppressive motifs, TTAGGG and TCAAGCTTGA (125), present in 3.0% and 0.02% of the 56,950 human milk-contigs, respectively (Table 4.2; 96). The same two immune suppressive motifs were detected in the fecal metagenomes of both infants and mothers, suggesting that having a diverse community of microbes may lead to similar abundance of immune suppressive motifs. It is likely that infant formulas do not harbour any immune-modulatory DNA motifs, except for formulas specifically supplemented with bacteria (for example, Nestlé Good

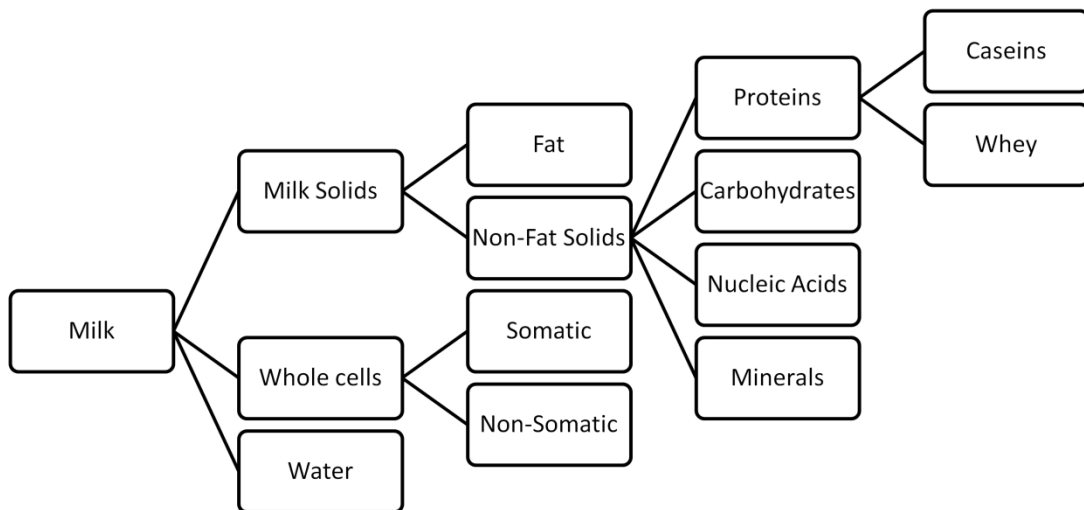
Start Probiotic with *Bifidobacterium lactis*) (http://www.nestle-baby.ca/en/products/formula/starter/goodstart_probiotic.htm, website accessed January 9, 2014)

Due to the immune modulatory potential of DNA sequences elsewhere in the body (51, 52), it would be of interest to determine if the DNA found in human milk is able to alter the immune response of cells along the GI tract, including those of the mucosal immune system. For example, future studies should first determine the quantity of cell free DNA within human milk, which was indistinguishable from within-cell DNA using the methods of this thesis. DNA from human milk can then be used to treat cells in vitro to determine if the DNA elicits an immune response, or perhaps suppresses an immune response to pathogenic bacteria if used as a pre-treatment. Further studies could be completed in rodent or porcine models to determine the in vivo effects of human milk DNA on the immune system. Such a study would be of value, given that infant formulas have recently been fortified with probiotics but the safety and efficacy of these products has yet to be established (168). Also, although infant formulas contain nucleic acids, no commercially available infant formula currently contains oligonucleotides (e.g. Nestlé, Enfamil and Similac, amongst others). Therefore, determining which bacteria and DNA motifs an infant is exposed to and the effects of those motifs on the GI tract and immune system would provide guidance to commercial infant formula manufacturers on how best to further humanize the bovine milk and plant based powders.

5.3 Conclusions

The complex and dynamic composition of human milk in comparison to infant formula warrants the humanization of infant formulas to better suit the immunological needs of the developing infant. The data presented here show that bioactives in human milk, such as sCD14, can survive GI transit, are transferred intact across the GI tract, and contribute to the functional pool of circulating blood proteins (90, 161). Furthermore, given the breadth of microbial diversity within human milk and the immune modulatory DNA that the microbes harbour (96), the supplementation of infant formulas should not be limited to human proteins. The current understanding of the gross composition of human milk (Figure 5.1) should be altered to include other milk components now better understood, such as cells of somatic and non-somatic origin as well as nucleic acids (Figure 5.1). Based on this new human milk model, the humanization of infant formulas to better emulate the composition of human milk should include the addition of human proteins, DNA and bacteria.

Figure 5.1. The gross composition of human milk, modified to include cells of somatic and non-somatic origin as well as nucleic acids. (Modified from 11, 12)



References

1. Hinde K, German JB. (2012). Food in an evolutionary context: insights from mother's milk. *J Sci Food Agric* 92, 2219–2223.
2. World Health Organization, UNICEF. (2009). Baby-friendly hospital initiative: revised, updated and expanded for integrated care.
3. Black RE, Victora CG, Walker SP, Bhutta ZA, Christian P, de Onis M, Ezzati M, Grantham-McGregor S, Katz J, Martorell R, Uauy R. (2013). Maternal and child undernutrition and overweight in low-income and middle-income countries. *The Lancet* 382, 427–451.
4. Wardlaw T, Salama P, Brocklehurst C, Chopra M, Mason E. (2010). Diarrhoea: why children are still dying and what can be done. *The Lancet* 375, 870–872.
5. Lamberti LM, Fischer Walker CL, Noiman A, Victora C, Black RE. (2011). Breastfeeding and the risk for diarrhea morbidity and mortality. *BMC Public Health* 11, S15.
6. Ballard O, Morrow AL. (2013). Human milk composition. *Pediatr Clin North Am* 60, 49–74.
7. Hester SN, Hustead DS, Mackey AD, Singhal A, Marriage BJ. (2012). Is the macronutrient intake of formula-fed infants greater than breast-fed infants in early infancy? *J Nutr Metab* 2012, 1–13.
8. Liao Y, Alvarado R, Phinney B, Lönnerdal B. (2011). Proteomic characterization of human milk fat globule membrane proteins during a 12 month lactation period. *J Proteome Res* 10, 3530–3541.
9. Liao Y, Alvarado R, Phinney B, Lönnerdal B. (2011). Proteomic characterization of human milk whey proteins during a twelve-month lactation period. *J Proteome Res* 10, 1746–1754.
10. Dalglish DG, Corredig M. (2012). The structure of the casein micelle of milk and its changes during processing. *Annu Rev Food Sci Technol* 3, 449–467.
11. Chandan RC. (1997). Dairy-based ingredients. Egan Press. St. Paul, MN, USA.
12. Perry SE. (2010). Maternal child nursing care. Mosby Elsevier. Maryland Heights, MO, USA.
13. Levieux D, Ollier A. (1999). Bovine immunoglobulin G, beta-lactoglobulin, alpha-lactalbumin and serum albumin in colostrum and milk during the early post partum period. *J Dairy Res* 66, 421–430.
14. Lien EL. (2003). Infant formulas with increased concentrations of alpha-lactalbumin. *Am J Clin Nutr* 77, 1555S–1558S.
15. Levay PF, Viljoen M. (1995). Lactoferrin: a general review. *Haematologica* 80, 252–267.
16. Nuijens JH, van Berkel PH, Schanbacher FL. (1996). Structure and biological actions of lactoferrin. *J Mammary Gland Biol Neoplasia* 1, 285–295.
17. Sánchez L, Calvo M, Brock JH. (1992). Biological role of lactoferrin. *Arch Dis Child* 67, 657–661.
18. Haneberg B. (1974). Immunoglobulins in feces from infants fed human or bovine milk. *Scand J Immunol* 3, 191–197.
19. Prentice A, Ewing G, Roberts SB, Lucas A, MacCarthy A, Jarjou LM, Whitehead RG. (1987). The nutritional role of breast-milk IgA and lactoferrin. *Acta Paediatr Scand* 76, 592–598.
20. Davidson LA, Lönnerdal B. (1987). Persistence of human milk proteins in the breast-fed infant. *Acta Paediatr Scand* 76, 733–740.
21. Goldman AS, Garza C, Schanler RJ, Goldblum RM. (1990). Molecular forms of lactoferrin in stool and urine from infants fed human milk. *Pediatr Res* 27, 252–255.

22. Ladinsky MS, Huey-Tubman KE, Bjorkman PJ. (2012). Electron tomography of late stages of FcRn-mediated antibody transcytosis in neonatal rat small intestine. *Mol Biol Cell* 23, 2537–2545.
23. Lönnerdal B, Jiang R, Du X. (2011). Bovine lactoferrin can be taken up by the human intestinal lactoferrin receptor and exert bioactivities. *J Pediatr Gastroenterol Nutr* 53, 606–614.
24. Walker A. (2010). Breast milk as the gold standard for protective nutrients. *J Pediatr* 156, S3–S7.
25. Manzoni P. (2009). Bovine lactoferrin supplementation for prevention of late-onset sepsis in very low-birth-weight neonates: a randomized trial. *JAMA* 302, 1421.
26. Manzoni P, Mostert M, Stronati M. (2011). Lactoferrin for prevention of neonatal infections. *Curr Opin Infect Dis* 24, 177–182.
27. Manzoni P, Stolfi I, Messner H, Cattani S, Laforgia N, Romeo MG, Bollani L, Rinaldi M, Gallo E, Quercia M, Maule M, Mostert M, Decembrino L, Magaldi R, Mosca F, Vagnarelli F, Memo L, Betta PM, Stronati M, Farina D. (2011). Bovine lactoferrin prevents invasive fungal infections in very low birth weight infants: a randomized controlled trial. *Pediatrics* 129, 116–123.
28. Labeta MO, Vidal K, Nores JE, Arias M, Vita N, Morgan BP, Guillemot JC, Loyaux D, Ferrara P, Schmid D, Affolter M, Borysiewicz LK, Donnet-Hughes A, Schiffrin EJ. (2000). Innate recognition of bacteria in human milk is mediated by a milk-derived highly expressed pattern recognition receptor, soluble CD14. *J Exp Med* 191, 1807–1812.
29. Vidal K, Labeta MO, Schiffrin EJ, Donnet-Hughes A. (2001). Soluble CD14 in human breast milk and its role in innate immune responses. *Acta Odontol Scand* 59, 330–334.
30. Kitchens RL, Thompson PA, Viriyakosol S, O'Keefe GE, Munford RS. (2001). Plasma CD14 decreases monocyte responses to LPS by transferring cell-bound LPS to plasma lipoproteins. *J Clin Invest* 108, 485–493.
31. Eggesbo M, Moen B, Peddada S, Baird D, Rugtveit J, Midtvedt T, Bushel PR, Sekelja M, Rudi K. (2011). Development of gut microbiota in infants not exposed to medical interventions. *APMIS* 119, 17–35.
32. Fan W, Huo G, Li X, Yang L, Duan C. (2013). Impact of diet in shaping gut microbiota revealed by a comparative study in infants during the six months of life. *J Microbiol Biotechnol* 24, 133–143.
33. Matamoros S, Gras-Leguen C, Le Vacon F, Potel G, de La Cochetiere MF. (2013). Development of intestinal microbiota in infants and its impact on health. *Trends Microbiol* 21, 167–173.
34. Palmer C, Bik EM, DiGiulio DB, Relman DA, Brown PO. (2007). Development of the human infant intestinal microbiota. *Plos Biol* 5, 1556–1573.
35. Warren HS, Fitting C, Hoff E, Adib-Conquy M, Beasley-Topliffe L, Tesini B, Liang X, Valentine C, Hellman J, Hayden D, Cavaillon JM. (2010). Resilience to bacterial infection: difference between species could be due to proteins in serum. *J Infect Dis* 201, 223–232.
36. Blais DR, Harrold J, Altosaar I. (2006). Killing the messenger in the nick of time: persistence of breast milk sCD14 in the neonatal gastrointestinal tract. *Pediatr Res* 59, 371–376.
37. Nizet V. (2009). *Essentials of glycobiology*, 2nd ed. Cold Spring Harbor Laboratory Press. Cold Spring Harbor, NY, USA.
38. Godowski PJ. (2005). A smooth operator for LPS responses. *Nat Immunol* 6, 544–546.
39. Jiang Z, Georgel P, Du X, Shamel L, Sovath S, Mudd S, Huber M, Kalis C, Keck S, Galanos C, Freudenberg M, Beutler B. (2005). CD14 is required for MyD88-independent LPS signaling. *Nat Immunol* 6, 565–570.

40. Spencer WJ, Binette A, Ward TL, Davis L, Blais DR, Harrold J, Mack DR, Altosaar I. (2010). Alpha-lactalbumin in human milk alters the proteolytic degradation of soluble CD14 by forming a complex. *Pediatr Res* 68, 490–493.
41. Zanon I, Ostuni R, Marek LR, Barresi S, Barbalat R, Barton GM, Granucci F, Kagan JC. (2011). CD14 controls the LPS-induced endocytosis of Toll-like receptor 4. *Cell* 147, 868–880.
42. Vamadevan AS, Fukata M, Arnold ET, Thomas LS, Hsu D, Abreu MT. (2010). Regulation of Toll-like receptor 4-associated MD-2 in intestinal epithelial cells: a comprehensive analysis. *Innate Immun* 16, 93–103.
43. Hassiotou F, Geddes DT, Hartmann PE. (2013). Cells in human milk: state of the science. *J Hum Lact Off J Int Lact Consult Assoc* 29, 171–182.
44. Jin YY, Wei Z, Cao RM, Xi W, Wu SM, Chen TX. (2011). Characterization of immunocompetent cells in human milk of Han Chinese. *J Hum Lact* 27, 155–162.
45. Martin R, Jimenez E, Heilig H, Fernandez L, Marin ML, Zoetendal EG, Rodriguez JM. (2009). Isolation of bifidobacteria from breast milk and assessment of the bifidobacterial population by PCR-denaturing gradient gel electrophoresis and quantitative real-time PCR. *Appl Env. Microbiol* 75, 965–969.
46. Martin V, Manes-Lazaro R, Rodriguez JM, Maldonado-Barragan A. (2011). *Streptococcus lactarius* sp. nov., isolated from breast milk of healthy women. *Int J Syst Evol Microbiol* 61, 1048–1052.
47. Hunt KM, Foster JA, Forney LJ, Schutte UM, Beck DL, Abdo Z, Fox LK, Williams JE, McGuire MK, McGuire MA. (2011). Characterization of the diversity and temporal stability of bacterial communities in human milk. *PLoS One* 6, e21313.
48. Cabrera-Rubio R, Collado MC, Laitinen K, Salminen S, Isolauri E, Mira A. (2012). The human milk microbiome changes over lactation and is shaped by maternal weight and mode of delivery. *Am J Clin Nutr* 96, 544–551.
49. Klindworth A, Pruesse E, Schweer T, Peplies J, Quast C, Horn M, Glockner FO. (2012). Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies. *Nucleic Acids Res* 41, e1–e1.
50. Sharon I, Banfield JF. (2013). Genomes from metagenomics. *Science* 342, 1057–1058.
51. He X, Jia H, Jing Z, Liu D. (2013). Recognition of pathogen-associated nucleic acids by endosomal nucleic acid-sensing Toll-like receptors. *Acta Biochim Biophys Sin* 45, 241–258.
52. Holm CK, Paludan SR, Fitzgerald KA. (2013). DNA recognition in immunity and disease. *Curr Opin Immunol* 25, 13–18.
53. Hofmann C, Dunger N, Doser K, Lippert E, Siller S, Edinger M, Falk W, Obermeier F. (2014). Physiologic TLR9-CpG-DNA interaction is essential for the homeostasis of the intestinal immune system. *Inflamm Bowel Dis* 20, 136–143.
54. Yu Y-Z, Li N, Ma Y, Wang S, Yu W-Y, Sun Z-W. (2013). Three types of human CpG motifs differentially modulate and augment immunogenicity of nonviral and viral replicon DNA vaccines as built-in adjuvants. *Eur J Immunol* 43, 228–239.
55. Frey EA, Miller DS, Jahr TG, Sundan A, Bazil V, Espevik T, Finlay BB, Wright SD. (1992). Soluble CD14 participates in the response of cells to lipopolysaccharide. *J Exp Med* 176, 1665–1671.
56. Collado MC, Laitinen K, Salminen S, Isolauri E. (2012). Maternal weight and excessive weight gain during pregnancy modify the immunomodulatory potential of breast milk. *Pediatr Res* 72, 77–85.
57. Roncon-Albuquerque R, Moreira-Rodrigues M, Faria B, Ferreira AP, Cerqueira C, Lourenco AP, Pestana M, von Hafe P, Leite-Moreira AF. (2008). Attenuation of the

- cardiovascular and metabolic complications of obesity in CD14 knockout mice. *Life Sci* 83, 502–510.
58. Fernandez-Real JM, Perez del Pulgar S, Luche E, Moreno-Navarrete JM, Waget A, Serino M, Sorianello E, Sanchez-Pla A, Pontaque FC, Vendrell J, Chacon MR, Ricart W, Burcelin R, Zorzano A. (2011). CD14 modulates inflammation-driven insulin resistance. *Diabetes* 60, 2179–2186.
 59. Lundell AC, Adlerberth I, Lindberg E, Karlsson H, Ekberg S, Aberg N, Saalman R, Hock B, Steinkasserer A, Hesselmar B, Wold AE, Rudin A. (2007). Increased levels of circulating soluble CD14 but not CD83 in infants are associated with early intestinal colonization with *Staphylococcus aureus*. *Clin Exp Allergy* 37, 62–71.
 60. Chatterton DE, Nguyen DN, Bering SB, Sangild PT. (2013). Anti-inflammatory mechanisms of bioactive milk proteins in the intestine of newborns. *Int J Biochem Cell Biol* 45, 1730–1747.
 61. Kuo TT, Baker K, Yoshida M, Qiao SW, Aveson VG, Lencer WI, Blumberg RS. (2010). Neonatal Fc receptor: from immunity to therapeutics. *J Clin Immunol* 30, 777–789.
 62. Taylor SN, Basile LA, Ebeling M, Wagner CL. (2009). Intestinal permeability in preterm infants by feeding type: mother's milk versus formula. *Breastfeed Med* 4, 13–17.
 63. Teichberg S, Isolauri E, Wapnir RA, Roberts B, Lifshitz F. (1990). Development of the neonatal rat small intestinal barrier to nonspecific macromolecular absorption: effect of early weaning to artificial diets. *Pediatr Res* 28, 31–37.
 64. Davis LD, Spencer WJ, Pham VT, Ward TL, Blais DR, Mack DR, Kaplan H, Altosaar I. (2011). (14)C radiolabeling of proteins to monitor biodistribution of ingested proteins. *Anal Biochem* 410, 57–61.
 65. Davis LD, Spencer WJ, Mack DR, Altosaar I. (2011). Maternal separation and gastrointestinal transit time in neonate rats. *Lab Anim* 45, 280–282.
 66. Eriksen EK, Holm H, Jensen E, Aaboe R, Devold TG, Jacobsen M, Vegarud GE. (2010). Different digestion of caprine whey proteins by human and porcine gastrointestinal enzymes. *Br J Nutr* 104, 374–381.
 67. DeSesso JM, Williams AL, John EM. (2008). Annual reports in Medical Chemistry. Chapter 21: Contrasting the gastrointestinal tracts of mammals: factors that influence absorption. Academic Press. Salt Lake City, UT, USA.
 68. Nicholas KR, Hartmann PE, McDonald BL. (1981). Alpha-Lactalbumin and lactose concentrations in rat milk during lactation. *Biochem J* 194, 149–154.
 69. Jacque B, Stephan K, Smirnova I, Kim B, Gilling D, Poltorak A. (2006). Mice expressing high levels of soluble CD14 retain LPS in the circulation and are resistant to LPS-induced lethality. *Eur J Immunol* 36, 3007–3016.
 70. Voss S, Welte S, Fotin-Mleczek M, Fischer R, Ulmer AJ, Jung G, Wiesmuller KH, Brock R. (2006). A CD14 domain with lipopolysaccharide-binding and -neutralizing activity. *ChemBioChem* 7, 275–286.
 71. Tartakovsky B, Sredni B, Zigman-Hoffman E, Senyor G, Naparstek E. (2012). A peptide of CD14 protects human lymphocytes from gliotoxin-induced apoptosis. *Int J Pept Res Ther* 18, 249–258.
 72. Seiffert M, Schulz A, Ohl S, Dohner H, Stilgenbauer S, Lichter P. (2010). Soluble CD14 is a novel monocyte-derived survival factor for chronic lymphocytic leukemia cells, which is induced by CLL cells in vitro and present at abnormally high levels in vivo. *Blood* 116, 4223–4230.
 73. Penn AH, Altshuler AE, Small JW, Taylor SF, Dobkins KR, Schmid-Schonbein GW. (2012). Digested formula but not digested fresh human milk causes death of intestinal cells in vitro: implications for necrotizing enterocolitis. *Pediatr Res* 72, 560–567.

74. Weltzin R, Lucia-Jandris P, Michetti P, Fields BN, Kraehenbuhl JP, Neutra MR. (1989). Binding and transepithelial transport of immunoglobulins by intestinal M cells: demonstration using monoclonal IgA antibodies against enteric viral proteins. *J Cell Biol* 108, 1673–1685.
75. Aoyama K, Koyama M, Matsuoka K, Hashimoto D, Ichinohe T, Harada M, Akashi K, Tanimoto M, Teshima T. (2009). Improved outcome of allogeneic bone marrow transplantation due to breastfeeding-induced tolerance to maternal antigens. *Blood* 113, 1829–1833.
76. Tanimura N, Saitoh S, Matsumoto F, Akashi-Takamura S, Miyake K. (2008). Roles for LPS-dependent interaction and relocation of TLR4 and TRAM in TRIF-signaling. *Biochem Biophys Res Commun* 368, 94–99.
77. Neal MD, Leaphart C, Levy R, Prince J, Billiar TR, Watkins S, Li J, Cetin S, Ford H, Schreiber A, Hackam DJ. (2006). Enterocyte TLR4 mediates phagocytosis and translocation of bacteria across the intestinal barrier. *J Immunol* 176, 3070–3079.
78. Lundell AC, Andersson K, Josefsson E, Steinkasserer A, Rudin A. (2007). Soluble CD14 and CD83 from human neonatal antigen-presenting cells are inducible by commensal bacteria and suppress allergen-induced human neonatal Th2 differentiation. *Infect Immun* 75, 4097–4104.
79. Jones CA, Holloway JA, Popplewell EJ, Diaper ND, Holloway JW, Vance GHS, Warner JA, Warner JO. (2002). Reduced soluble CD14 levels in amniotic fluid and breast milk are associated with the subsequent development of atopy, eczema, or both. *J Allergy Clin Immunol* 109, 858–866.
80. Savilahti E, Siltanen M, Kajosaari M, Vaarala O, Saarinen KM. (2005). IgA antibodies, TGF-beta1 and -beta2, and soluble CD14 in the colostrum and development of atopy by age 4. *Pediatr Res* 58, 1300–1305.
81. Ismail IH, Licciardi PV, Oppedisano F, Boyle RJ, Tang ML. (2013). Relationship between breast milk sCD14, TGF-beta1 and total IgA in the first month and development of eczema during infancy. *Pediatr Allergy Immunol* 24, 352–360.
82. Maayan-Metzger A, Itzhak A, Mazkereth R, Kuint J. (2004). Necrotizing enterocolitis in full-term infants: case-control study and review of the literature. *J Perinatol* 24, 494–499.
83. Sullivan S, Schanler RJ, Kim JH, Patel AL, Trawoger R, Kiechl-Kohlendorfer U, Chan GM, Blanco CL, Abrams S, Cotten CM, Laroia N, Ehrenkranz RA, Dudell G, Cristofalo EA, Meier P, Lee ML, Rechtman DJ, Lucas A. (2010). An exclusively human milk-based diet is associated with a lower rate of necrotizing enterocolitis than a diet of human milk and bovine milk-based products. *J Pediatr* 156, 562–567e1.
84. Blais DR, Vascotto SG, Griffith M, Altosaar I. (2005). LBP and CD14 secreted in tears by the lacrimal glands modulate the LPS response of corneal epithelial cells. *Invest Ophthalmol Vis Sci* 46, 4235–4244.
85. Wright SD, Ramos RA, Tobias PS, Ulevitch RJ, Mathison JC. (1990). CD14, a receptor for complexes of lipopolysaccharide (LPS) and LPS binding protein. *Science* 249, 1431–1433.
86. Kagan JC, Medzhitov R. (2006). Phosphoinositide-mediated adaptor recruitment controls Toll-like receptor signaling. *Cell* 125, 943–955.
87. Kagan JC, Su T, Horng T, Chow A, Akira S, Medzhitov R. (2008). TRAM couples endocytosis of Toll-like receptor 4 to the induction of interferon-beta. *Nat Immunol* 9, 361–368.
88. Deng M, Scott MJ, Loughran P, Gibson G, Sodhi C, Watkins S, Hackam D, Billiar TR. (2013). Lipopolysaccharide clearance, bacterial clearance, and systemic inflammatory

- responses are regulated by cell type-specific functions of TLR4 during sepsis. *J Immunol* 190, 5152–5160.
89. Haziot A, Hijiya N, Gangloff SC, Silver J, Goyert SM. (2001). Induction of a novel mechanism of accelerated bacterial clearance by lipopolysaccharide in CD14-deficient and Toll-like receptor 4-deficient mice. *J Immunol* 166, 1075–1078.
 90. Ward TL, Spencer WJ, Davis LD, Harrold J, Mack DR, Altosaar I. (2014). Ingested soluble CD14 from milk is transferred intact into the blood of newborn rats. *Pediatr Res* 75, 252–258.
 91. Zaroni I, Ostuni R, Marek LR, Barresi S, Barbalat R, Barton GM, Granucci F, Kagan JC. (2011). CD14 controls the LPS-induced endocytosis of Toll-like receptor 4. *Cell* 147, 868–880.
 92. Sanchez-Munoz F, Fonseca-Camarillo G, Villeda-Ramirez MA, Miranda-Perez E, Mendivil EJ, Barreto-Zuniga R, Uribe M, Bojalil R, Dominguez-Lopez A, Yamamoto-Furusho JK. (2011). Transcript levels of Toll-Like Receptors 5, 8 and 9 correlate with inflammatory activity in Ulcerative Colitis. *BMC Gastroenterol* 11, 138.
 93. Picariello G, Iacomino G, Mamone G, Ferranti P, Fierro O, Gianfrani C, Di Luccia A, Addeo F. (2013). Transport across Caco-2 monolayers of peptides arising from in vitro digestion of bovine milk proteins. *Food Chem* 139, 203–212.
 94. Martin-Latil S, Gnadig NF, Mallet A, Desdouts M, Guivel-Benhassine F, Jeannin P, Prevost MC, Schwartz O, Gessain A, Ozden S, Ceccaldi PE. (2012). Transcytosis of HTLV-1 across a tight human epithelial barrier and infection of subepithelial dendritic cells. *Blood* 120, 572–580.
 95. Giri CP, Shima K, Tall BD, Curtis S, Sathyamoorthy V, Hanisch B, Kim KS, Kopecko DJ. (2012). *Cronobacter* spp. (previously *Enterobacter sakazakii*) invade and translocate across both cultured human intestinal epithelial cells and human brain microvascular endothelial cells. *Microb Pathog* 52, 140–147.
 96. Ward T, Hosid S, Ioshikhes I, Altosaar I. (2013). Human milk metagenome: a functional capacity analysis. *BMC Microbiol* 13, 116.
 97. Norberg S, Stanton C, Ross RP, Hill C, Fitzgerald GF, Cotter PD. (2012). *Cronobacter* spp. in powdered infant formula. *J Food Prot* 75, 607–620.
 98. He W, Ladinsky MS, Huey-Tubman KE, Jensen GJ, McIntosh JR, Björkman PJ. (2008). FcRn-mediated antibody transport across epithelial cells revealed by electron tomography. *Nature* 455, 542–546.
 99. Haziot A, Ferrero E, Kontgen F, Hijiya N, Yamamoto S, Silver J, Stewart CL, Goyert SM. (1996). Resistance to endotoxin shock and reduced dissemination of gram-negative bacteria in CD14-deficient mice. *Immunity* 4, 407–414.
 100. Guo S, Al-Sadi R, Said HM, Ma TY. (2013). Lipopolysaccharide causes an increase in intestinal tight junction permeability in vitro and in vivo by inducing enterocyte membrane expression and localization of TLR-4 and CD14. *Am J Pathol* 182, 375–387.
 101. Wang SC, Klein RD, Wahl WL, Alarcon WH, Garg RJ, Remick DG, Su GL. (1998). Tissue coexpression of LBP and CD14 mRNA in a mouse model of sepsis. *J Surg Res* 76, 67–73.
 102. Tomita M, Ohkubo R, Hayashi M. (2004). Lipopolysaccharide transport system across colonic epithelial cells in normal and infective rat. *Drug Metab Pharmacokinet* 19, 33–40.
 103. Hornef MW, Frisan T, Vandewalle A, Normark S, Richter-Dahlfors A. (2002). Toll-like receptor 4 resides in the Golgi apparatus and colocalizes with internalized lipopolysaccharide in intestinal epithelial cells. *J Exp Med* 195, 559–570.
 104. Leaphart CL, Cavallo J, Gripar SC, Cetin S, Li J, Branca MF, Dubowski TD, Sodhi CP, Hackam DJ. (2007). A critical role for TLR4 in the pathogenesis of necrotizing enterocolitis by modulating intestinal injury and repair. *J Immunol* 179, 4808–4820.

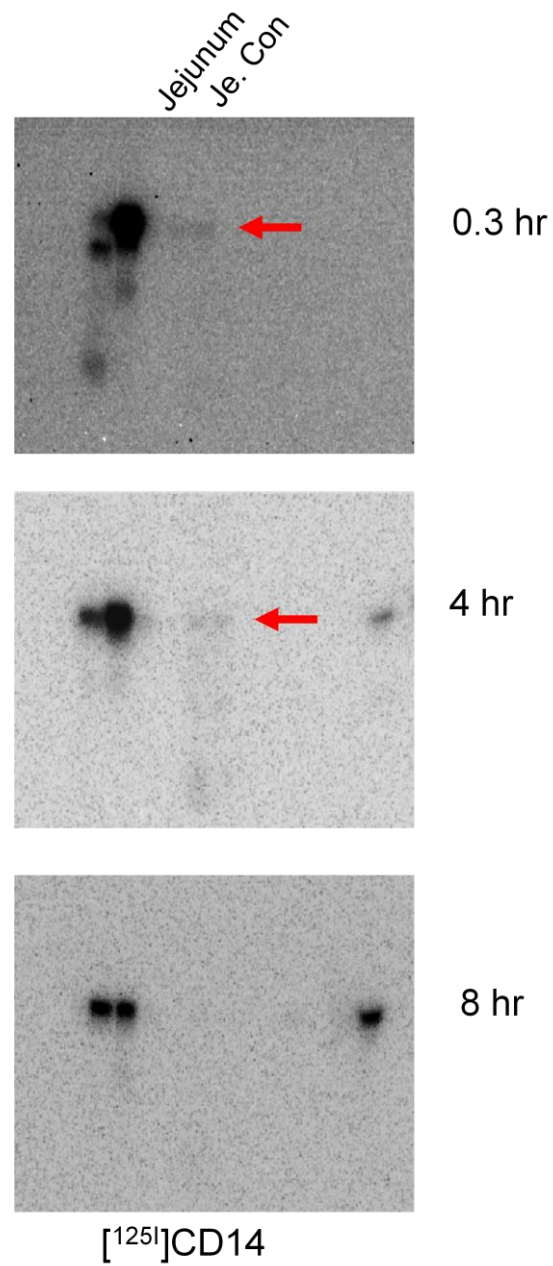
105. Nanthakumar N, Meng D, Goldstein AM, Zhu W, Lu L, Uauy R, Llanos A, Claud EC, Walker WA. (2011). The mechanism of excessive intestinal inflammation in necrotizing enterocolitis: an immature innate immune response. *PLoS One* 6, e17776.
106. Wolfs TG, Derikx JP, Hodin CM, Vanderlocht J, Driessen A, de Bruine AP, Bevins CL, Lasitschka F, Gassler N, van Gemert WG, Buurman WA. (2010). Localization of the lipopolysaccharide recognition complex in the human healthy and inflamed premature and adult gut. *Inflamm Bowel Dis* 16, 68–75.
107. Hackam DJ, Afrazi A, Good M, Sodhi CP. (2013). Innate immune signaling in the pathogenesis of necrotizing enterocolitis. *Clin Dev Immunol* 2013, 475415.
108. Neu J, Walker WA. (2011). Necrotizing enterocolitis. *N Engl J Med* 364, 255–64.
109. Fitzgibbons SC, Ching Y, Yu D, Carpenter J, Kenny M, Weldon C, Lillehei C, Valim C, Horbar JD, Jaksic T. (2009). Mortality of necrotizing enterocolitis expressed by birth weight categories. *J Pediatr Surg* 44, 1072–1075.
110. Neal MD, Sodhi CP, Dyer M, Craig BT, Good M, Jia H, Yazji I, Afrazi A, Richardson WM, Beer-Stolz D, Ma C, Prindle T, Grant Z, Branca MF, Ozolek J, Hackam DJ. (2013). A critical role for TLR4 induction of autophagy in the regulation of enterocyte migration and the pathogenesis of necrotizing enterocolitis. *J Immunol* 190, 3541–3551.
111. Yazji I, Sodhi CP, Lee EK, Good M, Egan CE, Afrazi A, Neal MD, Jia H, Lin J, Ma C, Branca MF, Prindle T, Richardson WM, Ozolek J, Billiar TR, Binion DG, Gladwin MT, Hackam DJ. (2013). Endothelial TLR4 activation impairs intestinal microcirculatory perfusion in necrotizing enterocolitis via eNOS-NO-nitrite signaling. *Proc Natl Acad Sci USA* 110, 9451–9456.
112. Chan KL, Wong KF, Luk JM. (2009). Role of LPS/CD14/TLR4-mediated inflammation in necrotizing enterocolitis: pathogenesis and therapeutic implications. *World J Gastroenterol* 15, 4745–4752.
113. Haziot A, Rong GW, Lin XY, Silver J, Goyert SM. (1995). Recombinant soluble CD14 prevents mortality in mice treated with endotoxin (lipopolysaccharide). *J Immunol* 154, 6529–6532.
114. Lee JW, Paape MJ, Zhao X. (2003). Recombinant bovine soluble CD14 reduces severity of experimental *Escherichia coli* mastitis in mice. *Vet Res* 34, 307–316.
115. Kramer MS, Guo T, Platt RW, Sevkovskaya Z, Dzikovich I, Collet JP, Shapiro S, Chalmers B, Hodnett E, Vanilovich I, Mezen I, Ducruet T, Shishko G, Bogdanovich N. (2003). Infant growth and health outcomes associated with 3 compared with 6 mo of exclusive breastfeeding. *Am J Clin Nutr* 78, 291–295.
116. Ladomenou F, Moschandreas J, Kafatos A, Tselentis Y, Galanakis E. (2010). Protective effect of exclusive breastfeeding against infections during infancy: a prospective study. *Arch Child* 95, 1004–1008.
117. Meinzen-Derr J, Poindexter B, Wrage L, Morrow AL, Stoll B, Donovan EF. (2009). Role of human milk in extremely low birth weight infants' risk of necrotizing enterocolitis or death. *J Perinatol* 29, 57–62.
118. Sangild PT, Siggers RH, Schmidt M, Elnif J, Bjornvad CR, Thymann T, Grondahl ML, Hansen AK, Jensen SK, Boye M, Moelbak L, Buddington RK, Westrom BR, Holst JJ, Burrin DG. (2006). Diet- and colonization-dependent intestinal dysfunction predisposes to necrotizing enterocolitis in preterm pigs. *Gastroenterol* 130, 1776–1792.
119. Sodhi C, Richardson W, Gribar S, Hackam DJ. (2008). The development of animal models for the study of necrotizing enterocolitis. *Model Mech* 1, 94–98.
120. Harmsen HJ, Wildeboer-Veloo AC, Raangs GC, Wagendorp AA, Klijn N, Bindels JG, Welling GW. (2000). Analysis of intestinal flora development in breast-fed and formula-fed infants by using molecular identification and detection methods. *J Pediatr Gastroenterol Nutr* 30, 61–67.

121. Sakata S, Tonooka T, Ishizeki S, Takada M, Sakamoto M, Fukuyama M, Benno Y. (2005). Culture-independent analysis of fecal microbiota in infants, with special reference to *Bifidobacterium* species. *FEMS Microbiol Lett* 243, 417–423.
122. Clemente JC, Ursell LK, Parfrey LW, Knight R. (2012). The impact of the gut microbiota on human health: an integrative view. *Cell* 148, 1258–1270.
123. Dalpke A, Frank J, Peter M, Heeg K. (2006). Activation of Toll-like receptor 9 by DNA from different bacterial species. *Infect Immun* 74, 940–946.
124. Gursel I, Gursel M, Yamada H, Ishii KJ, Takeshita F, Klinman DM. (2003). Repetitive elements in mammalian telomeres suppress bacterial DNA-induced immune activation. *J Immunol* 171, 1393–1400.
125. Bouladoux N, Hall JA, Grainger JR, dos Santos LM, Kann MG, Nagarajan V, Verthelyi D, Belkaid Y. (2012). Regulatory role of suppressive motifs from commensal DNA. *Mucosal Immunol* 5, 623–634.
126. Lin PW, Nasr TR, Stoll BJ. (2008). Necrotizing enterocolitis: recent scientific advances in pathophysiology and prevention. *Semin Perinatol* 32, 70–82.
127. Heikkilä MP, Saris PEJ. (2003). Inhibition of *Staphylococcus aureus* by the commensal bacteria of human milk. *J Appl Microbiol* 95, 471–478.
128. Martin R, Heilig HG, Zoetendal EG, Jimenez E, Fernandez L, Smidt H, Rodriguez JM. (2007). Cultivation-independent assessment of the bacterial diversity of breast milk among healthy women. *Res Microbiol* 158, 31–37.
129. Collado MC, Delgado S, Maldonado A, Rodriguez J. (2009). Assessment of the bacterial diversity of breast milk of healthy women by quantitative real-time PCR. *Lett Appl Microbiol* 48, 523–528.
130. Martin V, Maldonado-Barragan A, Moles L, Rodriguez-Banos M, Campo RD, Fernandez L, Rodriguez JM, Jimenez E. (2012). Sharing of bacterial strains between breast milk and infant feces. *J Hum Lact* 28, 36–44.
131. Meyer F, Paarmann D, D'Souza M, Olson R, Glass EM, Kubal M, Paczian T, Rodriguez A, Stevens R, Wilke A, Wilkening J, Edwards RA. (2008). The metagenomics RAST server - a public resource for the automatic phylogenetic and functional analysis of metagenomes. *BMC Bioinformatics* 9, 386.
132. Wilson K. (2001). Current protocols in molecular biology. Wiley, New York, NY, USA.
133. Boisvert S, Laviolette F, Corbeil J. (2010). Ray: simultaneous assembly of reads from a mix of high-throughput sequencing technologies. *J Comput Biol* 17, 1519–1533.
134. Everard A, Belzer C, Geurts L, Ouwerkerk JP, Druart C, Bindels LB, Guiot Y, Derrien M, Muccioli GG, Delzenne NM, de Vos WM, Cani PD. (2013). Cross-talk between *Akkermansia muciniphila* and intestinal epithelium controls diet-induced obesity. *Proc Natl Acad Sci* 110, 9066–9071.
135. Hartmann G, Krieg AM. (2000). Mechanism and function of a newly identified CpG DNA motif in human primary B cells. *J Immunol* 164, 944–953.
136. Stunz LL, Lenert P, Peckham D, Yi AK, Haxhinasto S, Chang M, Krieg AM, Ashman RF. (2002). Inhibitory oligonucleotides specifically block effects of stimulatory CpG oligonucleotides in B cells. *Eur J Immunol* 32, 1212–1222.
137. Peter M, Bode K, Lipford GB, Eberle F, Heeg K, Dalpke AH. (2008). Characterization of suppressive oligodeoxynucleotides that inhibit Toll-like receptor-9-mediated activation of innate immunity. *Immunology* 123, 118–128.
138. Ashman RF, Goeken JA, Latz E, Lenert P. (2011). Optimal oligonucleotide sequences for TLR9 inhibitory activity in human cells: lack of correlation with TLR9 binding. *Int Immunol* 23, 203–214.

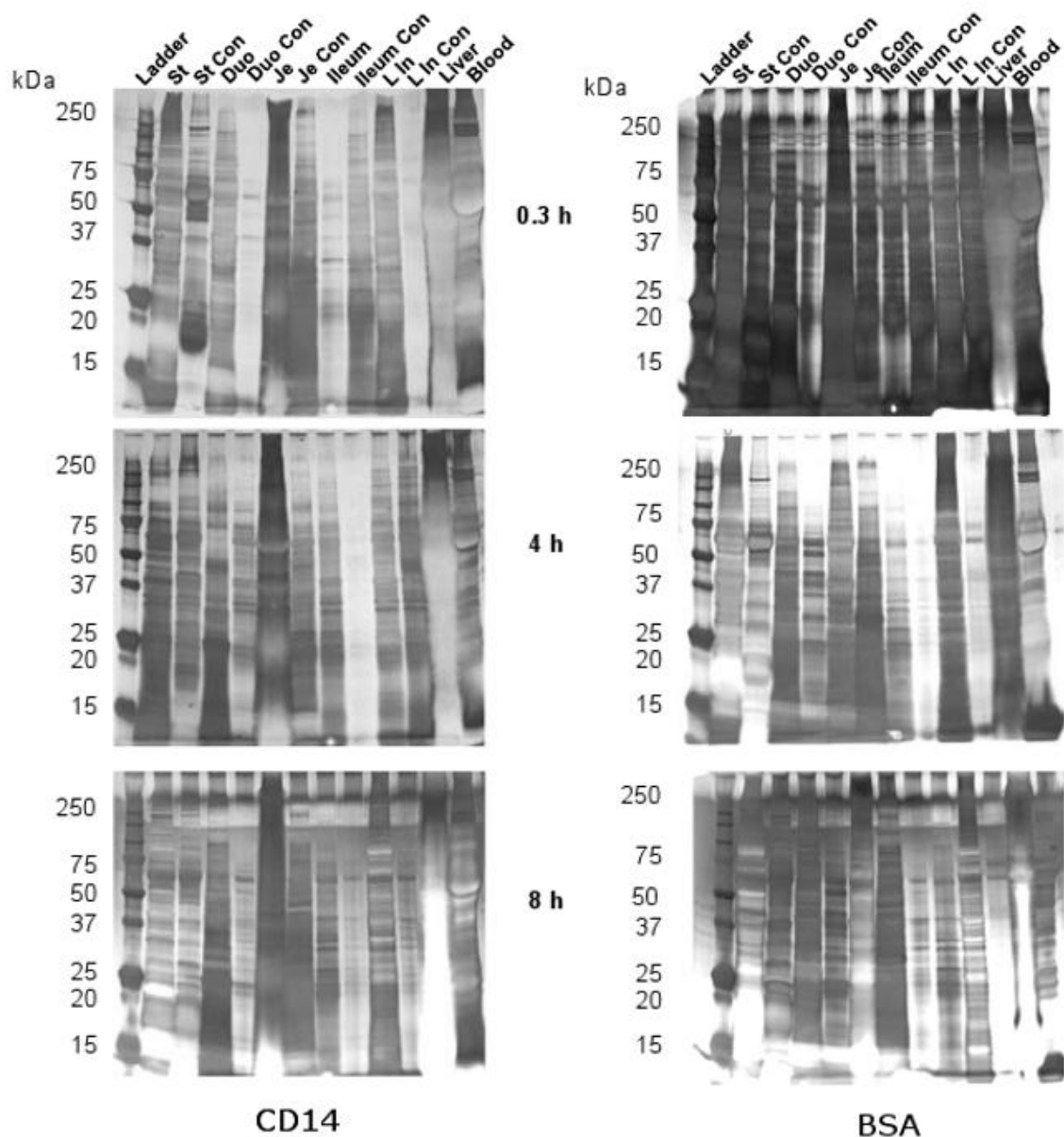
139. Zhang X, Gao M, Ha T, Kalbfleisch JH, Williams DL, Li C, Kao RL. (2012). The toll-like receptor 9 agonist, CpG-oligodeoxynucleotide 1826, ameliorates cardiac dysfunction after trauma-hemorrhage. *Shock* 38, 146–152.
140. Huttenhower C, Gevers D, Knight R, Abubucker A, Badger JH, Chinwalla AT, Creasy HH, Earl AM, FitzGerald MG, Fulton RS, Giglio MG, Hallworth-Pepin K. (2012). Structure, function and diversity of the healthy human microbiome. *Nature* 486, 207–214.
141. De Boer R, Peters R, Gierveld S, Schuurman T, Kooistra-Smid M, Savelkoul P. (2010). Improved detection of microbial DNA after bead-beating before DNA isolation. *J Microbiol Methods* 80, 209–211.
142. Yatsunenko T, Rey FE, Manary MJ, Trehan I, Dominguez-Bello MG, Contreras M, Magris M, Hidalgo G, Baldassano RN, Anokhin AP, Heath AC, Warner B, Reeder J, Kuczynski J, Caporaso JG, Lozupone CA, Lauber C, Clemente JC, Knights D, Knight R, Gordon JI. (2012). Human gut microbiome viewed across age and geography. *Nature* 486, 222–227.
143. Grice EA, Kong HH, Conlan S, Deming CB, Davis J, Young AC, Bouffard GG, Blakesley RW, Murray PR, Green ED, Turner ML, Segre JA. (2009). Topographical and temporal diversity of the human skin microbiome. *Science* 324, 1190–1192.
144. Costello EK, Lauber CL, Hamady M, Fierer N, Gordon JI, Knight R. (2009). Bacterial community variation in human body habitats across space and time. *Science* 326, 1694–1697.
145. Oh J, Conlan S, Polley EC, Segre JA, Kong HH. (2012). Shifts in human skin and nares microbiota of healthy children and adults. *Genome Med* 4, 77.
146. Dominguez-Bello MG, Costello EK, Contreras M, Magris M, Hidalgo G, Fierer N, Knight R. (2010). Delivery mode shapes the acquisition and structure of the initial microbiota across multiple body habitats in newborns. *Proc Natl Acad Sci USA* 107, 11971–11975.
147. Wharton BA, Balmer SE, Scott PH. (1994). Sorrento studies of diet and fecal flora in the newborn. *Acta Paediatr Jpn* 36, 579–584.
148. Perez PF, Dore J, Leclerc M, Levenez F, Benyacoub J, Serrant P, Segura-Roggero I, Schiffrin EJ, Donnet-Hughes A. (2007). Bacterial imprinting of the neonatal immune system: lessons from maternal cells? *Pediatrics* 119, e724–732.
149. Donnet-Hughes A, Perez PF, Dore J, Leclerc M, Levenez F, Benyacoub J, Serrant P, Segura-Roggero I, Schiffrin EJ. (2010). Potential role of the intestinal microbiota of the mother in neonatal immune education. *Proc Nutr Soc* 69, 407–415.
150. Rescigno M, Rotta G, Valzasina B, Ricciardi-Castagnoli P. (2001). Dendritic cells shuttle microbes across gut epithelial monolayers. *Immunobiology* 204, 572–581.
151. Engfer MB, Stahl B, Finke B, Sawatzki G, Daniel H. (2000). Human milk oligosaccharides are resistant to enzymatic hydrolysis in the upper gastrointestinal tract. *Am J Clin Nutr* 71, 1589–1596.
152. Zivkovic AM, German JB, Lebrilla CB, Mills DA. (2011). Human milk glycobiome and its impact on the infant gastrointestinal microbiota. *Proc Natl Acad Sci USA* 108 (Suppl 1), 4653–4658.
153. Hunt KM, Preuss J, Nissan C, Davlin CA, Williams JE, Shafii B, Richardson AD, McGuire MK, Bode L, McGuire MA. (2012). Human milk oligosaccharides promote the growth of staphylococci. *Appl Env. Microbiol* 78, 4763–4770.
154. Corvaglia L, Battistini B, Paoletti V, Aceti A, Capretti MG, Faldella G. (2008). Near-infrared reflectance analysis to evaluate the nitrogen and fat content of human milk in neonatal intensive care units. *Arch Child Fetal Neonatal Ed* 93, F372–375.
155. Lepage P, Van de Perre P. (2012). The immune system of breast milk: antimicrobial and anti-inflammatory properties. *Adv Exp Med Biol* 743, 121–137.

156. Urbaniak C, Burton JP, Reid G. (2012). Breast, milk and microbes: a complex relationship that does not end with lactation. *Womens Health Lond Engl* 8, 385–398.
157. Delgado S, Garcia P, Fernandez L, Jimenez E, Rodriguez-Banos M, del Campo R, Rodriguez JM. (2011). Characterization of *Staphylococcus aureus* strains involved in human and bovine mastitis. *FEMS Immunol Med Microbiol* 62, 225–235.
158. Espinosa-Martos I, Montilla A, Segura AG, Escuder D, Bustos G, Pallas C, Rodriguez JM, Corzo N, Fernandez L. (2013). Bacteriological, biochemical and immunological modifications in human colostrum after Holder pasteurisation. *J Pediatr Gastroenterol Nutr* 56, 560–568.
159. Goldman AS. (1993). The immune system of human milk: antimicrobial, antiinflammatory and immunomodulating properties. *Pediatr Infect J* 12, 664–671.
160. Jeurink PV, van Bergenhenegouwen J, Jimenez E, Knippels LM, Fernandez L, Garssen J, Knol J, Rodriguez JM, Martin R. (2013). Human milk: a source of more life than we imagine. *Benef Microbes* 4, 17–30.
161. Ward TL, Goto K, Altosaar I. (2013). Ingested soluble CD14 contributes to the functional pool of circulating sCD14 in mice. *Immunobiology In print*, doi:10.1016/j.imbio.2014.03.008
162. Blais DR, Altosaar I. (2006). Human CD14 expressed in seeds of transgenic tobacco displays similar proteolytic resistance and bioactivity with its mammalian-produced counterpart. *Transgenic Res* 15, 151–164.
163. Christou P. (2013). From medicinal plants to medicines in plants: plant factories for the production of valuable pharmaceuticals. *Curr Pharm Des* 19, 5469–5470.
164. Sabalza M, Vamvaka E, Christou P, Capell T. (2013). Seeds as a production system for molecular pharming applications: status and prospects. *Curr Pharm Des* 19, 5543–5552.
165. Greenham T, Altosaar I. (2013). Molecular strategies to engineer transgenic rice seed compartments for large-scale production of plant-made pharmaceuticals. *Methods Mol Biol* 956, 311–326.
166. Blais DR, Altosaar I. (2007). Humanizing infant milk formula to decrease postnatal HIV transmission. *Trends Biotechnol* 25, 376–384.
167. He Y, Ning T, Xie T, Qiu Q, Zhang L, Sun Y, Jiang D, Fu K, Yin F, Zhang W, Shen L, Wang H, Li J, Lin Q, Sun Y, Li H, Zhu Y, Yang D. (2011). Large-scale production of functional human serum albumin from transgenic rice seeds. *Proc Natl Acad Sci USA* 108, 19078–19083.
168. Braegger C, Chmielewska A, Decsi T, Kolacek S, Mihatsch W, Moreno L, Pieścik M, Puntis J, Shamir R, Szajewska H, Turck D, van Goudoever J. (2011). Supplementation of infant formula with probiotics and/or prebiotics: a systematic review and comment by the ESPGHAN committee on nutrition. *J Pediatr Gastroenterol Nutr* 52, 238–250.

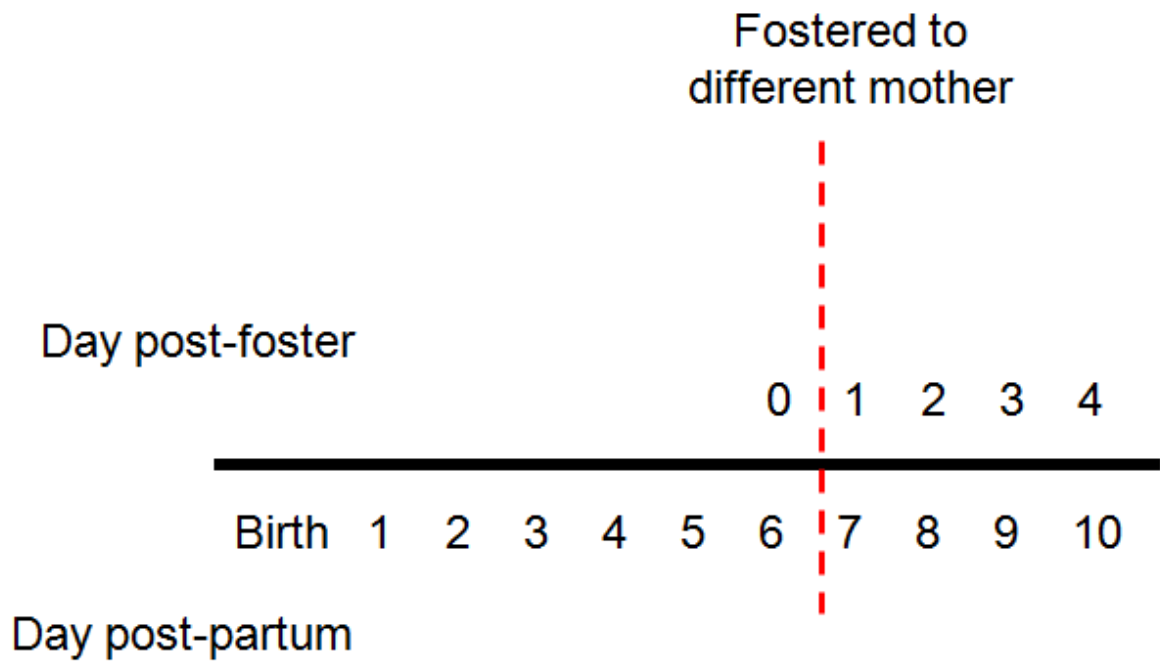
Appendix A. Supplemental Figures



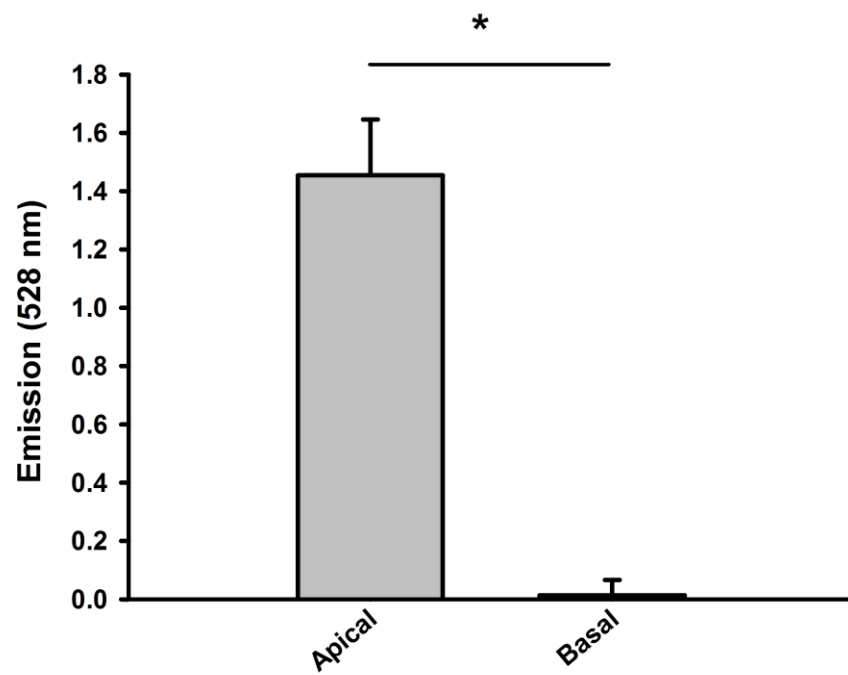
Supplemental Figure S2.1. Increased contrast of $[^{125}\text{I}]\text{CD14}$ phosphor-images used in Figure 2.2A. Red arrows point to intact $[^{125}\text{I}]\text{CD14}$ within the jejunum and its contents (Je. Con).



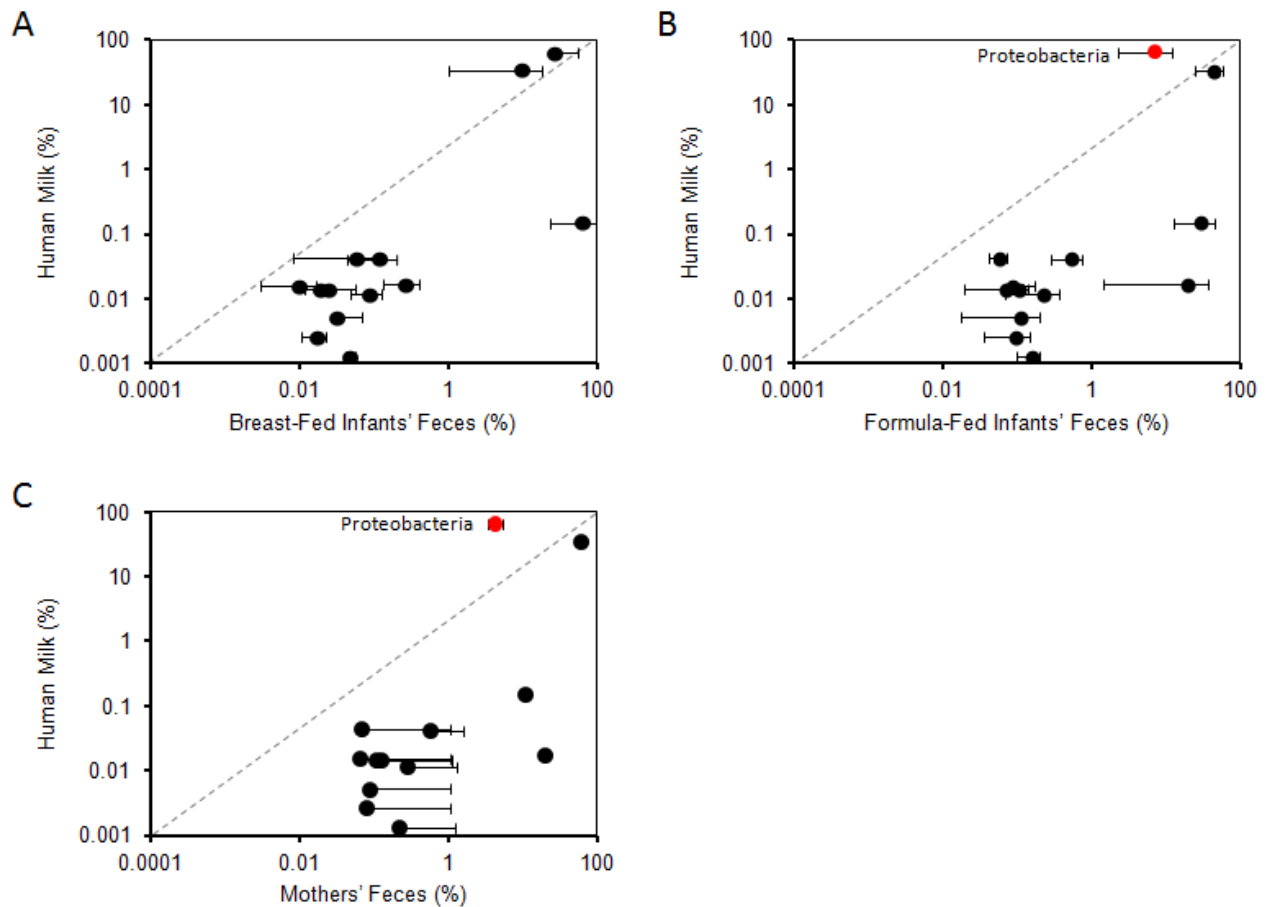
Supplemental Figure S2.2. Proteins isolated from the gastrointestinal tract and organs of rat pups fed $[^{125}\text{I}]\text{CD14}$ or $[^{125}\text{I}]\text{BSA}$. The same gels were used to produce the phosphor-images in Figure 2.2. Rat pups were gavage fed 25 μg of $[^{125}\text{I}]\text{sCD14}$ (12.16×10^6 dpm/ μg) or $[^{125}\text{I}]\text{BSA}$ (13.7×10^6 dpm/ μg). Post-euthanasia (0.3, 4 or 8 h), proteins were extracted from harvested organs, size separated by SDS-PAGE and silver stained.



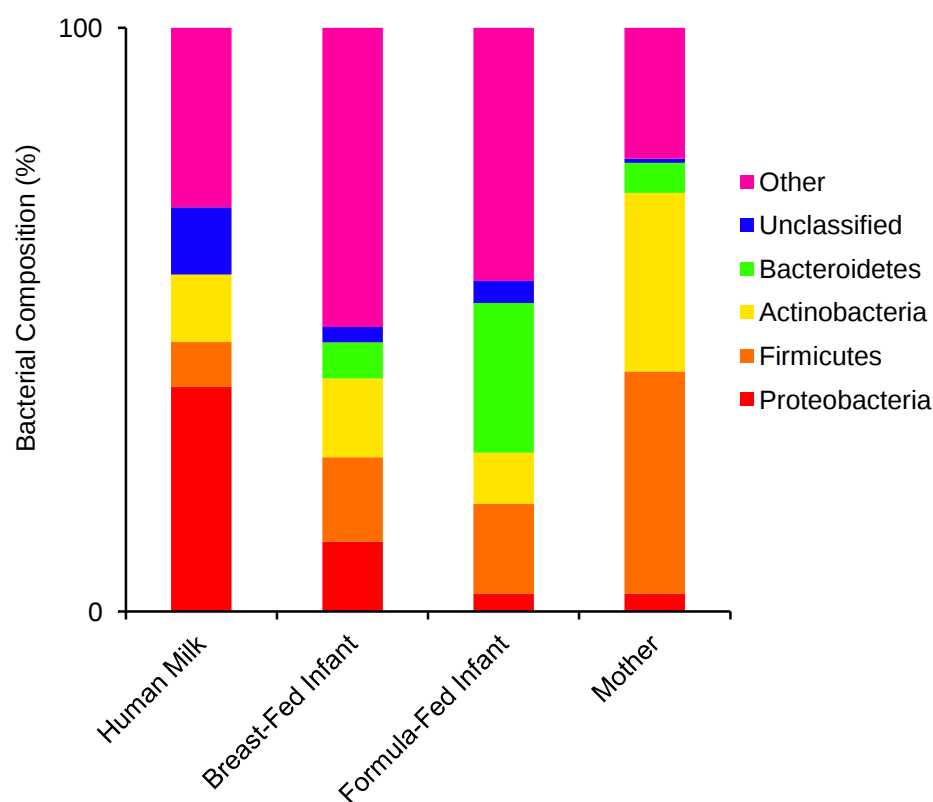
Supplemental Figure S3.1. Timeline of pup fostering. Pups of WT or CD14^{-/-} genotype were born and subsequently fostered to different mothers of either wild type or CD14^{-/-} genotype on day 6 PP (red dotted line).



Supplemental Figure S3.2. Passive transport of Lucifer Yellow across Caco-2 cell monolayers. Lucifer Yellow was added to the apical media and incubated 2 hours at 37°C.



Supplemental Figure S4.1. Pair-wise comparison of phyla abundance in human milk versus infants' and mothers' fecal metagenomes. Pair-wise comparisons for the human milk metagenome versus (A) breast-fed infants' feces, (B) formula-fed infants' feces and (C) mothers' feces are shown. Each point represents a different phylum and its relative abundance within the human milk metagenome compared to the fecal metagenomes. Points lying on or near the dotted line have equal or similar abundance in both metagenomes. Red dots signify those with significantly different proportions between the two metagenomes (Student's *t*-test, $P < 0.05$). Breast-fed and formula-fed infant feces values are an average of five individuals, and mothers' feces values are an average of three individuals. All subjects are unrelated.



Supplemental Figure S4.2. Lowest common ancestor comparison of bacterial phyla in human milk, infants' and mothers' feces. Contigs within each metagenome were assigned to a phylum within MG-RAST (maximum e-value of 1×10^{-5} , minimum identity of 60%, and minimum alignment length of 45 bp). Breast fed and formula-fed infant feces values are an average of five samples, and mothers' feces values are an average of three samples. All subjects are unrelated. Other contains phyla each representing <1% of the contigs.

Appendix B. Supplemental Tables

Supplemental Table S4.1. Abundance of DNA fragments in pooled human milk, sequenced seven times. Sequences of 51 bp were analyzed by Illumina sequencing and matched to human or prokaryotic genomes (≤ 2 bp mismatch) by BLAT.

Sequences per run				
	Total	Human	Prokaryotic	Other
	18,193,400	14,026,621	117,211	4,049,568
	35,809,228	26,466,812	211,199	9,131,217
	50,313,515	34,690,538	283,882	15,339,095
	39,021,577	27,432,706	123,626	11,465,245
	47,856,320	32,191,169	142,279	15,522,872
	27,612,520	20,864,809	188,516	6,559,195
	42,725,644	30,338,333	265,283	12,122,028
Total	261,532,204	186,010,988	1,331,996	74,189,220
Fraction of total $\pm$ SE		(72.01 $\pm$ 3.06%)	(0.53 $\pm$ 0.16%)	(27.46 $\pm$ 3.72%)

Supplemental Table S4.2. Classification of 51 bp DNA sequences derived from human milk by best hit analysis. Each match was characterized with a 95% sequence alignment with the known prokaryotic genome.

Genus	Percent of Sequences
Staphylococcus	74.96
Pseudomonas	14.74
Edwardsiella	2.34
Pantoea	1.43
Treponema	1.22
Streptococcus	1.07
Campylobacter	0.90
Corynebacterium	0.25
Thermoanaerobacter	0.23
Mycoplasma	0.22
Lactobacillus	0.20
Propionibacterium	0.19
Escherichia	0.11
Candidatus	0.11
Finegoldia	0.10
Riemerella	0.09
Bacillus	0.08
Stenotrophomonas	0.08
Erwinia	0.08
Sphingopyxis	0.08
Caldicellulosiruptor	0.07
Yersinia	0.06
Burkholderia	0.06
Acinetobacter	0.05
Listeria	0.04
Mycobacterium	0.04
Shewanella	0.04
Klebsiella	0.04
Brachyspira	0.04
Azotobacter	0.04
Agrobacterium	0.04
Salmonella	0.04
Enterococcus	0.04
Enterobacter	0.03
Verminephrobacter	0.03
Lactococcus	0.03

Clostridium	0.03
Thermoanaerobacterium	0.03
Pediococcus	0.03
Buchnera	0.03
Macrococcus	0.03
Methanococcus	0.02
Aeromonas	0.02
Shigella	0.02
Bacteroides	0.02
Cronobacter	0.02
Ralstonia	0.02
Borrelia	0.02
Desulfotomaculum	0.02
Helicobacter	0.02
Rothia	0.01
Xanthomonas	0.01
Serratia	0.01
Acidovorax	0.01
Haemophilus	0.01
Methanosarcina	0.01
Neisseria	0.01
Citrobacter	0.01
Anaerococcus	0.01
Legionella	0.01
Methylobacterium	0.01
Dickeya	0.01
Nitrosomonas	0.01
Geobacillus	0.01
Veillonella	0.01
Methanobrevibacter	0.01
Pectobacterium	0.01
Rhodothermus	0.01
Marinobacter	0.01
Rubrobacter	0.01
Syntrophomonas	0.01
Delftia	0.01
Leptospira	0.01
Chitinophaga	0.01
Synechococcus	0.01
Natrialba	4.73E-03
Sphingobium	4.65E-03
Trichodesmium	4.50E-03
Rhodococcus	4.35E-03

Micrococcus	4.28E-03
Spirochaeta	4.20E-03
Herbaspirillum	4.20E-03
Lawsonia	4.13E-03
Chlorobium	4.05E-03
Bradyrhizobium	3.98E-03
Methanosalsum	3.90E-03
Lysinibacillus	3.83E-03
Flavobacterium	3.83E-03
Leuconostoc	3.75E-03
Flavobacteriaceae	3.75E-03
Cupriavidus	3.75E-03
Myxococcus	3.60E-03
Xenorhabdus	3.53E-03
Pedobacter	3.53E-03
Oenococcus	3.53E-03
Alcanivorax	3.53E-03
Bordetella	3.45E-03
Azoarcus	3.45E-03
Rhizobium	3.38E-03
Halogeometricum	3.38E-03
Rahnella	3.30E-03
Polaromonas	3.30E-03
Ochrobactrum	3.23E-03
Geobacter	3.15E-03
Rickettsia	3.08E-03
Haliangium	3.08E-03
Brachybacterium	2.78E-03
Rhodopseudomonas	2.70E-03
Vibrio	2.55E-03
Thermodesulfobacterium	2.55E-03
Leptothrix	2.55E-03
Sphingomonas	2.48E-03
Dyadobacter	2.48E-03
Caulobacter	2.48E-03
Sinorhizobium	2.40E-03
Cytophaga	2.40E-03
Alkalilimnicola	2.33E-03
Gramella	2.25E-03
Brevundimonas	2.25E-03
Granulibacter	2.18E-03
Prevotella	2.10E-03
Gardnerella	2.10E-03

Acidiphilium	2.10E-03
Arthrobacter	2.03E-03
Actinobacillus	2.03E-03
Streptomyces	1.95E-03
Microbacterium	1.95E-03
Janthinobacterium	1.95E-03
Variovorax	1.80E-03
Pseudoalteromonas	1.80E-03
Alicyclophilus	1.80E-03
Pyrolobus	1.73E-03
Proteus	1.73E-03
Ramlibacter	1.65E-03
Cellvibrio	1.65E-03
Xylella	1.58E-03
Pseudoxanthomonas	1.58E-03
Chromobacterium	1.58E-03
Paracoccus	1.50E-03
Mobiluncus	1.50E-03
Marinomonas	1.43E-03
Laribacter	1.43E-03
Francisella	1.43E-03
Maricaulis	1.28E-03
Herminiimonas	1.28E-03
Butyrivibrio	1.28E-03
Xanthobacter	1.20E-03
Thioalkalivibrio	1.20E-03
Rhodobacter	1.20E-03
Frankia	1.20E-03
Alkaliphilus	1.20E-03
Halomonas	1.13E-03
Aggregatibacter	1.13E-03
Achromobacter	1.13E-03
Methylobacterium	1.05E-03
Dinoroseobacter	1.05E-03
Bifidobacterium	1.05E-03
Tolomonas	9.76E-04
Thauera	9.76E-04
Psychrobacter	9.76E-04
Leptotrichia	9.76E-04
Leadbetterella	9.76E-04
Pseudogulbenkiania	9.01E-04
Phenylobacterium	9.01E-04
Pelobacter	9.01E-04

Microcystis	9.01E-04
Kocuria	9.01E-04
Desulfovibrio	9.01E-04
Chelativorans	9.01E-04
Photorhabdus	8.26E-04
Ferrimonas	8.26E-04
Exiguobacterium	8.26E-04
Azospirillum	8.26E-04
Tetragenococcus	7.51E-04
Haloterrigena	7.51E-04
Halorhodospira	7.51E-04
Aerococcus	7.51E-04
Pelotomaculum	6.76E-04
Oceanobacillus	6.76E-04
Novosphingobium	6.76E-04
Fusobacterium	6.76E-04
Deinococcus	6.76E-04
Thiomonas	6.01E-04
Prosthecochloris	6.01E-04
Methanobacterium	6.01E-04
Saccharopolyspora	5.26E-04
Porphyromonas	5.26E-04
Paenibacillus	5.26E-04
Methanothermobacter	5.26E-04
Mesorhizobium	5.26E-04
Hyphomonas	5.26E-04
Desulfarculus	5.26E-04
Comamonas	5.26E-04
Collimonas	5.26E-04
Clostridiales	5.26E-04
Beijerinckia	5.26E-04
Arcanobacterium	5.26E-04
Sulfolobus	4.50E-04
Streptobacillus	4.50E-04
Prochlorococcus	4.50E-04
Pirellula	4.50E-04
Pelagibacterium	4.50E-04
Moraxella	4.50E-04
Methylibium	4.50E-04
Methanothermococcus	4.50E-04
Kytococcus	4.50E-04
Haloquadratum	4.50E-04
Gordonia	4.50E-04

Dechloromonas	4.50E-04
Clavibacter	4.50E-04
Ureaplasma	3.75E-04
Staphylothermus	3.75E-04
Sorangium	3.75E-04
Sodalis	3.75E-04
Roseobacter	3.75E-04
Rhodospirillum	3.75E-04
Methylobacillus	3.75E-04
Hahella	3.75E-04
Eubacterium	3.75E-04
Erythrobacter	3.75E-04
Ehrlichia	3.75E-04
Chromohalobacter	3.75E-04
Chloroflexus	3.75E-04
Capnocytophaga	3.75E-04
Bartonella	3.75E-04
Azorhizobium	3.75E-04
Actinosynnema	3.75E-04
Thiomicrospira	3.00E-04
Salinispora	3.00E-04
Pasteurella	3.00E-04
Parvibaculum	3.00E-04
Nocardioides	3.00E-04
Methylococcus	3.00E-04
Methanosphaera	3.00E-04
Methanopyrus	3.00E-04
Magnetococcus	3.00E-04
Jonesia	3.00E-04
Gallibacterium	3.00E-04
Catenulispora	3.00E-04
Brevibacillus	3.00E-04
Arcobacter	3.00E-04
Anoxybacillus	3.00E-04
Anabaena	3.00E-04
Alteromonas	3.00E-04
Zobellia	2.25E-04
Uncultured	2.25E-04
Thermus	2.25E-04
Thermofilum	2.25E-04
Streptosporangium	2.25E-04
Starkeya	2.25E-04
Roseburia	2.25E-04

Pusillimonas	2.25E-04
Parachlamydia	2.25E-04
Orientia	2.25E-04
Oligotropha	2.25E-04
Odoribacter	2.25E-04
Methylovorus	2.25E-04
Methanothermus	2.25E-04
Magnetospirillum	2.25E-04
Kangiella	2.25E-04
Jannaschia	2.25E-04
Hydrogenobaculum	2.25E-04
Gluconacetobacter	2.25E-04
Geodermatophilus	2.25E-04
Desulfobacterium	2.25E-04
Cellulomonas	2.25E-04
Carnobacterium	2.25E-04
Carboxydotherrnus	2.25E-04
Beutenbergia	2.25E-04
Atopobium	2.25E-04
Aromatoleum	2.25E-04
Zymomonas	1.50E-04
Verrucosispota	1.50E-04
Thermococcus	1.50E-04
Thermincola	1.50E-04
Teredinibacter	1.50E-04
Stackebrandtia	1.50E-04
Saccharophagus	1.50E-04
Ruegeria	1.50E-04
Roseiflexus	1.50E-04
Psychromonas	1.50E-04
Polynucleobacter	1.50E-04
Parabacteroides	1.50E-04
Oscillibacter	1.50E-04
Olsenella	1.50E-04
Nostoc	1.50E-04
Methylothera	1.50E-04
Methylocella	1.50E-04
Methanospirillum	1.50E-04
Megasphaera	1.50E-04
Kribbella	1.50E-04
Ketogulonicigenium	1.50E-04
Isoptericola	1.50E-04
Intrasporangium	1.50E-04

Halothiobacillus	1.50E-04
Coxiella	1.50E-04
Blattabacterium	1.50E-04
Asticcacaulis	1.50E-04
Anaeromyxobacter	1.50E-04
Allochromatium	1.50E-04
Acidaminococcus	1.50E-04
Zunongwangia	7.51E-05
Xylanimonas	7.51E-05
Weissella	7.51E-05
Weeksella	7.51E-05
Waddlia	7.51E-05
Tsukamurella	7.51E-05
Thermovibrio	7.51E-05
Thermosipho	7.51E-05
Thermobispora	7.51E-05
Tepidanaerobacter	7.51E-05
Syntrophobacter	7.51E-05
Synechocystis	7.51E-05
Sulfurospirillum	7.51E-05
Sphingobacterium	7.51E-05
Sideroxydans	7.51E-05
Segniliparus	7.51E-05
Sanguibacter	7.51E-05
Salinibacter	7.51E-05
Saccharomonospora	7.51E-05
Robiginitalea	7.51E-05
Rhodoferax	7.51E-05
Renibacterium	7.51E-05
Pyrococcus	7.51E-05
Polymorphum	7.51E-05
Parvularcula	7.51E-05
Nocardiopsis	7.51E-05
Nocardia	7.51E-05
Nitrosospira	7.51E-05
Nitrosococcus	7.51E-05
Nitrobacter	7.51E-05
Nitratifractor	7.51E-05
Nakamurella	7.51E-05
Microlunatus	7.51E-05
Methylomonas	7.51E-05
Methanosaeta	7.51E-05
Methanocaldococcus	7.51E-05

Mesoplasma	7.51E-05
Meiothermus	7.51E-05
Mannheimia	7.51E-05
Lacinutrix	7.51E-05
Krokinobacter	7.51E-05
Kosmotoga	7.51E-05
Kitasatospora	7.51E-05
Isosphaera	7.51E-05
Ilyobacter	7.51E-05
Idiomarina	7.51E-05
Halomicrobium	7.51E-05
Haliscomenobacter	7.51E-05
Fluviicola	7.51E-05
Ferroglobus	7.51E-05
Erysipelothrix	7.51E-05
Elusimicrobium	7.51E-05
Eggerthella	7.51E-05
Desulfurivibrio	7.51E-05
Desulfobacca	7.51E-05
Cyanothece	7.51E-05
Croceibacter	7.51E-05
Chlorobaculum	7.51E-05
Chlamydophila	7.51E-05
Brucella	7.51E-05
Bdellovibrio	7.51E-05
Archaeoglobus	7.51E-05
Amycolatopsis	7.51E-05
Alicyclobacillus	7.51E-05
Akkermansia	7.51E-05
Acidimicrobium	7.51E-05
Acetobacter	7.51E-05

Supplemental Table S4.3. Predicted open reading frames from human milk DNA sequences aligning to rRNA genes from known organisms. Minimum 99.5% identity and ORF length of 54 was used.

Genus	Open Reading Frames
Pseudomonas	73
Homo	40
Pantoea	30
Staphylococcus	21
Gorilla	16
Acholeplasma	8
Bacillus	8
Corynebacterium	8

Supplemental Table S4.4. Immune modulatory DNA motifs sought in DNA sequences derived from human milk or feces. Sequences were searched for in both 51 bp Illumina sequences, as well as in assembled contigs. No motifs listed here were observed in human milk contigs, whereas some were observed in contigs from breast-fed infants' feces (BF) formula-fed infants' feces (FF) and mothers' feces (MF).

Motif	Observed Hits	Immune Modification	Reference
TCCATGACGTTCTGACGTT	0	Suppressive	Zhang X, <i>et al.</i>
TCCATGACGTTCTGATGCT	0	Stimulatory	Zhang X, <i>et al.</i>
CTCCTATTGGGGGTTTCCTAT	0	Suppressive	Peter M, <i>et al.</i>
TCCTGGAGGGGAAGT	0	Suppressive	Ashman RF, <i>et al.</i>
TCCTGACGTTGAAGT	BF (4), MF (1)	Stimulatory	Ashman RF, <i>et al.</i>
TCCTGGAGGGGAAGT	0	Supressive	Ashman RF, <i>et al.</i>
TCCTGAAGGGGAAGT	0	Supressive	Ashman RF, <i>et al.</i>
TCCTGGCGGGGAAGT	MF (1)	Supressive	Ashman RF, <i>et al.</i>
TCCTGACGGGGAAGT	FF(1)	Supressive	Stunz LL, <i>et al.</i>
TCGTCGTTACGTAACGTCGTCGTT	0	Stimulatory	Stunz LL, <i>et al.</i>
TCGTCGTTACGTAACGACGTCGTT	0	Stimulatory	Stunz LL, <i>et al.</i>
TCGTCGTTCCCCCCCCCCCC	0	Stimulatory	Hartmann G, <i>et al.</i>

Appendix C. Curriculum Vitae

Tonya L Ward

PhD Candidate

Department of Biochemistry, Microbiology and Immunology

Faculty of Medicine

University of Ottawa

EDUCATION

PhD, Biochemistry

2011 – Present

University of Ottawa, Ottawa, ON

Thesis: Characterizing immune-modulatory components of human milk:
The fate and function of soluble CD14 and the human milk metagenome

Supervisor: Dr. Illimar Altosaar

BSc, Honors Specialization in Genetics

2004 – 2008

University of Western Ontario, London, ON

Thesis: Mutation frequency and pattern in the kidney of a mouse model of
elevated oxidative stress

Supervisor: Dr. Kathleen Hill

Secondary School

2000 – 2004

St. Patrick High School, Thunder Bay, ON

RESEARCH EXPERIENCE

Graduate Student Research

2009 – Present

University of Ottawa, Ottawa, ON

Department of Biochemistry, Microbiology and Immunology

Doctoral research in microbiology, innate immunity and bioinformatics in relation to human milk

- Developed expertise in protein biochemistry, cell culture, animal model development, microscopy, immunocytochemistry, radioisotope labeling, next generation sequencing, computational biology, statistical analysis, and scientific writing and communication

Undergraduate Student Research

2007 – 2008

University of Western Ontario, London, ON

Department of Biology

Honors thesis and two NSERC summer scholarships pertaining to environmental mutagens

- Developed expertise in DNA and RNA isolations, PCR, Sanger sequencing, gel electrophoresis, cloning, plaque assays, animal handling and molecular biology techniques

Lab Assistant

2005, 2006

Thunder Bay Public Health Lab, Thunder Bay Ontario
Departments of Serology and Water Testing

- Processed incoming clinical samples for various serological tests, processed drinking water to determine bacterial load

PUBLICATIONS**Peer Reviewed**

Ward TL, Goto K, Altosaar I. (2013) Ingested soluble CD14 contributes to the functional pool of circulating sCD14 in mice. *Immunobiology*. Under revision. IMBIO-D-13-00182, Submitted Oct 28, 2013.

Ward TL, Spencer WJ, Davis LDR, Harrold J, Mack DR, Altosaar I. (2013) Ingested soluble CD14 from milk is transferred intact into the blood of newborn rats. *Pediatric Research*. (Ahead of print) doi:10.1038/pr.2013.225.

Ward TL, Hosid S, Ioshikhes I, Altosaar I. (2013) Human milk metagenome: a functional capacity analysis. *BMC Microbiology*. 13:116. doi:10.1186/1471-2180-13-116.

Wan S, **Ward TL**, Altosaar I. (2012) Strategy and tactics of disarming Greenhouse Gases (GHG) at the source: N2O reductase-crops. *Trends in Biotechnology*. 30:410-415.

Zaidi M, El Bilali J, Koziol A, **Ward TL**, Styles G, Greenham T, Faiella W, Son H, Wan S, Taga I, Altosaar I. (2012) Gene technology in agriculture, environment and biopharming: Beyond Bt-rice and building better breeding budgets for crops. *Journal of Plant Biochemistry and Biotechnology*. 21(Suppl 1):2-9. doi: 10.1007/s13562-012-0128-z.

Davis LDR, Spencer WJ, Pham VT, **Ward TL**, Blais DR, Mack DR, Kaplan H, Altosaar I. (2011) ¹⁴C-Radiolabeling of proteins to monitor biodistribution of ingested proteins. *Analytical Biochemistry*. 410:57–61.

Spencer WJ, Binette A, **Ward TL**, Davis LR, Blais DR, Harrold J, Mack DR, Altosaar I. (2010) Alpha-lactalbumin in human milk alters the proteolytic degradation of soluble CD14 by forming a complex. *Pediatric Research*. 68:490-493.

Ward TL, Prtenjaca A, Hill KA. (2010) A novel *Escherichia coli*-derived mutation detected with the Big Blue cII mutant selectable assay. *Environmental and Molecular Mutagenesis*. 51:344-8.

In Preparation

Ward TL, Son H, Krishnan A, Altosaar I. Immune-modulatory potential of human milk derived DNA. *PLoS One*.

AWARDS

PhD

Teaching Assistant of the Year – Dept of Biochemistry	2013
NSERC Graduate Scholarship (CGS-D)	2012 - 2015
Leadership Award – Faculty of Medicine	2012
Ontario Graduate Scholarship	2012
Ontario Graduate Scholarship	2011

'MSc'

NSERC Graduate Scholarship (CGS-M)	2010
Ontario Graduate Scholarship	2010
Ontario Graduate Scholarship	2009
Excellence Scholarship – University of Ottawa	2009 - 2011

BSc

NSERC Undergraduate Student Research Award	2007, 2008
Dean's List	2004 - 2008
Queen Elizabeth II Aiming for the Top Scholarship	2004 - 2008
Entrance Scholarship, University of Western Ontario	2004

CONFERENCES

Canadian Society of Microbiologists 63rd Annual Meeting 2013
 Carleton University, Ottawa, ON
 Poster: **Ward TL**, Hosid S, Ioshikhes I, Altosaar I. Human milk metagenome: a functional capacity analysis.

Keystone Symposia: The Microbiome 2012
 Keystone Resort, Keystone, CO
 Poster: **Ward TL**, Hosid S, Ioshikhes I, Altosaar I. The human milk microbiome contains known and unknown bacteria.

The Ottawa Institute of Computational Biology and Bioinformatics Symposia
 University of Ottawa, Ottawa, ON 2011, 2012
 Posters: 1. **Ward TL**, Hosid S, Ioshikhes I, Altosaar I. Human milk microbiome contains an immunomodulatory landscape. (2012)
 2. **Ward TL**, Hosid S, Ioshikhes I, Altosaar I. The human milk microbiome contains known and unknown bacteria. (2011)

7th Milk Genomics and Human Health Symposium 2010
 University of California Davis, Davis, CA
 Poster: **Ward TL**, Spencer WJ, Altosaar I. The fate of human milk protein sCD14 in the infant.

39th Environmental Mutagen Society Meeting

2008

Rio Grande, Puerto Rico

Posters:

1. **Ward TL**, Van Osch FS, Hill KA. Mutation frequency and pattern in kidney tissue from a mouse model of elevated oxidative stress.
2. Crabbe RA, **Ward TL**, Hill KA. Mutation frequency and pattern in heart tissue of a mouse model of elevated oxidative stress and stress-induced heart disease.

TEACHING EXPERIENCE**Teaching Assistant**

2010-Present

University of Ottawa, Ottawa, ON

Faculty of Science

Biochemistry and Molecular Biology labs

- Taught basic and advanced molecular biology techniques, supervised and instructed labs, graded assignments (class sizes of 16-20 students, course codes: BCH 2333, 3356, 3346)

Instructor

2011 - 2013

University of Ottawa, Ottawa, ON

Faculty of Science

Detectives in Genes Mini Course

- Designed and implemented lectures to supplement organized experiments and taught molecular biology techniques to students in grades 8-11 during an annual week long course (class sizes of 22-24)

Research Supervisor

2012 - 2013

University of Ottawa, Ottawa, ON

Department of Biochemistry, Microbiology and Immunology

Undergraduate Research Opportunity Program

- Instructed and supervised undergraduate students as they completed small research projects throughout the school year (3 students)

Scientific Educator

2008

Thunder Bay Art Gallery, Thunder Bay, ON

Ontario Genomics Institute's (OGI) Gee in Genome Exhibit

- Designed and implemented experiments and lectures to supplement exhibit tours taken by students in grades 5-12 (class sizes of 15-35 students)

Instructor

2007, 2008

University of Western Ontario, London ON

Teachers' Science and Technology Outreach Program (TSTOP)

- Taught basic molecular biology techniques to secondary school teachers so they could then teach those skills to their students (one-on-one sessions)

Tutor

2006

NoKee Kwe, London, ON

- Taught members of the aboriginal community mathematics in preparation for high school equivalency exams (class sizes of 3-4)

ADMINISTRATIVE EXPERIENCE

Desk Coordinator, Member at Large/VP Social

2010-Present

University of Ottawa, Ottawa, ON

Biochemistry, Microbiology and Immunology Graduate Student Association

- Initiated and implemented the organization and distribution of desk space to postdoctoral fellows and graduate students of the department
- Organized speakers for a career day, served on the 'harmonization of procedures committee' for the department and attended monthly meetings
- Organized networking events for students, postdoctoral fellows and the faculty of the Biochemistry, Microbiology and Immunology Department