Human Milk Oligosaccharide Compositions Illustrate Global Variations in Early Nutrition

Article *in* Journal of Nutrition · February 2022

DOI: 10.1093/jn/nxac027

23 authors, including:

Jasmine Davis
Fitbit
21 PUBLICATIONS **1,219** CITATIONS
SEE PROFILE

Lauren Dimapasoc Wu
Thermo Fisher Scientific
21 PUBLICATIONS **857** CITATIONS
SEE PROFILE

Melanie A Martin
University of Washington Seattle
46 PUBLICATIONS **580** CITATIONS
SEE PROFILE

Michael Gurven
University of California, Santa Barbara
362 PUBLICATIONS **18,646** CITATIONS
SEE PROFILE

Human Milk Oligosaccharide Compositions Illustrate Global Variations in Early Nutrition

Anita Vinjamuri[1,2], Jasmine C. C. Davis[1,2,], Sarah M. Totten[1,2†], Lauren D. Wu[1,2], Laura D. Klein[3], Melanie Martin[4], EA Quinn[5], Brooke Scelza[6], Alicia Breakey[7], Michael Gurven[8], Grazyna Jasienska[9], Hillard Kaplan[10], Claudia Valeggia[11], Katie Hinde[12], Jennifer T. Smilowitz[2,13], Robin M. Bernstein[14,15], Angela M. Zivkovic[2,16], Michael J. Barratt[17,18], Jeffrey I. Gordon[17,18], Mark A. Underwood[2,19],David A. Mills[2,13], J. Bruce German[2,13], Carlito B. Lebrilla[1,2]*

[1]Department of Chemistry, [2]Foods for Health Institute, [13]Department of Food Science and Technology, [16]Department of Nutrition, and [19]Department of Pediatrics, University of California, Davis, CA 95616.
[3] Department of Human Evolutionary Biology, Harvard University, Cambridge, MA 02138
[4]Department of Anthropology, University of Washington
[5] Department of Anthropology, Washington University in St. Louis, Campus Box 1114, One Brookings Drive, St. Louis, MO 63130-4899, USA
[6] Department of Anthropology, University of California Los Angeles, 341 Haines Hall, Box 951553, Los Angeles, CA 90095, USA
[7] Wildwood School, Los Angeles CA
[8] Department of Anthropology, University of California Santa Barbara, Santa Barbara, CA 93106-3210, USA
[9] Department of Environmental Health, Faculty of Health Sciences, Jagiellonian University Medical College, ul. Skawińska 8, 31-066 Krakow, Poland [10] Health Economics and Anthropology, Chapman University, Orange CA
[11]Department of Anthropology, Yale University, 10 Sachem Street, New Haven, CT 06511, USA
[12]School of Human Evolution and Social Change, Arizona State University, AZ 85004
[14]Department of Anthropology, Institute of behavioral science[15], University of Colorado, Boulder, CO, United States
[17]Edison Family Center for Genome Sciences and Systems Biology and [18]Center for Gut Microbiome and Nutrition Research, Washington University School of Medicine, MO 6310
[†]Current association: Canary Center at Stanford, Stanford University School of Medicine, CA 94305
*Corresponding Author, Email: cblebrilla@ucdavis.edu

Funding: This work was supported by grants from the Bill and Melinda Gates Foundation (J.C.C.D., C.B.L.), the NIH [grants AT007079 to D.A.M., HD061923 to C.B.L. , R01AG024119

to MG and HK , and AT008759 to D.A.M.], the Peter J. Shields Endowed Chair in Dairy Food Science (D.A.M.), and NSF DDIG 1233270 to MM. The Etiology, Risk Factors and Interactions of Enteric Infections and Malnutrition and the Consequences for Child Health and Development (MAL-ED) Project was carried out as a collaborative project supported by the Bill & Melinda Gates Foundation, the Foundation for the National Institutes of Health, and the Fogarty International Center. The iLiNS-DYAD study was funded by the Bill & Melinda Gates Foundation and FANTA-2 project, AED. The ENID trial was supported by the UK Medical Research Council (MRC) (MC-A760-5QX00) and the UK Department for International Development (DFID), under the MRC/DFID Concordat agreement.

Competing interests: C.B.L., D.A.M, and J.B.G, are co-founders of Evolve Biosystems, Inc., a company focused on diet-based manipulation of the gut microbiota. .J.I.G. is a founder of Matatu, a company focused on microbial-based manipulation for improvement of health.

Data and materials availability: Materials and correspondence should be directed to Carlito B. Lebrilla, cblebrilla@ucdavis.edu.

Abbreviations:

2'FL-2'-Fucosyllactose
3'FL-3'-Fucosyllactose
3'SL-3'-Sialyllactose
3'SLN-3'-Sialyl-N-acetyllactosamine
6'SL-6'-Sialyllactose
6'SLN-6'-Sialyl-N-acetyllactosamine
DFLNHa-Difucosyllacto-N-hexaose a
DFLNHb-Difucosyllacto-N-hexaose b
DFLNHc-Difucosyllacto-N-hexaose b
DFLNnO I-Difucosyllacto-N-neooctaose I
DFLNnO III-difucosyllacto-N-neooctaose II
DFLNO I-Difucosyllacto-N-octaose I
DFLNO II-Difucosyllacto-N-octaose II
DFpLNH II-Difucosyl-para-lacto-N-hexaose II
DFS-LNH-Difucosylmonosialyllacto-N-hexaose

DFS-LNnH-Difucosylmonosialyllacto-N-neohexaose
DS-LNT-Disialyllacto-N-tetraose
EIC-Extracted Ion Chromatogram
F-LNH I-Fucosyl-lacto-N-hexaose I
F-LNH II-Fucosyl-lacto-N-hexaose II
F-LNO-Monofucosyllacto-N-octaose
F-LSTc-Monofucosylmonosialyllacto-N-neotetraose
FS-LNH I-Monofucosylmonosialyllacto-N-hexaose I
FS-LNH II-Monofucosylmonosialyllacto-N-hexaose II
FS-LNH III-Monofucosylmonosialyllacto-N-hexaose III
FS-LNH-Monofucosylmonosialyllacto-N-hexaose
FS-LNnH I-Monofucosylmonosialyllacto-N-neohexaose II
FS-LNO-Monofucosylmonosialyllacto-N-octaose
Fuc-Fucose
Gal-Galactose
GDP-Gross domestic product
Glc-Glucose
GlcNAc-N-acetylglucosamine
HMO-Human Milk Oligosaccharide
HPLC-High performance liquid chromatography
IFLNH I-Isomeric Fucosyl-lacto-N-hexaose I
IFLNH III-Isomeric Fucosyl-lacto-N-hexaose III
LDFT-Lactodifucotetraose
LNDFH I-Lacto-N-difucohexaose I
LNDFH II-Lacto-N-difucohexaose II
LNFP I-Lacto-N-fucopentaose I
LNFP II-Lacto-N-fucopentaose II
LNFP III-Lacto-N-fucopentaose III
LNFP V-Lacto-N-fucopentaose V
LNH-Lacto-N-hexaose
LNnH-Lacto-N-neohexaose
LNnT-Lacto-N-neotetraose
LNT-Lacto-N-tetraose
LSTa-Sialyllacto-N-tetraose a
LSTb-Sialyllacto-N-tetraose b
LSTc-Sialyllacto-N-neotetraose
MFpLNH IV-Fucosyl-para-lacto-N-hexaose
MS-Mass spectrometry
Neu5Ac-N-acetylneuraminic acid
PGC-Porous graphitized carbon

p-LNH-para-Lacto-N-hexaose
qTOF-quadropole Time of Flight
S-LNH-Monosialyllacto-N-hexaose
S-LNnH II-Monosialyllacto-N-neohexaose II
TFLNH-Trifucosyllacto-N-hexaose

ABSTRACT

Background. Human milk oligosaccharides (HMOs) are an abundant class of compounds found in human milk and have been linked to the development of the infant and specifically the brain, immune system, and gut microbiome.

Objective. Advanced analytical methods were used to obtain relative quantitation of many structures in approximately 2000 samples from over 1000 mothers in urban, semi-rural and rural sites across geographically diverse countries.

Methods. Liquid chromatography mass spectrometry-based analytical methods were used to profile the compounds with broad structural coverage and quantitative information. The profiles revealed their structural heterogeneity and their potential biological roles. Comparisons of HMO compositions were made between mothers of different age groups, lactation periods, infant's sex, and residing geographical locations.

Results. A common behavior found among all sites was a decrease in HMO abundances during lactation until approximately month six postnatal, where they remained relatively constant. The greatest variations in structural abundances were associated with the presence of α(1,2)-fucosylated species. Genomic analyses of the mothers were not performed, instead milk was phenotyped according to the abundances of α(1,2)-fucosylated structures. Mothers from the South American sites tended to have higher proportions of phenotypic secretors (mothers with relatively high concentrations α(1,2)-fucosylated structures) in their populations compared to the rest of the globe, with Bolivia~ 100%, Peru ~97%, Brazil~ 90%, and Argentina~ 85%. Conversely, the cohort sampled in Africa manifested the lowest proportion of secretors (South Africa ~ 63%, The Gambia~ 64%, and Malawi~75%). Furthermore, we compared total abundance of HMOs in secretors versus non-secretors and found that non secretors have lower abundances of HMOs

compared to secretors regardless of geographical location. We also observed compositional differences of the 50+ most abundant HMOs between milk types and geographical locations.

Conclusion. This study represents the largest structural HMO study to date and reveals the general behavior of HMOs during lactation among different populations.

KEY WORDS

Human Milk Oligosaccharides

Oligosaccharides

Mass Spectrometry

Glycans

Breast Milk

Carbohydrates

Lactose

Secretor

FUT2

INTRODUCTION

Human milk for infants is a manifestation of a highly adapted, dynamic and personalized process of human postnatal development. Human milk is 'dynamic' in the sense that time-dependent changes occur in the presence and concentration of milk bioactives within and across mothers and 'personalized' in the sense that maternal genotype, health status, and environmental exposures, including diet, can impact milk compositions.[1-5]

Human milk oligosaccharides (HMOs) are among the most abundant and diverse components of breast milk, with hundreds of unique structures identified to date.[6-9] HMOs serve as nutrients for highly-adapted early bacterial colonizers of the infant gut, including specific strains of bifidobacteria endowed with suites of gene encoded proteins dedicated to the import and utilization of HMOs [10-14].While the prebiotic effect is believed to be a major function, a small fraction of HMOs are absorbed in the small intestine and detectable in plasma [15] and urine[16] suggesting the potential for direct effects on host physiology including immunomodulation[17-19] and brain development.[20]

HMOs are assembled by glycosyltransferases to form either branched or linear structures. HMOs generally consist of a lactose [glucose (Glc) and galactose (Gal)] core, with variable combinations of N-acetylglucosamine (GlcNAc) and can be further bound to monosaccharides including sialic acids (N-acetylneuraminic acid or Neu5Ac) and fucose (Fuc).[21] The process yields an extensive number of oligosaccharides that in many cases are unique to human milk.[22]

Variations in HMOs are greatest among secretor genotypes. Secretors are individuals with a functioning FUT2 gene encoding α(1,2)-fucosyltransferase that attaches fucose via an α(1,2)-linkage to terminal Gal residues thereby producing blood antigens into secreted fluids (e.g., sweat, tears, semen, and milk) [21, 23-25]. Non-secretors have diminished ability to produce ABH or Lewis b

antigens (Le^b; Fucα1,2Galβ1,3[Fucα1,4]GlcNAcβ) due to mutations in the *FUT2* gene.[21, 23-25] Genetic studies have documented variations in the prevalence of the wild-type and mutant *FUT2* alleles around the world.[26-28] The prevailing existence of different genotypes in populations show that there are unique advantages to individuals. This notion is consistent with evidence suggesting a protective effect against, for example, otitis media[29]and autoimmune diseases[30, 31] for secretors and against viral diarrhea for non-secretors[32]. While different alterations in the FUT2 gene among various populations determine the secretor status of the mother, the milk and its function is guided by the abundances of HMO structures. Thus, classifying the milk phenotype, for example the amount of sialylation and fucosylation, is a more direct approach in surveying the differences between mothers' milks and correlating infant health outcomes. Advanced analytical methods now make it possible to accurately determine the phenotypic secretor status by directly quantitating the abundances of α(1,2)-fucosylated structures present in the breast milk.[23] For the purpose of this study, we refer to milk that corresponds to high amounts of α(1,2)-fucosylated structures as S+ milk, and milk corresponding to low abundances as S- milk. HMOs can therefore be used to type the milk according to the presence or absence of α(1,2)-fucosylated structures regardless of the genotype.[8, 23]

Mass spectrometry-based analytical methods have enabled more rapid and precise characterization of human milk, including the capability to simultaneously determine the abundances of hundreds of distinct carbohydrate structures. However, determination of HMOs are largely limited to studies of cohorts living in a small number of sites or geographic locales.[33-35] In this report, we describe the results of a cross-sectional analysis that quantified the abundances of HMOs to determine the natural variations during lactation among various sites involving mothers from different ethnic groups. Milk samples were obtained and analyzed from over 1000 mothers

living in 15 countries encompassing six continents and representing a diverse set of ecological and cultural backgrounds.

MATERIALS AND METHODS

Sources of breast milk samples

Breast milk samples (N=2234) were collected from mothers (N=1090) in 16 global sites, including urban, rural and semi-rural communities in 15 sites spanning Africa, Eurasia, the Americas, and Australia. Samples and resulting data were collected from different studies as detailed in the **Supplementary Text**. However, the quality controls and analytical methods for each sample remained the same. A detailed summary of sample information including country, population, collection procedures, sample size, and infant age is provided in **Table 1** and **Supplementary Table 1**. All infants were delivered full term. Written informed consent was obtained from all parents/guardians prior to study enrollment. Samples were collected using standardized protocols for all populations detailed in the **Supplementary Methods**. Details on the Ethical approval identifiers, trial registration information, and participant inclusion/exclusion criteria for each study site can also be found in the **Supplementary Methods.**

HMO extraction and mass spectrometric analysis

HMOs were extracted from breast milk samples using previously reported methods.[7, 8, 36, 37] Briefly, whole milk samples were aliquoted into 96-well plates, diluted, and then defatted via centrifugation. Proteins were precipitated with ethanol, and the resulting glycans were reduced with sodium borohydride (Sigma-Aldrich, St. Louis, MO). Solid phase extraction was performed

on graphitized carbon cartridges (Glygen, Columbia, MD) to remove lactose and salts. After the solvent was evaporated, purified HMOs were reconstituted and diluted prior to analysis. Standard solutions were made of HMO in water, with concentrations ranging from 0.05 mg/mL-0.2 mg/mL.

Extracted HMOs were analyzed on a nano-HPLC-TOF-MS The HPLC unit (Model series 1200, Agilent Technologies, Santa Clara, CA) that utilizes a capillary pump for sample loading (4 µL/min) and a nano pump for analyte separation (0.3 µL/min). Loading and separation were performed on a microfluidic chip packed with porous graphitized carbon via enrichment and analytical columns, respectively, using a binary gradient of solvent A [3% acetonitrile (ACN) in 0.1% formic acid] and solvent B (90% ACN in 0.1% formic acid). This system was coupled to an Agilent 6220 series TOF mass spectrometer. Detection was performed in the positive mode, and calibration was achieved with a dual nebulizer electrospray source with calibrant ions ranging from mass-to-charge (*m/z*) 118.086 to 2721.895.

Structural annotation of HMOs

Data were collected using Agilent MassHunter Workstation Data Acquisition software (B.02.01) and analyzed using Agilent MassHunter Qualitative Analysis software B.03.01 and B.06.00. The 'Find Compound by Molecular Feature' function was used to extract ion abundances to within 20 ppm of theoretical HMO masses. Individual HMOs were identified by accurate mass, retention time, and elution order defined previously developed HMO libraries [7, 8]. An in-house software program was used to align peaks due to minor retention time shift [37]. HMOs were grouped into classes as follows: fucosylated HMOs (any structure with Fuc), sialylated HMOs (any structure with Neu5Ac), undecorated HMOs (neither Fuc nor Neu5Ac present), and fucosylated plus sialylated HMOs (both Fuc and Neu5Ac present). Relative abundances (%) were calculated

by normalizing class and individual compound abundances to total HMO abundance in each breast milk sample. Compounds that were not identified in individual samples but were present in at least 50% of all samples were given a LOD/2 abundance.

Classification of milk secretor phenotype as S+ and S- based on HMO abundances

Phenotypic secretor status was determined following our previously published method [23]. Structures with known α(1,2)-Fuc linkages were identified by matching exact masses and retention times to previously developed annotated HMO libraries [7, 8]. The abundances of the most abundant α(1,2)-fucosylated structures, namely 2'-fucosyllactose(2'FL), lactodifucotetraose(LDFT), difucosyllacto-N-hexaose a(DFLNHa) and trifucosyllacto-N-hexaose (TFLNH) were summed and normalized to the total HMO abundances in a given sample, so that a relative α(1,2)-fucosylation value could be determined. Secretor status was assigned based on a previously established and validated threshold of 6%.[23] If this value exceeded the threshold. the mother was deemed a secretor (S+), conversely if the value fell below this threshold the mother was deemed a non-secretor (S-). The number of mothers producing S+ and S- milk in each location was determined and the proportion of S- mothers was calculated for each location. If a mother provided multiple samples from different postpartum time points, her secretor status was determined based on the secretor status determination in the majority of her samples. If there was no 'majority milk type', she was excluded from the statistical analysis (one mother from Davis, USA, two mothers from Malawi, and three mothers from Perth, Australia).

Statistical analyses

Mann-Whitney tests were used to determine the differences between absolute and relative abundances of the HMO classes. Furthermore, the data was grouped to show how secretor status, geographical location, age, sex, and lactation month affected the HMO profiles. An alpha correction of α=0.05 was used for the statistical analysis. Differences were determined when all samples from all time points were combined (S+ *N*=1709, S- *N*=524), and when samples were split by location including all time points (Argentina: S+ *N*=17, S- *N*=3; Boliva: S+ *N*=52, S- *N*=0; Bangladesh: S+ *N*=72, S- *N*=20; USA[Boston]: S+ *N*=15, S- *N*=5; Brazil: S+ *N*=97, S- *N*=12; USA[Davis]: S+ *N*=194, S- *N*=58; Gambia: S+ *N*=63, S- *N*=36; India: S+ *N*=166, S- *N*=81; Malawi: S+ *N*=601, S- *N*=208; Namibia: S+ *N*=10, S- *N*=2; Nepal: S+ *N*=20, S- *N*=8; Australia: S+ *N*=42, S- *N*=15; Peru: S+ *N*=227, S- *N*=8; Philippines: S+ *N*=13, S- *N*=5; Poland: S+ *N*=18, S- *N*=5; South Africa: S+ *N*=93, S- *N*=58). Due to the vast changes in milk composition throughout lactation, samples were binned based on lactation month, therefore milk collected from a single mother at different time points were treated as independent samples.

Data Availability

Data that support the findings of this study are available upon request from the corresponding author (C.B.L.).

RESULTS

HMO-based classification of milk into S+ and S- phenotypes

Over 2000 breast milk samples were collected from 1,090 mothers in 15 geographical sites. The samples obtained from six continents were analyzed under one protocol allowing direct comparison of abundances by classes and individual structures (**Table 1**). Using nano-HPLC-qTOF-MS, we identified 60 structures that were common to most samples. However, the total

number of unique structures varied for each mother, with the average count of nearly 100 structures in a single mother. The abundances for these structures varied widely, spanning four orders of magnitude.

HMOs containing Lewis b structures (α(1,2)-fucose), a feature of the secretor genotype (homozygous or heterozygous for the functional *FUT2* allele)[23,38,25], were most variable between mothers. We defined milk rich in Lewis b structures [2'-fucosyllactose(2'FL), lactodifucotetraose(LDFT), difucosyllacto-*N*-hexaose a(DFLNHa) and trifucosyllacto-*N*-hexaose (TFLNH)] that were consistently represented in samples collected within and across the different geographic sites S+ milk, belonging to a secretor mother. Milk containing a total relative abundance of <6% of these four Lewis b structures was defined as S+ milk, thus belonging to a non-secretor mother. This criterion was developed previously and has been validated with genomic data.[23] Phenotyping the milk addresses what the infant receives, while genotyping the mother does not necessarily translate to HMO abundances.

Figure 1 shows extracted ion chromatograms (EICs) of the most abundant α(1,2)-fucosylated compounds in human milk, namely 2'-fucosyllactose (2'FL) and lactodifucotetraose (LDFT) (**Figure 1a, b**)**.** These compounds were consistently higher in S+ milk (blue) compared to S- (red, near baseline) across all sites ($P<0.05$). Five other α(1,2)-Fuc-containing structures, namely lacto-*N*-fucopentaose I (LNFP I), lacto-*N*-difucohexaose I (LNDFH I), monofucosyllacto-*N*-hexaose I (MFLNH I), isomer I fucosyl-paralacto-*N*-hexaose (IFLNH I), and difucosyllacto-*N*-hexaose c (DFLNH c), were much less ubiquitous, and when detected, were present at much lower abundances than the four structures used for determination of secretor status (**Table 2**). Conversely, there were structures, particularly α(1,3)- and α(1,4)-fucosylated HMOs, that were significantly higher in S- milk (**Figure 1c-f, Table 2**). Three isomers with composition

3Hex:1HexNAc:1Fuc and having α(1,3/4)-Fuc linkages were produced in higher abundances in mothers with S- milk (**Figure 1e**).

The relative abundances of the 60 most common structures were shown for S+ milk and S- in **Figure 2** (see **Table 2**). The abundance of each HMO was normalized to the total abundances of the selected group, which made up approximately 97% of all abundances. For mothers that produce S+ milk, the most abundant HMOs were Lacto-N-Tetraose/Lacto-N-Neotetraose (LNT/LNnT). These two compounds are isomers and were difficult to resolve chromatographically, and the data presented are the sum of their combined abundances. At all sites, mothers with S- milk had higher abundances of LNFP II, suggesting this HMO is a potential marker of S- milk ($P<0.0001$; Mann-Whitney). Other HMOs, including MFLNH I, MFLNH III and DFpLNH II were also higher in S- milk achieving statistical significance ($P< 0.05$; Mann-Whitney) at all sites except Brazil ($P<0.38$; Mann-Whitney).

Total HMO abundances between sites and during lactation

The total abundances of HMOs between mothers from various sites were compared as a function of months postpartum. **Figure 3** compared total HMOs from mothers living in Brazil, Bangladesh, Davis-USA, Peru, India and South Africa sampled between postpartum months 1-6. These sites were selected because they provided the most extensive longitudinal sampling. The highest abundances were observed at month 1 and decreased uniformly thereafter. Each site showed similar behavior and similar decreases in total abundances with some variations in months 3 and 4. For example, in both the Peru and India sites total abundances decreased at month 3 and remained nearly constant thereafter, while in the other sites HMO abundances dropped uniformly

between month 1 and 6. When comparing total absolute abundance in secretors between month 1 and 6 statistical significance was achieved across all selected sites, were higher significance was achieved in sites with more sampling **(Figure 3b).** Similarly, statistical significance was similarly achieved when comparing abundances in non-secretors between month 1 and 6 where there were sufficient number of samples. HMO abundances were also comparing between S+ and S- milk across these sites by averaging the means of all time points within one site. In all counties except Brazil, S+ milk possessed higher abundances of HMOs ($P<0.0001$, Mann-Whitney).

Total HMO abundances and maternal age

Variations in HMOs by maternal age were examined at sites with large numbers of samples [Malawi, Davis-USA, India, Peru]. The milk samples were binned into age groups <20, 21-30, 31-40 and 41-50 years. Total absolute abundances of HMOs were not different between the age groups **(Supplementary Figure 1)** or between S+ and S- milk (Data not shown). There were also no maternal age-related differences in total fucosylation, total sialylation or levels of 2'FL (the latter among S+ mothers).

Total HMO abundances and infant sex

HMOs were compared to determine whether the sex of the infant affected total HMO abundances. Comparisons were made within sites (**Supplementary Figure 2**). We found no statistically significant differences in absolute HMO abundances in milk from mothers of male

compared to mothers of female infants. Similarly, total fucosylation and total sialylation yielded no differences based on sex of the offspring (not shown).

Variations in HMO subtypes across study sites and secretor status

Fucosylated HMOs. Fucosylated oligosaccharides comprised the largest and most abundant groups of HMO structures. We first compared the total abundances of α(1-2)-fucosylated HMOs in milk collected monthly during the first 6 months from mothers at study sites with longitudinal sampling, namely Bangladesh, Bolivia, India, Peru, and South Africa (**Figure 4a and Table 3)**. The proportion of α(1,2)-fucosylated structures in secretors increased significantly between postnatal months 1-6, with statistical significance across all selected sites ($P<0.05$-$P<0.0001$, Mann-Whitney). However for S- milk, only the site with large sample size (India) achieved statistical significance. S+ milk had consistently higher abundances of α(1,2)-fucosylated HMOs (ranging from 15-20% of total HMOs) compared to S- milk (<5% of total HMOs; $P<0.0001$, Mann-Whitney)). Interestingly S- milk, while containing very low abundances of α(1,2)-fucosylated HMOs, at times had nonzero abundances slightly above the baseline ($P<0.5$; Mann-Whitney). As 2-fucosyllactose is the most abundant of these structures, it increased over time as expected (**Supplementary Figure 3**). Total fucosylation (oligosaccharides that contain fucose regardless of the linkage) increased between months 1 and 6 in all sites, however significance was only achieved in the India site and only for S+ milk ($P<0.05$, Mann-Whitney) (**Figure 4b**). Total fucosylation differed slightly between S+ and S- milk (**Figure 4b**). In general, S- milk had low abundances of fucosylated structures compared to S+ milk, reflecting the deficit in α(1,2)-fucosylated HMOs, but partially compensated by increases in (1,3/4)-fucose linkages.

Variations in fucosylated structures were observed between milk from different countries within each secretor phenotype (**Figure 2**). Mothers with S+ milk from Nepal had significantly higher levels of DFLNHa ($P<0.0002$, Mann-Whitney) and lower levels of 2'FL ($P<0.0001$) (both S+, secretor markers) compared to the other sites. Human milk obtained from mothers residing in the Boston-USA and Gambia sites had high relative abundances of LNFP I and LNFP III compared to all other sites, and low abundances of LNFP II. Milk from mothers residing in Perth-Australia had relatively high abundances of LSTc and LSTb compared to other sites. Interestingly, 2-fucosyllactose was not generally the most abundant in S+ milk among all countries. Australia, Boston-USA, and Namibia all had S+ milk with the most abundant α(1-2)-fucosylated compound being LDFT ($P = 0.005$, Mann-Whitney).

Sialylated HMOs. Sialylated oligosaccharides generally had much lower abundances in human breast milk than fucosylated species. The mean abundances of sialylated HMOs in samples collected over the first 6 months postnatal from different study sites are shown in **Figure 5** and **Table 3.** Total abundances of sialylated HMOs were relatively constant throughout lactation for both S+ and S- phenotypes. Likewise, the relative abundances of sialylated HMOs were comparable across study sites with extensive sampling (Bangladesh, Bolivia, India, Peru, and South Africa) averaging 12.3% ± 3.4 of the total HMOs in S- milk and 15.2% ± 3.8 in S+ milk ($P<0.05$-$P<0.0001$, Mann-Whitney). The most abundant sialylated HMOs when considering all samples were the combined group of LSTc+LSTb (in both S+ and S- samples) (**Figure 2**). Note however that the values from Peru for S- milk were obtained from samples provided by a single mother due to the very low prevalence of S- milk in this population. In populations with greater frequencies of S- milk, there was a greater relative abundance of sialylated structures in S- milk

compared to S+ milk. In samples from India, Bangladesh and South Africa, the absolute abundances of sialylated HMOs in month 1 was nearly 15% greater in S- compared to S+ milk.

It is noteworthy that the sialylated HMO DSLNT, which has been shown to be protective against necrotizing enterocolitis in an animal model[39] and in a human preterm infant cohort [40] was often found in low abundance and was not one of the 60 HMOs found in all sites.

Undecorated (nonfucosylated and nonsialylated) HMOs. The total abundances of undecorated HMOs (lacking both fucose and sialic acid) in milk from mothers at the various study sites were compared in **Supplementary Figure 4** and **Table 3.** Undecorated HMOs were lower in S+ milk compared to S- milk (P=0.002, Mann-Whitney).[41] Between sites, overall relative abundances of undecorated HMOs were comparable (31.6±8.9%) with the exception of Peru, which exhibited significantly lower levels (20.5± 5.2%) (P=<0.0001) due to the higher amounts of fucosylation and diminished numbers of S- milk. Across all samples the most abundant undecorated HMO was the LNT/LNnT group (Figure 2).

Global Variations of S+ and S- milk

With the samples from different locations and the ability to distinguish S+ and S- milk, we determined the distribution of the two phenotypes globally. The sites are located on the map along with the fraction of S+ (blue) and S- (red) milk (**Figure 6)**. The fraction of the mothers that produced S- milk was lower than those that produced S+ at every geographical site. High proportions of S- milk was found in Africa (37% in South Africa and 36% in The Gambia). These

values contrast with other sites in Africa including Namibia (17%) and Malawi (25%) that had lower proportions of S- milk. Other sites with high levels of S- milk included those in India 35% and Nepal 29%. Bangladesh was lower with only 20% S- mothers. The USA sites (Davis, CA (22%) and Boston MA, 25%) were similar to the Poland (22%) site and the Australian site (19%). Low values were found in South America sites with Bolivia (0%), Peru (3%), Brazil (10%). The samples from these sites were obtained from indigenous populations. The other South American site, Argentina (15%) included nonindigenous populations.

DISCUSSION

The analyses of the HMO compositions of human milk samples collected from 16 sites around the world using a sensitive HPLC-qTOF-MS-based approach provided the most extensive dataset reported to date. While the samples were collected from different studies, the analytical method for each sample was the same. Unfortunately, absolute quantitation of total HMO and specific structures was not possible due to the low availability of standards in the earliest analysis.

This study significantly expands on our previous work at a single site in The Gambia[23]. The large number of samples from various geographical locations allowed us to further explore factors that may affect HMO production during lactation. Factors such as the age of the mother and the sex of the infant do not affect HMO abundances. The sex of the infant had been previously reported to affect milk compositions, suggesting that some components of human milk may be tailored to sex-differentiated developmental priorities.[42] However, examination of milk among all populations and within each site yielded no significant variations in total HMO abundances based on sex of a mother's child. Because neither dietary data nor maternal nutritional status were

available for this study, we were not able to determine the extent to which these factors might influence abundances of individual structures nor levels of fucosylation and sialylation. However, the comparison of per capita gross domestic product (GDP) was obtained by comparing total abundances with published GDP. Interestingly, we observed some correlation between per capita GDP and total levels of HMOs from mothers living at the different sites. (**Supplementary Figure 5**). As shown, countries with high GDP per capita tended to have milk with the most abundant levels of HMOs. Likewise, mothers from sites with the low GDP per capita tended to have lower levels of HMOs. Among countries with low GDPs, there was a common minimum level of HMOs, while the trend towards higher abundances of HMOs did not appear to manifest until significantly higher GDPs were obtained. It would be difficult to make conclusions regarding GDP and HMO production as the amounts of HMOs fed to the infant can vary depending on feeding local feeding practices. Furthermore, the sample size, though large in totality, is still small at the local level and cannot fully represent the respective nation. However, we encourage further studies on societal effects on milk production.

Genomic analysis has been previously used to obtain global distributions of secretors and nonsecretors. [43] While this approach uses obtains the genotypic status, the phenotype, i.e., the actual abundances of different structures and structural types in the milk, are those that affect infant health and developmental outcomes.[44] Hence, the concentrations of HMOs are crucial to understanding the health outcomes of infants providing information of distinct nutritional components that align with secretor status. Additionally, among secretors there are large variations in the abundances of fucosylated and sialylated species that are important contributors to infant development. The concentrations can then be used to phenotype milk based on the abundances of specific structures, namely $\alpha(1,2)$-fucosylated species, without genomic data. For the purposes of

this study, we therefore refer to milk that corresponds to high amounts of α(1,2)-fucosylated structures as S+ milk, and milk corresponding to low abundances as S- milk.

The fraction of mothers who produce S- milk was nearly 40% in West and South Africa and South Asia, while nearly zero in parts of Latin America. The fraction of nonsecretor mothers (and hence S- milk) is often cited to be ~20%. However, this number is based primarily on studies of European and Euro-American mothers, which is consistent with our own results for sites in Europe and the USA. The USA sites (Davis CA, Boston MA), and Poland (the sole European site) had S- milk in proportions similar to those previously reported. [34] The rarity of S- milk among indigenous populations in South America suggests either founder effects during human migration into Beringia or selection from pathogen pressure. Many infectious diseases including cholera, whose severity is associated with blood type, can have devastating effects on populations. [17] As previously discussed, blood typing is similar to human milk in that the presence or absence of α(1,2)-fucosylated structures is the key determinant. A study of the cholera outbreak in Peru in 1991 found that those with blood type O, which occurs at very high frequency among South American indigenous populations, had more severe symptoms and were eight times more likely to be hospitalized, emphasizing the relevance of glycosylation on infectious diseases. [45]

A distinguishing feature of humans and primates compared to other mammals is the high level of fucosylated structures with humans having the highest abundances.[46] Fucosylated structures, or the presence of at least one fucose, increased in relative abundances throughout lactation. Previous studies similarly noted that fucosylation increased throughout lactation for the first six months regardless of secretor status.[41] Fucosylation was generally higher in S+ milk. As a consequence of higher fucosylation in S+ milk, the amount of undecorated HMOs was lower compared to S- negative milk, consistent with earlier findings. [41] In contrast, sialylation is the

lowest among humans but is significantly greater in bovine and porcine milk [47] . In this study, though sialylated oligosaccharides in the human milk samples were still low, they were relatively higher in S- compared to S+ milk. Previous studies have also reported the higher abundances of sialylated HMOs in nonsecretor (S-) milk.[41]

The structural variety of HMOs and their different relative abundances have the potential to endow human milk from different mothers with distinct functional properties, including modulating its effect on the developing microbiota and its effects on infection/colonization with enteropathogens.[48] Members of the infant gut microbiota have specific glycosyl hydrolases and glycan-binding proteins that either promote the fitness of specific community members or block the host from infection. For example, HMOs promote the growth of *Bifidobacterium longum* subsp. *infantis*, a gut symbiont that is richly endowed with a suite of genes specifically adapted to import and utilize HMOs. The HMO composition of breastmilk thus has the potential to influence the fitness of strain-level variants of this and other related bifidobacterial species and to shape a program of normal postnatal community development (succession) that has been identified in healthy individuals and that is impaired in those with undernutrition [13, 49] Preclinical studies in gnotobiotic animals and clinical studies of the effects of repairing microbial community immaturity in children with acute malnutrition support the notion that healthy development of the microbiota is causally linked to healthy growth [13, 49] Similarly, HMOs with α(1-2) linked Fuc (S+ milk) are associated with decreased incidence and severity of diarrhea caused by *Campylobacter jejuni* and enteropathogenic *E. coli* [50, 51] - enteropathogens that are ubiquitous in many low income countries where childhood undernutrition is prevalent. [52] S- milk is enriched in HMOs that bind to *Helicobacter pylori*, and enteropathogenic *E. coli* preventing their attachment to gut epithelial cells[53]. Individuals homozygous for *FUT2* mutations (non-secretors) also show resistance to

399 norovirus infection [54, 55] or considering this relationship from the viewpoint of the pathogen, both
400 rotavirus and norovirus (two of the most common causes of viral diarrhea in infants) appear to
401 prefer the secretor host. [32] These mothers could in turn provide protection to their infant by
402 delivering S- milk. Determining infant infectious disease risk should ideally include consideration
403 of the secretor status of both the mother and the infant to address such key question as whether S+
404 milk is particularly beneficial for the non-secretor infant and vice versa.

405 The global results suggest that secretor status and the complement of HMO produced by a
406 mother during lactation are influenced in part by adaptations shaped by ancestral nutritional and
407 disease ecologies experienced by diverse human populations. Indeed, immunofactors in breastmilk
408 are associated with subsistence practices that affect nutritional intake and pathogen exposure of
409 diverse traditional societies and demonstrate the importance of considering populations within
410 their contemporary, historical, and prehistorical contexts.[45] The "first-step" findings described
411 here highlight the importance of multi-population studies to better characterize the relationships
412 among maternal characteristics, HMO composition, early gut community development, the
413 products of microbial metabolism of these HMOs, and measures of infant health status.[46]
414 Delineation of these relationships, along with those mediated by other key constituents of
415 breastmilk ,e.g., secreted immunoglobulins and antibacterial proteins, will help guide the design
416 of future prebiotic approaches based on purified milk components (or synthetically-produced
417 mimetics) and/or synbiotics (prebiotics combined with a probiotic micro-organisms) that promote
418 healthy gut community development healthy growth of infants , and even healthy immune and
419 inflammatory responses over a lifetime.
420 .
421

ACKNOWLEDGMENTS

We thank all of the mothers who participated, the communities who allowed the research and contributed samples. We also thank all the fieldworkers, staff and researchers, clinical staff, and laboratory technicians who collected the samples and data, and field assistant/translator John Jakurama.

We are grateful to the PIs and their colleagues who oversaw the studies at MAL-ED sites [Tahmeed Ahmed (icddr,b, Bangladesh), Aldo Lima (Universidade Federal do Ceará, Brazil), Margaret Kosek (Johns Hopkins School of Public Health, Iquitos, Peru), Gagandeep Kang (Christian Medical College, Vellore, India) and Pascale Bessong (University of Venda, South Africa)], We are grateful to the iLNS-DYAD-Malawi PIs and study team [Ken Maleta (University of Malawi College of Medicine), Per Ashorn (University of Tampere School of Medicine, Finland), Robin Bernstein (University of Colorado)and Jennifer T. Smilowitz and Kathryn Dewey (University of California, Davis)], We would like to thank Andrzej Galbarzyck for critical organizing at the Mogielica Human Ecology Study Site (Poland).The Argentine samples from Qom mothers (Qom field assistants and Fundación ECO). Milk samples from Nepal were collected among ethnic Tibetan mothers in northern Gorkha district (Jhangchuk Sangmo, Nyima Sangmo, Tsewang Palton Tibetan field assistants, Dr. Diki Bista, co-PI); milk samples from the Philippines were collected from mothers in rural and urban Cebu, Philippines (Office of Population Studies, University of San Carlos). The Argentine samples from Qom mothers (Qom field assistants and Fundación ECO).

Author contributions: L.D.K., M.M., E.A.Q., A.B., K.H., J.T.S., A.M.Z., M.J.B., J.I.G., M.A.U., C.V. and J.B.G. oversaw sample collection and provided samples. L.D.W., S.M.T., and J.C.C.D prepared samples. A.V., L.D.W., S.M.T., and J.C.C.D analyzed HMO data. A.V. and J.C.C.D. performed statistical analyses on HMO data and infant disease metadata. C.B.L. designed the study and oversaw the analyses. A.V. and C.B.L. interpreted results and wrote the manuscript together with J.C.C.D., M.B. and J.I.G. All authors have read and approved the final version: A.V., J.C.D., S.M.T., L.D.W., L.D.K., M.M., E.Q., B.S., A.B., M.G., G.J., H.K., C.V., K.H., J.T.S., R.M.B., A.M.Z., M.J.B., J.I.G., M.A.U., D.A.M., J.B.G., C.B.L.

REFERENCES

1. Ballard, O.; Morrow, A., Human Milk Composition Nutrients and Bioactive Factors. *Pediatric clinics of North America* **2013,** *60*, 49-74.
2. Dorea, J. G., Selenium and breast-feeding. *British Journal of Nutrition* **2002,** *88* (5), 443-461.
3. Innis, S. M., Impact of maternal diet on human milk composition and neurological development of infants. *The American Journal of Clinical Nutrition* **2014,** *99* (3), 734S-741S.
4. Picciano, M. F., Nutrient Composition of Human Milk. *Pediatric Clinics of North America* **2001,** *48* (1), 53-67.
5. Hinde, K.; Milligan, L. A., Primate milk: Proximate mechanisms and ultimate perspectives. *Evolutionary Anthropology: Issues, News, and Reviews* **2011,** *20* (1), 9-23.
6. Ruhaak, L. R.; Lebrilla, C. B., Analysis and role of oligosaccharides in milk. *BMB Reports* **2012,** *45* (8), 442-451.
7. Wu, S.; Grimm, R.; German, J. B.; Lebrilla, C. B., Annotation and Structural Analysis of Sialylated Human Milk Oligosaccharides. *Journal of Proteome Research* **2011,** *10*, 856-868.
8. Wu, S.; Tao, N.; German, J. B.; Grimm, R.; Lebrilla, C. B., Development of an Annotated Library of Neutral Human Milk Oligosaccharides. *Journal of Proteome Research* **2010,** *9*, 4138-4151.
9. Newburg, D. S.; Neubauer, S. H., *Handbook of Milk Composition*. Academic Press, Inc.: 1995; p 919.
10. Engfer, M. B.; Stahl, B.; Finke, B.; Sawatzki, G.; Daniel, H., Human milk oligosaccharides are resistant to enzymatic hydrolysis in the upper gastrointestinal tract. *The American Journal of Clinical Nutrition* **2000,** *71*, 1589-1596.
11. Garrido, D.; Dallas, D. C.; Mills, D. A., Consumption of human milk glycoconjugates by infant-associated bifidobacteria: mechanisms and implications. *Microbiology* **2013,** *159*, 649-664.
12. LoCascio, R. G.; Ninonuevo, M. R.; Freeman, S. L.; Sela, D. A.; Grimm, R.; Lebrilla, C. B.; Mills, D. A.; German, J. B., Glycoprofiling of Bifodobaterial Consumption of Human Milk Oligosaccharides Demonstrates Strain Specific, Preferential Consumption of Small Chain Glycans Secreted in Early Human Lactation. *Journal of Agricultural and Food Chemistry* **2007,** *55*, 8914-8919.
13. Gehrig, J. L.; Venkatesh, S.; Chang, H.-W.; Hibberd, M. C.; Kung, V. L.; Cheng, J.; Chen, R. Y.; Subramanian, S.; Cowardin, C. A.; Meier, M. F.; O'Donnell, D.; et al., Effects of microbiota-directed foods in gnotobiotic animals and undernourished children. *Science* **2019,** *365* (6449), eaau4732.
14. Hinde, K.; Lewis, Z. T., Mother's littlest helpers. *Science* **2015,** *348* (6242), 1427-1428.
15. Ruhaak, L. R.; Stroble, C.; Underwood, M. A.; Lebrilla, C. B., Detection of milk oligosaccharides in plasma of infants. *Analytical and Bioanalytical Chemistry* **2014,** *406*, 5775-5784.
16. De Leoz, M. L. A.; Wu, S.; Strum, J. S.; Ninonuevo, M. R.; Gaerlan, S. C.; Mirmiran, M.; German, J. B.; Mills, D. A.; Lebrilla, C. B.; Underwood, M. A., A quantitative and comprehensive method to analyze human milk oligosaccharide structures in the urine and feces of infants. *Analytical and Bioanalytical Chemistry* **2013,** *405*, 4089-4105.
17. Coppa, G. V.; Zampini, L.; Galeazzi, T.; Facinelli, B.; Ferrante, L.; Capretti, R.; Orazio, G., Human Milk Oligosaccharides Inhibit the Adhesion to Caco-2 Cells of Diarrheal

Pathogens: Escherichia coli, Vibrio cholerae, and Salmonella fyris. *Pediatric Research* **2006,** *59* (3), 377-382.
18. Hong, P.; Ninonuevo, M. R.; Lee, B.; Lebrilla, C.; Bode, L., Human milk oligosaccharides reduce HIV-1-gp120 binding to dendritic cell-specific ICAM3-grabbing non-integrin (DC-SIGN). *British Journal of Nutrition* **2008,** *101* (4), 482-486.
19. Eiwegger, T.; Stahl, B.; Haidl, P.; Schmitt, J.; Boehm, G.; Dehlink, E.; Urbanek, R.; Szépfalusi, Z., Prebiotic oligosaccharides: In vitro evidence for gastrointestinal epithelial transfer and immunomodulatory properties. *Pediatric Allergy and Immunology* **2010,** *21*, 1179-1188.
20. Wang, B., Sialic Acid is an Essential Nutrient for Brain Development and Cognition. *Annu. Rev. Nutr.* **2009,** *29*, 177-222.
21. Kunz, C.; Rudloff, S., Biological functions of oligosaccharides in human milk. *Acta Paediatr* **1993,** *82*, 903-912.
22. Urashima, T.; Kitaoka, M.; Asakuma, S.; Messer, M., Milk Oligosaccharides. In *Advanced Dairy Chemistry: Lactose, Water, Salts, and Minor Constituents*, McSweeney, P.; Fox, P., Eds. Springer Science: New York, 2009; pp 295-348.
23. Totten, S. M.; Zivkovic, A. M.; Wu, S.; Ngyuen, U.; Freeman, S. L.; Ruhaak, R. L.; Darboe, M. K.; German, J. B.; Prentice, A. M.; Lebrilla, C. B., Comprehensive Profiles of Human Milk Oligosaccharides Yield Highly Sensitive and Specific Markers for Determining Secretor Status in Lactating Mothers. *Journal of Proteome Research* **2012,** *11* (12), 6124-6133.
24. Brooks, S. A.; Dwek, M. V.; Schumacher, U., *Functional and Molecular Glycobiology*. BIOS Scientific Publishers Ltd: Oxford OX4 1RE, UK, 2002.
25. Thurl, S.; Henker, J.; Siegel, M.; Tovar, K.; Sawatzki, G., Detection of four human milk groups with respect to Lewis blood group dependent oligosaccharides. *Glycoconjugate Journal* **1997,** *14*, 795-799.
26. Mottram, L.; Wiklund, G.; Larson, G.; Qadri, F.; Svennerholm, A.-M., FUT2 non-secretor status is associated with altered susceptibility to symptomatic enterotoxigenic Escherichia coli infection in Bangladeshis. *Scientific Reports* **2017,** *7* (1), 10649.
27. Liu, Y.; Koda, Y.; Soejima, M.; Pang, H.; Schlaphoff, T.; du Toit, E. D.; Kimura, H., Extensive polymorphism of the FUT2 gene in an African (Xhosa) population of South Africa.
28. Kudo, T.; Iwasaki, H.; Nishihara, S.; Shinya, N.; Ando, T.; Narimatsu, I.; Narimatsu, H., Molecular Genetic Analysis of the Human Lewis Histo-blood Group System: II. SECRETOR GENE INACTIVATION BY A NOVEL SINGLE MISSENSE MUTATION A385T IN JAPANESE NONSECRETOR INDIVIDUALS. *Journal of Biological Chemistry* **1996,** *271* (16), 9830-9837.
29. Santos-Cortez, R. L. P.; Chiong, C. M.; Frank, D. N.; Ryan, A. F.; Giese, A. P. J.; Bootpetch Roberts, T.; Daly, K. A.; Steritz, M. J.; Szeremeta, W.; Pedro, M., et al., FUT2 Variants Confer Susceptibility to Familial Otitis Media. *Am J Hum Genet* **2018,** *103* (5), 679-690.
30. Smyth, D. J.; Cooper, J. D.; Howson, J. M. M.; Clarke, P.; Downes, K.; Mistry, T.; Stevens, H.; Walker, N. M.; Todd, J. A., FUT2 Nonsecretor Status Links Type 1 Diabetes Susceptibility and Resistance to Infection. *Diabetes* **2011,** *60* (11), 3081.
31. Ye, B. D.; Kim, B. M.; Jung, S.; Lee, H. S.; Hong, M.; Kim, K.; Moon, J. W.; Baek, J.; Oh, E. H.; Hwang, S. W.; Park, S. H.; Yang, S. K.; Song, K., Association of FUT2 and ABO with Crohn's disease in Koreans. *J Gastroenterol Hepatol* **2020,** *35* (1), 104-109.

32. Rossouw, E.; Brauer, M.; Meyer, P.; du Plessis, N. M.; Avenant, T.; Mans, J., Virus Etiology, Diversity and Clinical Characteristics in South African Children Hospitalised with Gastroenteritis. *Viruses* **2021,** *13* (2).
33. Erney, R. M.; Malone, W. T.; Skelding, M. B.; Marcon, A. A.; Kleman–Leyer, K. M.; O'Ryan, M. L.; Ruiz–Palacios, G.; Hilty, M. D.; Pickering, L. K.; Prieto, P. A., Variability of Human Milk Neutral Oligosaccharides in a Diverse Population. *Journal of Pediatric Gastroenterology and Nutrition* **2000,** *30* (2), 181-192.
34. McGuire, M. K.; Meehan, C. L.; McGuire, M. A.; Williams, J. E.; Foster, J.; Sellen, D. W.; Kamau-Mbuthia, E. W.; Kamundia, E. W.; Mbugua, S.; Moore, S. E.; et al., What's normal? Oligosaccharide concentrations and profiles in milk produced by healthy women vary geographically. *The American Journal of Clinical Nutrition* **2017**.
35. Plows, J. F.; Berger, P. K.; Jones, R. B.; Alderete, T. L.; Yonemitsu, C.; Najera, J. A.; Khwajazada, S.; Bode, L.; Goran, M. I., Longitudinal Changes in Human Milk Oligosaccharides (HMOs) Over the Course of 24 Months of Lactation. *J Nutr* **2021,** *151* (4), 876-882.
36. Ninonuevo, M. R.; Park, Y.; Yin, H.; Zhang, J.; Ward, R. E.; Clowers, B. H.; German, J. B.; Freeman, S. L.; Killeen, K.; Grimm, R.; Lebrilla, C. B., A Strategy for Annotating the Human Milk Glycome. *Journal of Agricultural and Food Chemistry* **2006,** *54*, 7471-7480.
37. Totten, S. M.; Wu, L. D.; Parker, E. A.; Davis, J. C. C.; Hua, S.; Stroble, C.; Ruhaak, L. R.; Smilowitz, J. T.; German, J. B.; Lebrilla, C. B., Rapid-throughput glycomics applied to human milk oligosaccharide profiling for large human studies. *Analytical and Bioanalytical Chemistry* **2014,** *406*, 7925-7935.
38. Kumazaki, T.; Yoshida, A., Biochemical evidences that secretor gene, Se, is a structural gene, coding a specific fucosyltransferase. *Proceedings of the National Academy of Sciences of the United States of America* **1984,** *81*, 4193-7.
39. Zhang, W.; He-Yang, J.; Zhuang, W.; Liu, J.; Zhou, X., Causative role of mast cell and mast cell-regulatory function of disialyllacto-N-tetraose in necrotizing enterocolitis. *Int Immunopharmacol* **2021,** *96*, 107597.
40. Masi, A. C.; Embleton, N. D.; Lamb, C. A.; Young, G.; Granger, C. L.; Najera, J.; Smith, D. P.; Hoffman, K. L.; Petrosino, J. F.; Bode, L.; Berrington, J. E.; Stewart, C. J., Human milk oligosaccharide DSLNT and gut microbiome in preterm infants predicts necrotising enterocolitis. *Gut* **2020**.
41. Xu, G.; Davis, J. C. C.; Goonatilleke, E.; Smilowitz, J. T.; German, J. B.; Lebrilla, C. B., Absolute Quantitation of Human Milk Oligosaccharides Reveals Phenotypic Variations during Lactation. *The Journal of Nutrition* **2016,** *147* (1), 117-124.
42. Galante, L.; Milan, A. M.; Reynolds, C. M.; Cameron-Smith, D.; Vickers, M. H.; Pundir, S. Sex-Specific Human Milk Composition: The Role of Infant Sex in Determining Early Life Nutrition *Nutrients* [Online], 2018. PubMed. http://europepmc.org/abstract/MED/30200404 (accessed 2018/09//).

43. Arrouzet, C. J.; Ellis, K.; Kambhampati, A.; Chen, Y.; Steele, M.; Lopman, B., Population-Level Human Secretor Status Is Associated With Genogroup 2 Type 4 Norovirus Predominance. *The Journal of Infectious Diseases* **2020,** *221* (11), 1855-1863.

44. Klein, L. D.; Huang, J.; Quinn, E. A.; Martin, M. A.; Breakey, A. A.; Gurven, M.; Kaplan, H.; Valeggia, C.; Jasienska, G.; Scelza, B.; Lebrilla, C. B.; Hinde, K., Variation

among populations in the immune protein composition of mother's milk reflects subsistence pattern. *Evolution, Medicine, and Public Health* **2018,** *2018* (1), 230-245.

45. Swerdlow, D. L.; Mintz, E. D.; Rodriguez, M.; Tejada, E.; Ocampo, C.; Espejo, L.; Barrett, T. J.; Petzelt, J.; Bean, N. H.; Seminario, L.; Tauxe, R. V.
46. Tao, N.; Wu, S.; Kim, J.; An, H. J.; Hinde, K.; Power, M. L.; Gagneux, P.; German, J. B.; Lebrilla, C. B., Evolutionary glycomics: characterization of milk oligosaccharides in primates. *J Proteome Res* **2011,** *10* (4), 1548-57.
47. Tao, N.; DePeters, E. J.; German, J. B.; Grimm, R.; Lebrilla, C. B., Variations in bovine milk oligosaccharides during early and middle lactation stages analyzed by high-performance liquid chromatography-chip/mass spectrometry. *J Dairy Sci* **2009,** *92* (7), 2991-3001.
48. Underwood, M. A., Impact of probiotics on necrotizing enterocolitis. *Semin Perinatol* **2017,** *41* (1), 41-51.
49. Raman, A. S.; Gehrig, J. L.; Venkatesh, S.; Chang, H.-W.; Hibberd, M. C.; Subramanian, S.; Kang, G.; Bessong, P. O.; Lima, A. A. M.; Kosek, M. N.; et al., A sparse covarying unit that describes healthy and impaired human gut microbiota development. *Science* **2019,** *365* (6449), eaau4735.
50. Newburg, D. S.; Ruiz-Palacios, G.; Altaye, M.; Chaturvedi, P.; Meinzen-Derr, J.; de Lourdes guerrero, M.; Morrow, A. L., Innate protection conferred by fucosylated oligosaccharides of human milk against diarrhea in breastfed infants. *Glycobiology* **2004,** *14* (3), 253-263.
51. Morrow, A. L.; Ruiz-Palacios, G. M.; Altaye, M.; Jiang, X.; Guerrero, M. L.; Meinzen-Derr, J. K.; Farkas, T.; Chaturvedi, P.; Pickering, L. K.; Newburg, D. S., Human Milk Oligosaccharide Blood Group Epitopes and Innate Immune Protection against Campylobacter and Calcivirus Diarrhea in Breastfed Infants. In *Protecting Infants through Human Milk*, Pickering, L. K.; Morrow, A. L.; Ruiz-Palacios, G. M.; Schanler, R. J., Eds. Springer US: 2004; pp 443-446.
52. Platts-Mills, J. A.; Liu, J.; Rogawski, E. T.; Kabir, F.; Lertsethtakarn, P.; Siguas, M.; Khan, S. S.; Praharaj, I.; Murei, A.; Nshama, R.; et al., Use of quantitative molecular diagnostic methods to assess the aetiology, burden, and clinical characteristics of diarrhoea in children in low-resource settings: a reanalysis of the MAL-ED cohort study. *The Lancet Global Health* **2018,** *6* (12), e1309-e1318.
53. Simon, P. M.; Goode, P. L.; Mobasseri, A.; Zopf, D., Inhibition of Helicobacter pylori binding to gastrointestinal epithelial cells by sialic acid-containing oligosaccharides. *Infection and Immunity* **1997,** *65* (2), 750-7.
54. Thorven, M.; Grahn, A.; Hedlund, K. O.; Johansson, H.; Wahlfrid, C.; Larson, G.; Svensson, L., A homozygous nonsense mutation (428G-->A) in the human secretor (FUT2) gene provides resistance to symptomatic norovirus (GGII) infections. *J Virol* **2005,** *79* (24), 15351-5.
55. Larsson, M. M.; Rydell, G. E. P.; Grahn, A.; Rodríguez-Díaz, J.; Åkerlind, B.; Hutson, A. M.; Estes, M. K.; Larson, G.; Svensson, L., Antibody Prevalence and Titer to Norovirus (Genogroup II) Correlate with Secretor (FUT2) but Not with ABO Phenotype or Lewis (FUT3) Genotype. *Journal of Infectious Diseases* **2006,** *194* (10), 1422-1427.

TABLES

Table 1. Summary of successful milk collections as a function of location and lactation month**.**

	Lactation Month																							
Location	-1	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	21	24	26	Total
Argentina			1	1		2	1	3	2		3	3	2	2											20
Bangladesh		15	15	14	12	12	12	12																	92
Bolivia		1	1	3	7	4	1	6	4	2	5	3	2	1	7	2	3	2	3	1	2			1	61
Boston				2	1	4	2	2	2	1	1		4	1											20
Brazil			22	18	14	20	18	17																	109
Davis	8	59	40	31	37	38		30					1	11											255
Gambia			33			33	33																		99
India			42	42	42	42	40	39																	247
Malawi			73		84			652																	809
Namibia			3	2		1			1	2					1	1				1					12
Nepal			1	3		3	2	4	4	2	3	2			1							1	2		28
Perth		28	29																						57
Peru		35	36	33	34	32	31	32																	233
Philippines		5	1	1	2	2	2	1	2	2	1	1	2	1											18
Poland			1	2	1	4	2	2	4	1		3		2	1										23
South Africa			26	26	27	26	25	26																	151
Total	8	143	324	178	261	223	169	816	19	10	13	12	11	18	10	3	3	2	3	2	2	1	2	1	2234

[1]Values presented correspond to the number of samples collected as a function of geographical location and month of lactation.

Table 2. The 60 most common HMO structures in S+ and S- milks across all study sites.

Neutral Mass (Da)	**Composition** (Hex_HexNAc_Fuc_Neu5Ac)	**HMO**	**S- type milk mean**	**S+ type milk mean**	***P*-value**
490.19	2010	2'FL	0.002 ±0.005	0.07 ±0.04	<0.0001
636.24	2020	LDFT	0.001 ± 0.003	0.04 ± 0.03	<0.0001
709.26	3100	LNT + LNnT	0.2 ± 0.09	0.1 ± 0.06	<0.0001
855.32	3110	LNFP II	0.04 ± 0.03	0.02 ± 0.02	<0.0001
855.32	3110	LNFP I + LNFP III	0.02 ± 0.02	0.06 ± 0.04	<0.0001
1074.39	4200	LNH	0.009 ± 0.01	0.009 ± 0.007	0.06
1074.39	4200	LNnH	0.009 ± 0.01	0.01 ± 0.01	<0.0001
1074.39	4200	p-LNH	0.004 ± 0.007	0.004 ± 0.005	<0.0001
1220.45	4210	MFpLNH IV	0.02 ± 0.01	0.02 ± 0.009	0.1
1220.45	4210	412Oa	0.007 ± 0.01	0.003 ± 0.006	<0.0001
1220.45	4210	MFLNH III + MFLNH I	0.03 ± 0.02	0.02 ± 0.01	0.02
1220.45	4210	IFLNH III	0.008 ± 0.006	0.01 ± 0.006	<0.0001
1220.45	4210	IFLNH I	0.001 ± 0.003	0.004 ± 0.004	<0.0001
1366.51	4220	DFpLNH II	0.01 ± 0.007	0.01 ± 0.006	<0.0001
1366.51	4220	DFLNH b	0.02 ± 0.01	0.01 ± 0.007	<0.0001
1366.51	4220	DFLNHa	0.001 ± 0.002	0.01 ± 0.01	<0.0001
1512.57	4230	TFLNH	0.004 ± 0.004	0.006 ± 0.005	<0.0001
1585.58	5310	5130a	0.006 ± 0.006	0.004 ± 0.003	<0.0001
1585.58	5310	F-LNO	0.004 ± 0.003	0.003 ± 0.002	0.002
1731.64	5320	DFLNO I	0.007 ± 0.005	0.003 ± 0.003	<0.0001
1731.64	5320	DFLNnO II	0.004 ± 0.005	0.003 ± 0.003	0.9
1731.64	5320	5230a + DFLNnO I/DFLNO II	0.004 ± 0.003	0.006 ± 0.004	<0.0001

635.22	2001	6'SL	0.003 ± 0.003	0.002 ± 0.003	<0.0001
635.22	2001	3'SL	0.01 ± 0.008	0.01 ± 0.007	0.4
1000.36	3101	LSTc + LSTb	0.03 ± 0.02	0.03 ± 0.01	<0.0001
1000.36	3101	LSTa	0.003 ± 0.002	0.002 ± 0.002	<0.0001
1365.49	4201	S-LNH	0.003 ± 0.004	0.002 ± 0.002	<0.0001
1365.49	4201	4021a + S-LNnH II	0.004 ± 0.005	0.006 ± 0.005	<0.0001
490.19	2010	%2'FL	0.4 ± 0.9	11.1 ± 4.6	<0.0001
636.24	2020	%LDFT	0.2 ± 0.5	5.8 ± 5.1	<0.0001
709.26	3100	%LNT + LNnT	30.3 ± 10	21.7 ± 5.9	<0.0001
855.32	3110	%LNFP II	6.1 ± 5	2.9 ± 2.6	<0.0001
855.32	3110	%LNFP I + LNFP III	3.8 ± 2.5	8.2 ± 5	<0.0001
1074.39	4200	%LNH	1.5 ± 1.6	1.2 ± 0.9	0.5
1074.39	4200	%LNnH	1.4 ± 1.8	1.7 ± 1.4	<0.0001
1074.39	4200	%p-LNH	0.7 ± 1.2	0.6 ± 0.7	<0.0001
1220.45	4210	%MFpLNH IV	2.8 ± 1.5	2.3 ± 1.2	<0.0001
1220.45	4210	%412Oa	1.1 ± 1.7	0.6 ± 1	<0.0001
1220.45	4210	%MFLNH III + MFLNH I	4.3 ± 2.6	3.3 ± 1.8	<0.0001
1220.45	4210	%IFLNH III	1.3 ± 0.9	1.4 ± 0.9	0.007
1220.45	4210	%IFLNH I	0.2 ± 0.5	0.5 ± 0.6	<0.0001
1366.51	4220	%DFpLNH II	2.2 ± 1.3	1.5 ± 0.8	<0.0001
1366.51	4220	%DFLNH b	3.3 ± 2.6	1.6 ± 1.2	<0.0001
1366.51	4220	%DFLNHa	0.2 ± 0.3	2 ± 1.8	<0.0001
1512.57	4230	%TFLNH	0.7 ± 0.6	0.9 ± 0.7	<0.0001
1585.58	5310	%5130a	0.9 ± 0.8	0.5 ± 0.4	<0.0001
1585.58	5310	%F-LNO	0.6 ± 0.4	0.4 ± 0.3	<0.0001
1731.64	5320	%DFLNO I	1.1 ± 0.7	0.5 ± 0.4	<0.0001

1731.64	5320	%DFLNnO II	0.7 ± 0.7	0.5 ± 0.4	0.0007
1731.64	5320	%5230a + DFLNnO I/DFLNO II	0.7 ± 0.5	0.8 ± 0.4	<0.0001
635.22	2001	%6'SL	0.5 ± 0.5	0.3 ± 0.4	<0.0001
635.22	2001	%3'SL	2.6 ± 1.7	2.2 ± 1.4	<0.0001
1000.36	3101	%LSTc + LSTb	5.2 ± 2	3.8 ± 1.6	<0.0001
1000.36	3101	%LSTa	0.5 ± 0.3	0.3 ± 0.3	<0.0001
1365.49	4201	%S-LNH	0.5 ± 0.5	0.3 ± 0.3	<0.0001
1365.49	4201	%4021a + S-LNnH II	0.6 ± 0.7	0.8 ± 0.6	<0.0001

[1]Values are presented as Mean Abundance ± SD. All data collected was used for this analysis including data from the same mother at different time points of lactation. Monosaccharide composition represented by four-digit code (Hex_HexNAc_Fuc_Neu5Ac). Common HMO abbreviations were used to name oligosaccharides. Oligosaccharides with two compound names are isomers that were difficult to resolve chromatographically, the data presented are the sum of their combined abundances. Oligosaccharide names proceeding a percentage indicates that the values presented are mean relative abundances. P values were obtaining using Mann-Whitney tests.

Table 3. Mean levels of total HMO, α(1,2)-fucosylated HMOs, total fucosylation, total sialylation and total undecorated HMOs by study site and secretor status

Country	Site Location	GDP per capita	Total HMO (Normalized Counts)		Total Fucosylation (%)		Total Sialylation (%)		Total Undecorated (%)		α(1,2)-Fucosylated (%)	
			B+	B-	B+	B-	B+	B-	B+	B-	B+	B-
Argentina	Namqom	14,508	0.71 ± 0.14	0.37 ± 0.07	69.04 ± 4.69	64.66 ± 5.62	10.40 ± 2.62	24.16 ± 2.86	23.47 ± 5.03	22.44 ± 5.07	24.48 ± 4.94	2.22 ± 0.53
Bolivia	Amazonian Lowlands	3,351	0.61 ± 0.13	0 ± 0	68.90 ± 6.39	0 ± 0	11.14 ± 2.69	0 ± 0	23.20 ± 6.39	0 ± 0	24.12 ± 8.87	0 ± 0
Brazil	Fortaleza, Ceará	9,881	0.59 ± 0.16	0.59 ± 0.12	65.18 ± 6.19	54.80 ± 10.01	11.74 ± 3.64	14.76 ± 5.66	27.68 ± 6.64	36.94 ± 11.55	20.51 ± 7.32	1.72 ± 1.01
Peru	Loreto	6,723	0.61 ± 0.19	0.48 ± 0.17	71.26 ± 4.79	64.15 ± 5.57	13.63 ± 3.22	15.99 ± 3.46	20.35 ± 4.92	26.74 ± 6.41	24.64 ± 7.65	1.98 ± 0.68
Australia	Perth	53,831	0.98 ± 0.24	0.94 ± 0.16	63.24 ± 4.58	58.02 ± 10.77	29.43 ± 6.73	34.17 ± 6.98	21.42 ± 6.08	26.32 ± 10.71	10.84 ± 5.06	2.61 ± 0.82
Bangladesh	Mirpur, Dhaka	1,564	0.58 ± 0.15	0.47 ± 0.15	65.30 ± 5.42	56.16 ± 12.40	11.44 ± 2.95	15.17 ± 3.19	27.76 ± 5.79	34.31 ± 12.66	18.45 ± 7.10	1.28 ± 0.90
India	Vellore, Tamil Nadu	1,980	0.62 ± 0.16	0.54 ± 0.16	62.94 ± 7.43	54.15 ± 9.78	10.78 ± 2.87	15.13 ± 4.15	29.92 ± 7.18	36.37 ± 9.16	19.32 ± 6.49	1.17 ± 2.07
Nepal	Nubri Valley	900	0.64 ± 0.13	0.63 ± 0.24	62.23 ± 6.64	49.66 ± 19.50	14.31 ± 2.88	19.19 ± 4.04	28.42 ± 6.12	38.71 ± 19.37	12.82 ± 3.68	0.99 ± 0.37
Philippines	Cebu	2,982	0.62 ± 0.09	0.51 ± 0.11	64.11 ± 9.35	60.90 ± 5.89	11.54 ± 2.34	16.11 ± 4.35	27.65 ± 9.94	28.74 ± 6.34	20.09 ± 5.37	2.24 ± 0.36
Gambia		673	0.58 ± 0.18	0.55 ± 0.17	53.63 ± 8.49	42.60 ± 16.81	13.85 ± 5.07	16.34 ± 5.76	36.62 ± 7.68	45.96 ± 16.20	18.49 ± 65.85	1.69 ± 1.47
Malawi	Mangochi district	357	0.72 ± 0.20	0.66 ± 0.20	62.01 ± 9.91	54.22 ± 16.59	12.21 ± 3.37	14.31 ± 5.61	29.92 ± 9.32	36.37 ± 15.11	20.54 ± 7.20	1.32 ± 0.74
Namibia	Omuhonga Basin	5,516	0.52 ± 0.25	0.47 ± 0.02	63.63 ± 10.08	57.52 ± 1.67	17.92 ± 4.0	25.43 ± 1.54	26.07 ± 8.41	26.93 ± 0.67	15.10 ± 10.04	1.11 ± 0.03
South Africa	Thohoyandou, Limpopo Province	6,120	0.56 ± 0.15	0.51 ± 0.18	60.15 ± 7.71	50.95 ± 10.45	13.16 ± 3.26	15.20 ± 3.08	31.32 ± 7.69	39.58 ± 9.85	17.88 ± 5.84	1.03 ± 0.74
Poland	Beskid Wyspowy Mtns	13,871	0.70 ± 0.22	0.54 ± 009	63.58 ± 6.60	51.46 ± 6.98	10.69 ± 3.66	17.35 ± 6.01	29.13 ± 7.25	36.68 ± 9.3	18.86 ± 5.06	1.20 ± 0.29
United States	Boston, Massachusetts	59,939	0.76 ± 0.25	0.58 ± 0.27	62.44 ± 3.88	58.95 ± 4.89	12.73 ± 3.47	16.45 ± 2.88	29.40 ± 5.35	30.87 ± 5.70	18.41 ± 6.77	2.87 ± 0.97
United States Davis	Davis, California	59,939	0.72 ± 0.24	0.65 ± 0.24	64.00 ± 5.86	59.92 ± 9.23	13.57 ± 4.40	19.41 ± 6.36	27.50 ± 5.98	29.10 ± 9.19	15.44 ± 7.19	2.11 ± 1.52

[1]Gross Domestic Product (GDP) values for each country are representative of the year of collection. Values are presented as mean ± SD. All data collected was used for this analysis including data from the same mother at different time points of lactation. P values were obtaining using Mann-Whitney tests.

FIGURE LEGENDS

Figure 1. Extracted ion chromatograms (EICs) displaying differences in abundances of HMO markers with α(1-2)-linked Fuc between mothers with S+ (–) and S- (–) milk from different locations around the world. Locations were chosen to represent different areas with both S+ and S- milk producers. Bolivia was also chosen to display EICs of 100% S+ milk producers. Monosaccharide composition of structures are given as Hex_HexNAc_Fuc_Neu5Ac and represented as glucose (●), galactose (●), N-acetylglucosamine (■), and fucose (▲). (**a**) EIC of 2'-fucosyllactose (2'FL) with *m/z* 491.19. (**b**) EIC of lactodifucotetraose (LDFT) with *m/z* 637.25. (**c**) EIC of isomers difucosyl-*parap*-*lacto*-*N*-hexaose (DFpLNH II), difucosyllacto-*N*-hexaose (b) (DFLNH b), difucosyllacto-*N*-hexaose (a) (DFLNHa), and difucosyllacto-*N*-hexaose (c) (DFLNHc) with *m/z* 684.27. (**d**) EIC of trifucosyllacto-*N*-hexaose (TFLNH) with *m/z* 757.29. (**e**) EIC of isomers lacto-*N*-fucopentaose II (LNFP II), lacto-*N*-fucopentaose I (LNFP I), and lacto-*N*-fucopentaose III (LNFP III) with *m/z* 856.33. (**f**) EIC of isomers fucosyl-*para*-lacto-*N*-hexaose (MFpLNH IV), 4120a, monofucosyllacto-*N*-hexaose III (MFLNH III), monofucosyllacto-*N*-hexaose I (MFLNH I), isomer III fucosyl-*para*-lacto-*N*-hexaose (IFLNH III), and isomer I fucosyl-*para*-lacto-*N*-hexaose (IFLNH I) with *m/z* 611.24.

Figure 2. Heatmap of relative abundances of the most common (60) HMOs across 15 geographically diverse sites. Comparison of abundances from mothers who are (A) S- producers and (B) S+ producers. HMO abundance values correspond to HPLC-qTOF MS spectral abundance normalized to the mean of the total abundance of counts from each sample. HMOs that were not baseline separated (resolution >1.5) were grouped together and labeled accordingly.

Figure 3. (A) Mean changes in HMO concentrations in breastmilk samples collected monthly during the first 6 months postpartum at sites with extensive longitudinal sampling. (B) Total HMO abundances as a function of location, lactation month, and secretor status. HMO abundance values correspond to HPLC-qTOF MS spectral abundance normalized to the mean of the total abundance of ion counts from each sample. N values correspond to the number of samples. Error bars represent standard deviation. P values were obtaining using Mann-Whitney tests with an α correction of α=0.05.

Figure 4. Variations in Fucosylated HMOs during Lactation Between Geographical Sites. (A) Mean relative abundance of total α(1,2)-fucose-containing HMOs in breastmilk samples as a function of lactation month (child's postnatal age) at sites with extensive longitudinal sampling. (B) Summed mean relative abundance of all fucose containing HMOs in the same samples (% Total Fucosylation). S- type producers (■) and S+ type producers (■). HMO abundance values correspond to HPLC-qTOF MS spectral abundance normalized to the mean of the total abundance of counts from each sample. Samples from multiple timepoints provided by a single mother as well as samples which only one time point was provided were included in this analysis. N values correspond to the number of samples. Error bars represent standard deviation. P values were obtaining using Mann-Whitney tests with an α correction of α=0.05.

Figure 5. Mean relative abundances of total sialic acid (Neu5Ac)-containing HMOs (% Total Sialylation) in breastmilk samples as a function of lactation month (child's postnatal age) at sites with extensive longitudinal sampling. Samples from multiple timepoints provided by a single

mother as well as samples from multiple mothers with one time point were all included in this analysis. HMO abundances corresponded to HPLC-qTOF MS abundances normalized to the mean of the total ion counts from each sample. Error bars represented standard deviation. P values were obtaining using Mann-Whitney tests with an α correction of α=0.05.

Figure 6. Proportion of samples tested from each study location that were S+ type (i.e. secretor mothers, ■) and S- type (non-secretor mothers, ■). Labels presented as Country-Number of Mothers. If samples from multiple timepoints were provided by the same mother, the secretor status determination was concluded based on the majority of her samples. If there was no 'majority milk type', she was excluded from the statistical analysis (one mother from Davis, USA, two mothers from Malawi, and three mothers from Perth, Australia).