

# *Bifidobacterium infantis* utilizes N-acetylglucosamine-containing human milk oligosaccharides as a nitrogen source

Shuqi Li <sup>a\*</sup>, Xiaomeng You <sup>a\*</sup>, Asha Rani <sup>a</sup>, Ezgi Özcan <sup>a</sup>, and David A. Sela <sup>a,b,c,d</sup>

<sup>a</sup>Department of Food Science, University of Massachusetts, Amherst, MA, USA; <sup>b</sup>Department of Microbiology, University of Massachusetts Amherst, Amherst, MA, USA; <sup>c</sup>Department of Nutrition, University of Massachusetts Amherst, Amherst, MA, USA; <sup>d</sup>Department of Microbiology & Physiological Systems and Center for Microbiome Research, University of Massachusetts Medical School, Worcester, MA, USA

## ABSTRACT

*Bifidobacterium longum* subsp. *infantis* (*B. infantis*) utilizes oligosaccharides secreted in human milk as a carbohydrate source. These human milk oligosaccharides (HMOs) integrate the nitrogenous residue N-acetylglucosamine (NAG), although HMO nitrogen utilization has not been described to date. Herein, we characterize the *B. infantis* nitrogen utilization phenotype on two NAG-containing HMO species, LNT and LNnT. This was characterized through *in vitro* growth kinetics, incorporation of isotopically labeled NAG nitrogen into the proteome, as well as modulation of intracellular 2-oxoglutarate levels while utilizing HMO nitrogen. Further support is provided by comparative transcriptomics and proteomics that identified global regulatory networks deployed during HMO nitrogen utilization. The aggregate data demonstrate that *B. infantis* strains utilize HMO nitrogen with the potential to significantly impact fundamental and clinical studies, as well as enable applications.

## ARTICLE HISTORY

Received 4 August 2022  
Revised 26 July 2023  
Accepted 1 August 2023

## KEYWORDS

human milk  
oligosaccharides; microbiota;  
bifidobacteria; nitrogen  
metabolism; 2-oxoglutarate

## Introduction

Microbial colonization of the infant gastrointestinal tract plays a fundamental role in infant health, which is a critical window for gut maturation<sup>1</sup> and immune system development<sup>2</sup>. Diet is widely recognized as an important determinant of gut microbiota colonization during infancy. Human milk is often the first food introduced to the infant and delivers oligosaccharides which transit intact to the colon as they are indigestible.<sup>3</sup> These human milk oligosaccharides (HMOs) are utilized by the small fraction of gut commensals that are capable of doing so, to date mostly characterized as specific taxa within the genus *Bifidobacterium*.<sup>4</sup> Accordingly, *Bifidobacterium longum* subsp. *infantis* (*B. infantis*) encodes a suite of HMO utilization genes typically arrayed in co-linear clusters, which is linked to its ability to utilize HMOs.<sup>5–7</sup>


Bacteria catabolize various nitrogenous molecules to ammonia which subsequently is synthesized into glutamate and glutamine, ultimately to incorporate nitrogen within the cell.<sup>8</sup> A common

bacterial system to accomplish this is the glutamine synthetase/glutamate synthetase (GS-GOGAT) pathway, which produces glutamate from ammonia and 2-oxoglutarate (2-OG).<sup>8</sup> In some bacteria, however, an alternate ammonium assimilation pathway proceeds through glutamate dehydrogenase (GDH). The production of glutamate through GDH also requires ammonia and 2-OG as substrates.<sup>8</sup> Ammonia is derived from nitrogenous molecules, and it is likely that *B. infantis* produces 2-OG from acetyl-CoA *via* citrate synthase, aconitase, and isocitrate dehydrogenase in its incomplete TCA cycle.<sup>9</sup> 2-OG typically resides at the intersection of carbohydrate and nitrogen metabolism, as it provides the carbon skeleton for nitrogen assimilatory reactions in GS-GOGAT or GDH pathways.<sup>10</sup>

Intracellular 2-OG levels fluctuate according to nitrogen availability in *Escherichia coli* subjected to nitrogen starvation, with intracellular 2-OG levels increasing fivefold after ammonium is removed from the growth medium<sup>11</sup> and decreases 20-fold when ammonium chloride is added.<sup>12</sup> This inverse

**CONTACT** David A. Sela  [davidsela@umass.edu](mailto:davidsela@umass.edu)  Department of Food Science, University of Massachusetts, Amherst, MA, USA

\*SL and XY contributed equally.

 Supplemental data for this article can be accessed online at <https://doi.org/10.1080/19490976.2023.2244721>.

© 2023 The Author(s). Published with license by Taylor & Francis Group, LLC.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

relationship is also exhibited by other microbes, including *Herbaspirillum seropedicae*,<sup>13</sup> *Cyanobacteria*,<sup>14</sup> and *Saccharomyces cerevisiae*.<sup>15</sup> However, in *Bacillus subtilis*, there was no relationship between 2-OG and nitrogen accessibility.<sup>16</sup> Due to the conservation of nitrogen assimilation pathways and the relationship between 2-OG and nitrogen in bacteria, we hypothesized that intracellular 2-OG levels in *B. infantis* are inversely related with how efficiently the cells utilize the nitrogen source. Therefore, 2-OG was investigated as a potential marker for the efficiency that *B. infantis* metabolizes nitrogen.

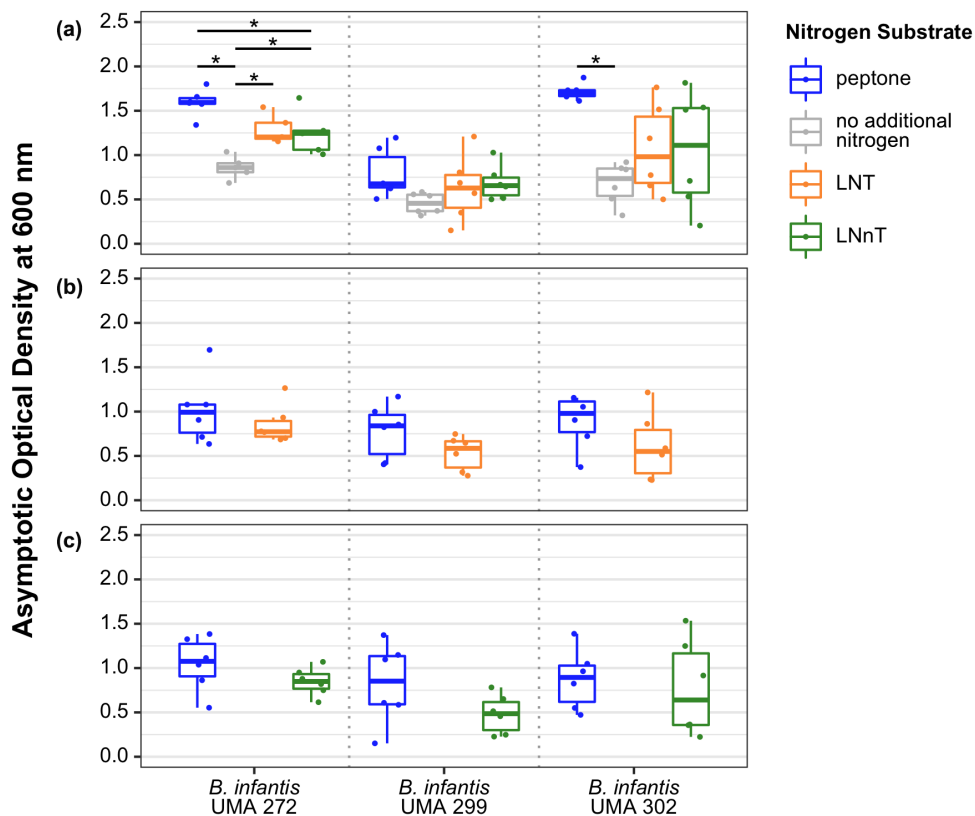
HMOs have been primarily studied as a prebiotic carbon source, although they incorporate nitrogen in the aminosugar residue *N*-acetyl glucosamine (NAG), and acidic HMO species incorporate *N*-acetylneuraminic acid (i.e., sialic acid).<sup>3</sup> Whether these nitrogen-containing HMOs and/or their constituents are utilized as a nitrogen source for *B. infantis* is currently unresolved. In this study, lacto-*N*-tetraose (LNT), the most abundant non-fucosylated neutral HMO species<sup>17</sup>, and its isomer,

which vary by a single glycosidic bond, lacto-*N*-neotetraose (LNnT), was evaluated as a primary nitrogen source for *B. infantis* along with features of their metabolism at the system level.

## Results

### *B. infantis* utilizes HMO and constituent NAG as a nitrogen source

*B. infantis* UMA 272 (parent strain ATCC 15697<sup>T</sup>) accumulated biomass while subsisting on LNT or LNnT as the primary nitrogen source which was significantly higher relative to the negative control (Figure 1A). Also, UMA 272 biomass while subsisting on LNnT was significantly lower compared to the positive control (i.e., peptone) (Figure 1A). Therefore, UMA 272 utilizes LNT or LNnT (albeit inefficiently) as a nitrogen source (Figure 1A). Despite differences in biomass accumulation, growth rates on LNT, LNnT, and peptone remained similar (Figure S1A). This suggests an



**Figure 1.** Terminal asymptotic optical density at 600<sub>nm</sub> while *B. infantis* utilizes LNT or LNnT as a nitrogen source. Data were grouped into three panels according to carbohydrate conditions: (A) lactose, (B) LNT, and (C) LNnT. Adjusted  $P < 0.05$  was designated significant.

equivalent preference for LNT, LNnT, and peptone as nitrogen sources.<sup>18</sup> UMA 302 (parent strain JCM 7009) exhibited a higher growth rate while subsisting on either of the HMOs relative to the negative control (Figure S1A), although UMA 302 does not display appreciable growth beyond the negative control (Figure 1A). UMA 299 (parent strain JCM 1260), in contrast, does not accumulate any biomass or alter its growth rate while subsisting on HMO nitrogen. UMA299 is not an efficient HMO carbohydrate consumer and was analyzed specifically to characterize strain diversity. UMA299 often yields phenotypes inconsistent with other *B. infantis* strains, in particular, as it pertains to HMO utilization.

UMA 272 and UMA 302 achieved similar final biomass when utilizing LNT or LNnT as both the carbohydrate and nitrogen sources. This is in comparison to the positive control that had peptone as the nitrogen source and the corresponding HMO as the carbohydrate source (Figure 1B, 1C). This may indicate that LNT or LNnT could be utilized by *B. infantis* as a nitrogen source equivalent to peptone under specific conditions (e.g., while utilizing LNT as a carbon source).

Interestingly, when LNT is metabolized as the primary nitrogen source, trace amounts of NAG were detected in the supernatant for all strains (Figure S2A). This indicates that soluble NAG is secreted during LNT utilization in contrast to LNnT. Alternatively, it is possible that LNT is more available to extracellular hydrolysis, although this is contrary to what is currently understood regarding *B. infantis* HMO metabolism which proceeds intracellularly. Regardless, there is a clear phenotypic difference attributable to structural variation between the two tetrasaccharides. LNT and LNnT diverge solely by a terminal linkage between D-galactose and NAG at the non-reducing end,  $\beta$ -(1-3) and  $\beta$ -(1-4), respectively.

NAG is the nitrogen-containing residue in LNT and LNnT, thus *B. infantis* was evaluated for its ability to utilize the nitrogen from this soluble amino sugar. Accordingly, all three strains utilize NAG as a nitrogen source (Figure S3A). The growth rate is mixed among the strains with UMA 272 exhibiting a similar rate between NAG and peptone and UMA 302 displaying a reduced rate on NAG. Interestingly, UMA 299 had an

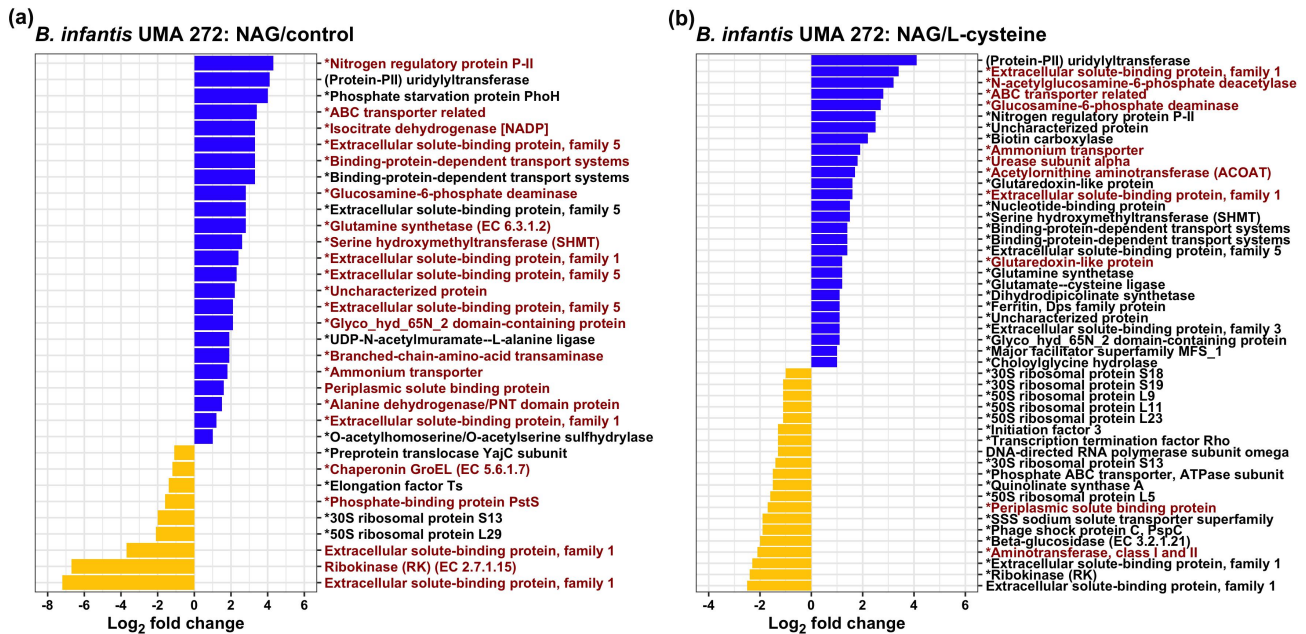
elevated growth rate on 1% NAG and a depressed rate on 2%. In addition, *B. infantis* UMA 299 and UMA 302 exhibited higher consumption rates on 2% NAG nitrogen compared to UMA 272 (Figure S2B).

### **<sup>15</sup>N labeled NAG nitrogen is incorporated into the *B. infantis* proteome**

In order to demonstrate the potential that HMO nitrogen incorporated in NAG is harnessed for anabolic reactions, <sup>15</sup>N labeled NAG was tracked within the *B. infantis* proteome with isotopically labeled proteins analyzed by mass spectrometry. As anticipated, biological replicates clustered together (Figure S4A, S4B). <sup>15</sup>N labeled NAG nitrogen was identified in Differentially Regulated Proteins (DRPs) in comparing NAG with either the complex control or L-cysteine (Figure 2, S5, S6). This approach efficiently quantified the fraction of the *B. infantis* proteome which only utilized <sup>15</sup>N labeled NAG nitrogen beyond the contribution from L-cysteine to address cysteine auxotrophy.

<sup>15</sup>N derived from NAG was broadly distributed throughout the proteome, including core nitrogen metabolic processes. This includes glutamine/glutamate biosynthesis and urea catabolic proteins in addition to carbohydrate metabolism, cell wall biogenesis, ABC transporters, SBPs, and arginine biosynthesis. 433 proteins were identified while *B. infantis* utilized NAG or labeled NAG as a nitrogen source (Figure S5A). <sup>15</sup>N was incorporated into 88% of the proteome (Figure S5A), other identified proteins (12%) included unlabeled NAG nitrogen only and thus were not included in the final incorporation analysis.

Furthermore, <sup>15</sup>N labeled L-cysteine was incorporated into the proteome in order to compare this with the <sup>15</sup>N labeled NAG proteome (Figure S5B) as L-cysteine must be added to the growth media in limiting amounts. A minority of <sup>15</sup>N labeled proteins 38.6% (n = 151) were shared among the NAG and L-cysteine proteomes with an additional 2.3% (n = 9) and 59.1% (n = 231) unique to <sup>15</sup>N-labeled L-cysteine and <sup>15</sup>N-labeled NAG proteomes, respectively (Figure S5B). Clearly, NAG nitrogen is catabolized as a nitrogen source and incorporated into the proteome. This is beyond the contribution that L-cysteine alone makes as a nitrogen



**Figure 2.** *B. infantis* differentially expressed proteins and genes while utilizing NAG as a nitrogen source. Differentially expressed genes during NAG utilization consistently within the transcriptome are colored in red. (A) differentially regulated proteins (DRPs) during NAG or complex nitrogen utilization. (B) differentially expressed proteins during NAG or L-cysteine utilization. Upregulated DRPs are colored in blue, and downregulated DRPs are colored in yellow. Asterisks indicate that the protein contain a  $^{15}\text{N}$  label. Log<sub>2</sub> fold change ( $\text{Log}_2\text{fc} > 1.5$ ,  $p < 0.1$  FDR corrected) values are indicated on x-axis.

source when it is added to address cysteine auxotrophy.

### Intracellular 2-OG concentrations increase while metabolizing LNT or LNnT as a nitrogen source

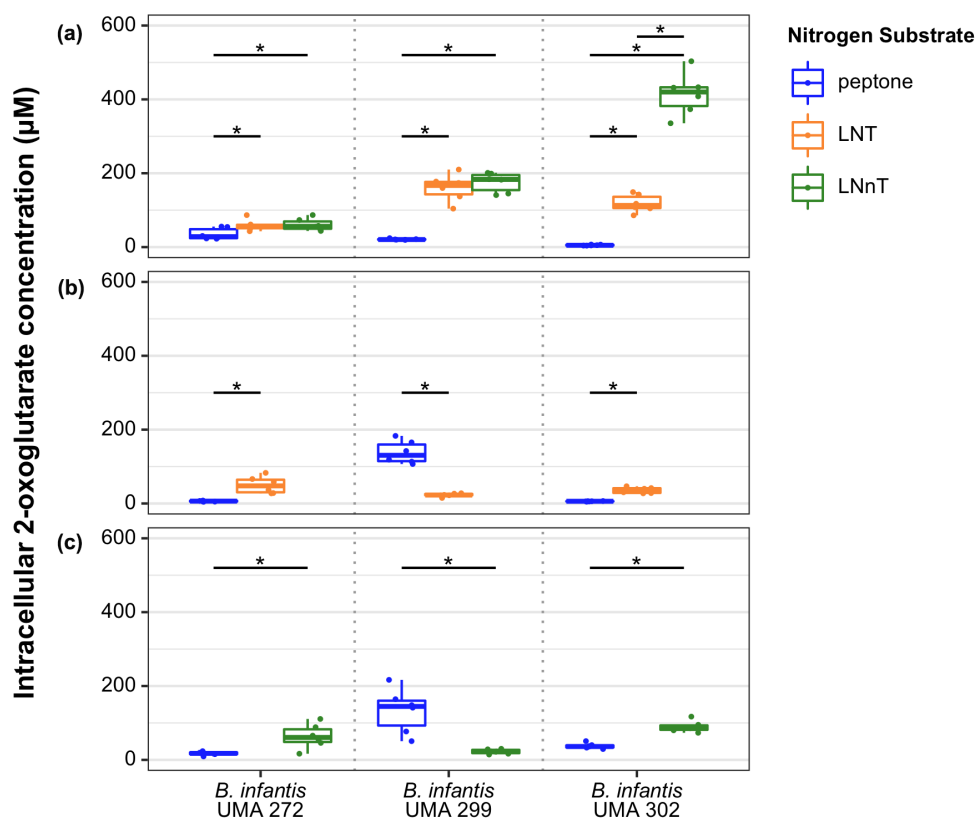
Intracellular 2-OG concentrations were quantified while *B. infantis* strains utilized lactose as a carbohydrate source with LNT or LNnT as a nitrogen source. As a result, intracellular 2-OG concentrations were significantly higher during LNT or LNnT nitrogen utilization relative to the control in *B. infantis* UMA 272, UMA 299, and UMA 302 (Figure 3A). 2-OG concentrations were similar regardless of LNT or LNnT nitrogen utilization for strains UMA 272 and UMA 299. Interestingly, UMA 302 exhibited significantly higher concentrations of 2-OG while utilizing LNnT relative to LNT (Figure 3A).

Similarly, when LNT or LNnT served as the carbohydrate source and a nitrogen source (i.e., no lactose), UMA 272 and UMA 302 exhibited significantly higher 2-OG concentrations compared to peptone nitrogen (Figure 3B, 3C). In general, *B. infantis* intracellular 2-OG levels increased when LNT or LNnT was utilized as nitrogen

source. The increased 2-OG concentrations are consistent with other bacterial systems that are enriched for this metabolism while utilizing an inefficient nitrogen substrate or while experiencing nitrogen starvation. Thus, there is evidence for an inverse relationship between 2-OG concentrations and nitrogen utilization efficiency in *B. infantis*.

### HMO carbohydrate metabolism does not impact intracellular 2-OG consistently

*B. infantis* utilizes LNT and LNnT as a carbohydrate source<sup>19</sup>, although it is unknown how this relates to intracellular 2-OG. Accordingly, it is known that 2-OG concentration is positively correlated with carbohydrate accessibility in *E. coli*<sup>10</sup>, which contrasts with its inverse relationship with nitrogen availability. Interestingly, intracellular 2-OG levels were significantly lower while UMA 272 utilized LNT or LNnT compared to lactose (Figure S7A). In contrast, 2-OG levels significantly increased while *B. infantis* UMA 299 utilized LNT or LNnT (Figure S7A). In general, *B. infantis* UMA 299 is a poor HMO utilizer<sup>6</sup>, although it is unclear how this impacts flux through 2-OG pathways while utilizing HMO as a carbon or nitrogen source. UMA 302



**Figure 3.** *B. infantis* intracellular 2-OG concentrations while *B. infantis* utilizes LNT, LNnT, or peptone nitrogen. Data are grouped into three panels according to carbohydrates fermented in the presence of a defined nitrogen: (A) lactose, (B) LNT, (C) LNnT. Adjusted  $P < 0.05$  was designated significant.

increased 2-OG concentrations while utilizing LNnT as the carbohydrate source, although there was not any difference observed for LNT (Figure S7A). Interestingly, when LNT or LNnT was defined as the carbohydrate source and nitrogen source (i.e., no peptone and no lactose), 2-OG levels in UMA 299 and UMA 302 were significantly lower compared to cells grown on lactose (carbon) and HMO (nitrogen) (Figure S7B, Figure S7C).

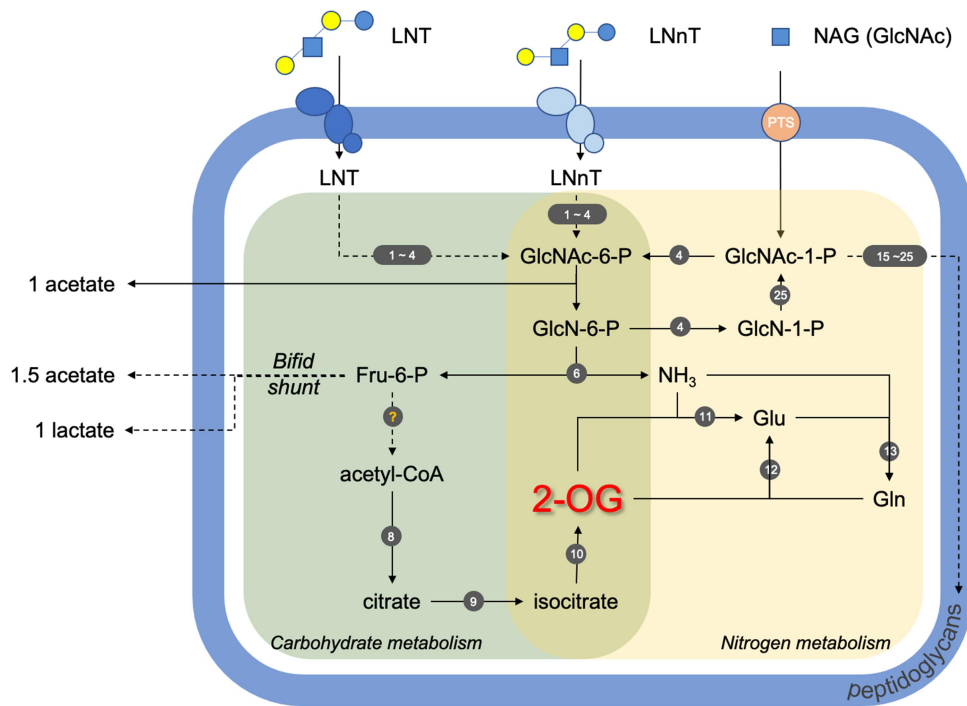
#### **Secreted acetate to lactate ratio increases while metabolizing HMO nitrogen**

*B. infantis* fermentation end-products during HMO nitrogen utilization were analyzed to characterize any deviation from typical metabolite profiles. HMO carbohydrate utilization has been previously studied, including secreted fermentation products.<sup>19</sup> Contrasting with HMO carbohydrate metabolism, *B. infantis* only secreted acetate and lactate during HMO nitrogen utilization as formate and ethanol were not

detected. *B. infantis* UMA 272 did not alter the acetate to lactate ratio relative to peptone nitrogen while metabolizing HMO nitrogen (Figure S2C). UMA 299 and UMA 302, however, increased the acetate to lactate ratio while utilizing LNT as a nitrogen source compared to peptone (Figure S2C). Interestingly, all strains exhibited a significant increase in the acetate to lactate ratio while NAG was utilized as a nitrogen source compared to peptone (Figure S2C). The increased ratio is likely due to additional acetate secretion through deacetylation of NAG and its residues within HMO.<sup>20</sup>

#### **The *B. infantis* transcriptome responds to HMO nitrogen utilization**

The global transcriptomes of *B. infantis* strains were resolved while utilizing LNT, LNnT, NAG, and pooled HMO nitrogen. As expected, low sample-to-sample distances between two biological replicates for each nitrogen substrate



**Figure 4.** Model of *B. infantis* HMO nitrogen utilization. Nitrogen metabolism is shaded in yellow, carbohydrate metabolism in green, and peptidoglycan synthesis pathways in purple. Solid arrows are for one-step reactions and dashed arrows for multi-step reactions. The genes involved in the pathways are designated by numbers (1–25) in dark gray circles and listed in Table S5. LNT, lacto-*N*-tetraose; LNnT, lacto-*N*-neotetraose; NAG or GlcNAc, *N*-acetyl glucosamine; GlcNAc-6-P, *N*-acetyl glucosamine-6-phosphate; GlcNAc-1-P, *N*-acetyl glucosamine-1-phosphate; GlcN-6-P, glucosamine-6-phosphate; GlcN-1-P, glucosamine-1-phosphate; Fru-6-P, fructose-6-phosphate; 2-OG, 2-oxoglutarate; Glu, glutamate; Gln, glutamine; UDP-GlcNAc, uridine diphosphate-*N*-acetyl glucosamine; UDP-MurNAc, uridine diphosphate-*N*-acetyl muramic acid.

were observed. Furthermore, the UMA 272 transcriptome while utilizing LNT nitrogen closely clustered with LNnT nitrogen metabolism (Figure S8A), which may be indicative of a similar *B. infantis* regulatory regime to harness nitrogen from these HMOs. *B. infantis* UMA 272 transcriptome profile during pooled HMO nitrogen utilization clustered with complex nitrogen (Figure S8A), which indicates that pooled HMO serves as a sufficient complex nitrogen source for *B. infantis* UMA 272. Interestingly, the three *B. infantis* strains exhibited similar dendrogram topologies (Figure S8B, S8C, S8D). The transcriptome during NAG nitrogen utilization closely clustered with L-cysteine and divergent to complex nitrogen group regardless of strain (Figure S8B, S8C, S8D). This may be due to nitrogen metabolism regulation being more similar for the monomeric NAG and L-cysteine in contrast to oligosaccharides or the complex control.

#### **NAG catabolism is upregulated during HMO and NAG nitrogen utilization**

Blon\_0881 (*nagB*) and Blon\_0882 (*nagA*) are essential for metabolizing NAG, and thus essential during HMO nitrogen utilization. Accordingly, Blon\_0881 and Blon\_0882 were significantly upregulated in the *B. infantis* UMA 272 transcriptome while LNT, LNnT, NAG, or pooled HMO was utilized as the primary nitrogen compared to the complex nitrogen control (Log<sub>2</sub>FC = 0.56–1.23,  $p < 0.001$ ) (Figure S9, Table S1). This was verified by qRT-PCR targeting these two genes (Figure S10). In addition, Blon\_0881 and Blon\_0882 were significantly upregulated in UMA 272 when NAG was supplied as the primary nitrogen compared to L-cysteine (Log<sub>2</sub>FC = 1.13,  $p < 0.001$ ) (Figure 2B, Figure S9, Table S1). *B. infantis* UMA 302 exhibited similar upregulation for Blon\_0881 and Blon\_0882 during NAG nitrogen utilization (Log<sub>2</sub>FC = 0.8–1.9,  $p < 0.001$ ) (Figure S9, Table S1). The proteins encoded by Blon\_0881 and Blon\_0882 exhibited similar upregulation in the proteome (Figure 2,

Table S2, S3). In summary, NAG catabolic genes and proteins are upregulated while *B. infantis* utilizes LNT, LNnT, NAG, and pooled HMO nitrogen.

### **Nitrogen assimilation pathways are upregulated while utilizing NAG nitrogen**

Gene and protein expression in pathways related to nitrogen metabolism were targeted, including the GS-GOGAT pathway, GDH pathway, and the PII nitrogen regulatory signaling (Figure 2, S9). In the GS-GOGAT pathway, glutamine synthetase (GS, Blon\_1893 and Blon\_0830) and glutamate synthetase (GOGAT, Blon\_1481, and Blon\_1482) facilitate the incorporation of nitrogen into glutamine and glutamate, respectively.<sup>10</sup> Blon\_1893 was significantly upregulated in UMA 302 (Log2FC = 5.04,  $p < 0.001$ ) by NAG nitrogen utilization compared to the complex nitrogen control and to L-cysteine (Log2FC = 3.55,  $p < 0.001$ ) (Figure S9, Table S1). Blon\_1893 and Blon\_0830 were significantly upregulated in *B. infantis* UMA 299 by NAG nitrogen compared to the complex nitrogen control (Log2FC = 2.90–3.82,  $p < 0.001$ ). Blon\_1481 and Blon\_1482 were significantly upregulated in *B. infantis* UMA 302 under NAG nitrogen compared to the complex nitrogen control (Log2FC > 2,  $p < 0.001$ ) and L-cysteine (Log2FC > 2,  $p < 0.001$ ) (Figure S9, Table S1). The proteomics were consistent with this observation as glutamine synthetase (Blon\_1893) was significantly upregulated by NAG compared to the complex nitrogen control in UMA 272 (Log2FC = 2.8,  $p < 0.05$ ) (Figure 2A, Table S2) and compared to L-cysteine (Log2FC = 1.2,  $p < 0.05$ ) (Figure 2B, Table S3). The other glutamate synthetase subunit (Blon\_1481) was upregulated by NAG as compared to control but not significantly (log2FC = 2.5,  $p > 0.1$ ) (data not shown). In general, the *B. infantis* GS-GOGAT system was upregulated while NAG was utilized as a nitrogen source.

In the GDH pathway, glutamate dehydrogenase (Blon\_0011) incorporates ammonium into 2-OG to generate glutamate. Blon\_0011 was significantly upregulated by NAG nitrogen compared to the complex nitrogen control in UMA 302 (Log2FC = 2.43,  $p < 0.001$ ); however, it is significantly downregulated in UMA 299 (Log2FC = -2.25,  $p < 0.001$ ) (Figure S9, Table S1) which may be a feature of this

strain's nitrogen utilization strategy favoring GS-GOGAT under these conditions. In addition, the protein glutamate dehydrogenase was not found to be significantly different under NAG nitrogen compared to control in proteomics profiles.

The PII nitrogen regulatory system involves an ammonium transporter (Blon\_0223), PII protein (Blon\_0224), and a uridylyltransferase (Blon\_0225).<sup>21</sup> Blon\_0223 was significantly upregulated in *B. infantis* UMA 272 during NAG nitrogen and pooled HMO nitrogen utilization relative to complex nitrogen control (Log2FC > 2,  $p < 0.01$ ) in UMA 299 and in UMA 302 compared to complex nitrogen control and L-cysteine (Log2FC > 2,  $p < 0.001$ ) (Figure S9, Table S1). Blon\_0224 was significantly upregulated in *B. infantis* UMA 272 and UMA 302 (Log2FC > 2,  $p < 0.001$ ) during NAG nitrogen utilization compared to the control (Figure S9, Table S1). Blon\_0225 was significantly upregulated in UMA 302 during NAG nitrogen utilization compared to the control (Log2FC > 2,  $p < 0.001$ ) (Figure S9, Table S1). The upregulated PII nitrogen regulatory gene levels were consistent with the proteomics results. In the proteomics profiles, an ammonium transporter (Blon\_0223), PII protein (Blon\_0224), and uridylyltransferase (Blon\_0225) were upregulated by NAG as compared to complex nitrogen control (log2FC > 1.8,  $p < 0.1$ ) and L-cysteine (log2FC > 1.9,  $p < 0.1$ ) (Figure 2, Table S2, S3). In general, the PII nitrogen regulatory system of *B. infantis* was upregulated by NAG nitrogen utilization, according to transcriptomics and proteomics results.

### **The *B. infantis* urease gene cluster is significantly induced by NAG nitrogen and HMO nitrogen**

*B. infantis* ATCC 17930 (UMA 299) utilizes urea as a nitrogen source, due to the expression of urease gene cluster.<sup>22,23</sup> The urease gene cluster (Blon\_0104–Blon\_0115) is conserved in *B. infantis* UMA 272 as well as other strains.<sup>5,6</sup> The transcriptome and proteome were interrogated to ascertain whether the urease gene cluster is co-regulated by another nitrogen source from human milk. Accordingly, Blon\_0104–Blon\_0115 were significantly upregulated in all strains while utilizing LNT, LNnT, and NAG nitrogen compared to complex nitrogen control (Log2FC > 2,  $p < 0.1$ )

or L-cysteine ( $\text{Log}_2\text{FC} > 2$ ,  $p < 0.1$ ) (Figure S9, Table S4). The urease subunit alpha (*ureC*, Blon\_0111) was significantly enriched in NAG nitrogen group compared to L-cysteine ( $\text{Log}_2\text{FC} = 1.8$ ,  $p < 0.05$ ) (Figure 2B, Table S3) in the proteome.

In addition, acetylornithine 5-aminotransferase (ACOAT) involved in arginine biosynthesis was also significantly upregulated by NAG utilization as compared to L-cysteine ( $\text{Log}_2\text{FC} = 1.7$ ,  $p < 0.05$ ) (Figure 2B, Table S3). The upregulation of glutamine synthetase, glutamate synthetase, PII system, urease, and ACOAT indicates that *B. infantis* co-regulates overlapping nitrogen utilization systems during HMO and potentially other milk nitrogen substrate metabolism.

## Discussion

LNT and LNnT contain *N*-acetylglucosamine and thus we hypothesized that *B. infantis* harness these HMOs as a nitrogen source.<sup>24</sup> Accordingly, *B. infantis* generates biomass on LNT or LNnT as a primary nitrogen source as well as incorporates NAG nitrogen into the proteome. The type strain *B. infantis* ATCC15697<sup>T</sup> (UMA 272) exhibits the growth phenotype most prominently, and all strains exhibit hallmarks of utilizing HMO nitrogen including NAG nitrogen metabolism, transcriptome responses, and modulating intracellular 2-OG levels. UMA 299 was evaluated as a control as this strain does not efficiently utilize HMO, and the growth phenotype is a reflection of this, at least in part. UMA 302 was selected to investigate strain-variation among HMO utilizers, as it diverges in its utilization phenotype from the type strain. Growth phenotypes are a crude assay and the negative control (i.e., no additional nitrogen) is confounded by the L-cysteine auxotroph<sup>25–27</sup> utilizing this essential amino acid as a nitrogen source as well, despite exhibiting metabolic hallmarks of HMO nitrogen utilization. The aggregate evidence demonstrates that *B. infantis* utilizes HMO as a nitrogen source with potential inter-strain variation. Figure 4 depicts a model of HMO nitrogen utilization that is supported by data presented herein.

We hypothesized that *B. infantis* intracellular 2-OG concentration is inversely related to the efficiency of nitrogen utilization as observed in other

microbial systems. Nitrogen efficiency is defined herein as nitrogen that is less accessible (e.g., requires more cellular resources) that will yield inefficient growth compared to more preferred nitrogen substrates. Accordingly, *B. infantis* UMA 272 increased intracellular 2-OG concentration while inefficiently utilizing LNT or LNnT as a primary nitrogen source. *B. infantis* UMA 299 and UMA 302 also increased intracellular 2-OG concentration while utilizing LNT or LNnT nitrogen. The same phenotype was not observed while utilizing these HMO tetrasaccharides as a carbohydrate source, further supporting the conclusion that *B. infantis* utilizes HMO nitrogen. Interestingly, UMA 302 exhibited significantly higher 2-OG concentration while metabolizing LNnT nitrogen relative to LNT, which is indicative of strain-specific phenotypes that are related to the single glycosidic linkage difference between LNT and LNnT<sup>19</sup>. It is clear that *B. infantis* intracellular 2-OG concentration is inversely related to nitrogen utilization efficiency through comparison to the complex nitrogen control. These results are consistent with *E. coli* intracellular 2-OG that fluctuates with nitrogen source or availability as 2-OG levels are higher in nitrogen-limiting conditions.<sup>10,28</sup>

Utilization of NAG-containing HMO as a nitrogen source is supported by the incorporation of <sup>15</sup>N labeled NAG into the *B. infantis* UMA 272 proteome, often into core nitrogen regulation processes. Accordingly, NAG deacetylase and deaminase were upregulated while utilizing LNT, LNnT, NAG, and pooled HMO as a nitrogen source as characterized in the transcriptome and proteome. Furthermore, isocitrate dehydrogenase, which facilitates the synthesis of 2-OG from isocitrate, was upregulated while utilizing HMO nitrogen. This is consistent with the molecular mechanisms involved in increasing intracellular 2-OG concentrations while utilizing HMO nitrogen. *B. infantis* nitrogen assimilation was upregulated in the transcriptome including glutamine synthetase and glutamate synthetase which incorporate nitrogen into the 2-OG carbon skeleton.

Interestingly, urease genes and proteins were upregulated under HMO nitrogen conditions. This suggests that the *B. infantis* urease gene cluster is co-regulated with ammonium assimilation and/or other nitrogen metabolism genes. Previous work



reported that urease genes are co-regulated with glutamate/glutamine synthetase by a master regulator in *Streptomyces coelicolor*<sup>29</sup> (GlnR) and in *Corynebacterium glutamicum*<sup>30</sup> (AmtR) nitrogen metabolism. That these two divergent species also belong to the phylum *Actinobacteria*, suggest that *B. infantis* may possess a similar master regulator.

This work represents a significant advance in the definition of the *B. infantis* HMO utilization phenotype(s) and nitrogen metabolism in general. Specifically, it is now understood that *B. infantis* strains utilize HMO nitrogen that is shunted into anabolic processes *via* intracellular 2-OG among other features in the proteome and transcriptome. This supports further research to evaluate these processes *in vivo* within the infant gut microbiome. Moreover, this enables innovation of strategies to impact infant health and nutrition early in life, as it pertains to nitrogen cycling within the microbiome. This includes the rational design of synthetic milk formulas and other applications that specifically target bifidobacterial metabolism beyond carbohydrates.

## Material and Methods

### Routine propagation of bacterial cultures

*Bifidobacterium longum subsp. infantis* UMA 272 (ATCC 15697), UMA 299 (JCM 1260), UMA 302 (JCM 7009) were routinely propagated on de Man, Rogosa, and Sharpe (MRS; BD Difco) and supplemented with L-cysteine hydrochloride (0.05%, *w/v*, Across Organics) at 37°C under anaerobic conditions (Coy Laboratory Products). The anaerobic gas condition was maintained at a composition of >5% CO<sub>2</sub>, >5% H<sub>2</sub>, and the remainder as N<sub>2</sub>.

### Bifidobacterial *in vitro* growth

Single colonies were precultured in MRS broth (0.05% *w/v* L-cysteine) overnight. This was inoculated 1% (*v/v*) to experimental media which contained a defined carbohydrate source and a defined nitrogen source. Carbohydrate sources were limited to lactose (2%), LNT (2%), and LNnT (2%), while nitrogen sources were limited to NAG (1% or 2%), LNT (2%), LNnT (2%), pooled HMO (2%), or complex nitrogen control (2%). The experimental

media shared a basic composition, including 0.2% (*w/v*) potassium phosphate, 0.02% (*w/v*) magnesium sulfate, 0.005% (*w/v*) manganese sulfate, 0.1% (*v/v*) Tween 80, and 0.03% (*w/v*) L-cysteine. 0.5% (*w/v*) sodium acetate was added when NAG was tested as the nitrogen source with lactose as the carbon source. The complex nitrogen control contained either 2% peptone or a combination of 1% peptone, 0.5% yeast extract, and 0.5% ammonium citrate. Each experimental group was evaluated in six biological replicates (at least) with technical triplicates. NAG was evaluated in biological triplicates. Growth kinetics were determined using the Growthcurver<sup>31</sup> package in R<sup>32</sup>. The asymptotic optical density at 600<sub>nm</sub> (asym OD<sub>600nm</sub>) and the growth rate constant of the first growth phase were extracted from the results for comparisons.

### Quantification of the intracellular intermediate metabolite 2-OG

*B. infantis* was inoculated from MRS preculture (0.03% L-cysteine) to the modified MRS (0.03% L-cysteine supplemented) media of defined carbohydrate and nitrogen conditions, each at 2% (*w/v*). Carbohydrate conditions were lactose, LNT, or LNnT and nitrogen conditions were peptone, LNT, or LNnT. The conditions are further defined in Table S6. The bacterial cells were harvested mid-log with OD<sub>600nm</sub> measured by a Nanodrop 2000 (Thermo Fisher Scientific Inc.). Cells were harvested by centrifugation at 5000 rpm for 5 min. The resultant pellets were vortexed and incubated in 300 µL of 80:20 methanol:water over dry ice (−75°C) for 15 min (with 50 µL 0.1 mM <sup>13</sup>C<sub>5</sub>-labeled 2-OG water solution spiked in). These samples were centrifuged (5000 rpm, 5 min, 4°C), and the supernatant was collected. The extraction was repeated twice. The final extract was concentrated, redissolved in a solvent system (1 mL) matching the mobile phase condition used in the liquid chromatography, filtered (0.22 µm, Sartorius), and stored at 4°C before injection into a UPLC/Xevo-TQD QQQ-MS system (Waters). The LC-MS method was adapted from Bajad et al.<sup>33</sup> and validated for the instrument. Unlabeled 2-OG (Sigma-Aldrich) and <sup>13</sup>C<sub>5</sub>-labeled 2-OG (Cambridge Isotope Laboratories, Inc.) were used as standards.

Liquid chromatography was performed on an ACQUITY UPLC BEH Amide Column (130Å, 1.7 µm, 2.1 mm × 150 mm, Waters). Flow rate: 0.15 mL/min. Inject volume: 10 µL. Column temperature: room temperature (25 ± 2°C). The LC solvents used were Phase A: 20 mM ammonium acetate + 20 mM ammonium hydroxide in 95:5 water: acetonitrile, pH 9.45; Phase B: acetonitrile. The gradients were as follows:  $t = 0$  min, 85% B;  $t = 5$  min, 0% B;  $t = 10$  min, 0% B;  $t = 10.01$  min, 85% B;  $t = 20$  min, 85% B. Areas under peaks were used to interpret 2-OG concentration in samples. The retention time of both molecules was at  $6.6 \pm 0.57$  min of each run. Finally, the measured 2-OG concentration was normalized by the corresponding OD<sub>600nm</sub> at the time of harvest to yield the final 2-OG concentration.

#### **Characterization of bifidobacterial extracellular metabolites**

The concentration of lactate, acetate, formate, ethanol, and NAG in the spent media were determined by high-performance liquid chromatography (HPLC) analysis. Samples were prepared from the supernatants of the 96-well microplate, at early stationary phase under defined nitrogen conditions in mMRS media. Nitrogen sources were either 2% LNT, 2% LNnT, 2% NAG, or 2% peptone. Lactose was used as the carbohydrate source 2% (w/v) for all experimental groups. The cell-free supernatants were filtered through a 0.22 µm filter (Sartorius) and stored at -20°C until analysis. Substrate residues and organic acids were quantified using a Shimadzu HPLC system equipped with a Refractive Index Detector 20A (Shimadzu Corp.). The separation was carried out on an Aminex HPX-87 H column (7.8 mm ID × 300 mm, Bio Rad Laboratories) at 30°C. The mobile phase was 5 mM H<sub>2</sub>SO<sub>4</sub>, flow rate was 0.6 ml/min, and the injection volume was 25 µL. Lactate, acetate, and NAG concentrations were calculated from standard curves (0.5, 1, 5, 10, 20, and 50 mM). Four biological replicates were performed for each experimental condition.

#### **Quantitative real-time PCR to measure relative gene expression**

One ml samples were harvested at mid-exponential phase (OD<sub>600nm</sub> ~0.4–0.6), pelleted at 12,000 × g for 2 min, and stored in 1 ml Ambion RNAlater (Life Technologies). RNA extraction and cDNA conversion were performed as previously described.<sup>34</sup> Samples were centrifuged at 12,000 × g for 2 min to collect the cell pellet. The pellet was washed twice with PBS buffer to remove residual RNAlater and centrifuged at 12,000 × g for 2 min. Total RNA was extracted using Ambion RNAqueous-Mini kit (Life Technologies) according to the manufacturer's instructions. Cells suspended in lysis buffer were transferred to Lysing Matrix E tubes (MP Biomedicals LLC) to disrupt cell walls through bead beating at 5.5 m/s for 30 s twice using FastPrep 24 bead beater (MP Biomedicals). Total RNA was eluted in 50 µl of EB solution and immediately subjected to DNase treatment with the Ambion Turbo DNA-free (Invitrogen) using 1 µL of DNase I for 30 min. Subsequently, total RNA was converted to cDNA using the High-Capacity cDNA Reverse Transcription Kit (Applied Biosystems) according to the manufacturer's instructions. The resultant cDNA was quantified by a Nanodrop 2000 Spectrophotometer (Thermo Fisher Scientific Inc.). The qRT-PCR was performed on a 7500 Fast Real-Time PCR System (Applied Biosystems) with PowerUP SYBR Green Master Mix (Applied Biosystems, Foster City, CA) using 200 ng of input cDNA. The reaction conditions were optimized for the specific target locus. qRT-PCR primers were designed using the Primer3 software (Table S7; <http://frodo.wi.mit.edu>). The gene Blon\_0393, encoding a cysteinyl-tRNA synthetase was used as an endogenous control as previously.<sup>34,35</sup> Growth on lactose (2% wt/v) served as a reference condition for gene expression. Results were expressed as fold change relative to the reference. These experiments were conducted on biological triplicates, and triplicate technical measurements were performed.

Following DNase treatment, the absence of genomic DNA was confirmed using the total RNA as template by qRT-PCR targeting the endogenous control.

### **Transcriptome analysis in response to human milk nitrogen sources**

#### **RNA isolation and purification**

Bacteria were grown in mMRS media supplemented with pooled HMOs (2%, w/v), LNT (2%), LNnT (2%), NAG (2%), complex nitrogen (1% peptone, 0.5% yeast extract, 0.5% ammonium citrate), or L-cysteine (0.15%) as the nitrogen source, while 2% lactose as supplemented as the carbohydrate source. Bacteria cells were harvested at exponential phase and pelleted immediately by centrifugation at  $12,000 \times g$  for 1 min and resuspended in 1 ml of RNAlater Ambion (Thermo Fisher Scientific) and incubated at 4°C overnight and then stored at -80°C until further analysis. RNA isolation was performed using an Ambion RNAqueous™ kit (Thermo Fisher Scientific) following manufacturer's instructions. An additional step of bead-beating for cell disruption (FastPrep-24™ 5 G, MP Biomedicals Inc) was added to the RNA isolation procedure. Total RNA was immediately treated with the Turbo DNase free kit (Ambion, Thermo Fisher Scientific). Samples with an RNA integrity number (RIN) above 7 were selected for further processing, as checked by the High Sensitivity RNA analysis ScreenTape Assay (Agilent Technologies). DNA contamination was assessed by qRT-PCR using the Fast SYBR Green master mix (Applied Biosystems) with the primers: Uni334F, 5'-ACTCCTACGGGAGGCAGCAGT-3' and Uni514R, 5'-ATTACCGCGGCTGCTGGC-3'.

#### **Library construction and sequencing**

Total RNA was treated with Ribo-Zero rRNA Removal kit for Bacteria (Illumina) to remove ribosomal RNA according to manufacturer's instructions. The depleted RNA was purified with the RNeasy MinElute Cleanup kit (Qiagen). The rRNA depletion efficiency was evaluated with the High Sensitivity RNA analysis ScreenTape Assay (Agilent Technologies). Whole transcriptome libraries were constructed using TruSeq Stranded

mRNA Library Preparation Kit (Illumina) or NEBNext Ultra II Directional RNA library Prep Kit for Illumina (New England Biolabs Inc) following manufacturer's instructions. Barcoded library quality was assessed by DNA analysis ScreenTape Assay (Agilent Technologies) and quantified by Qubit dsDNA BR Assay (Life technology). Libraries were diluted and pooled to equimolar concentration and denatured immediately prior to sequencing on an Illumina NextSeq 500 sequencing system (150 cycles, paired-end sequencing, two runs). Raw reads and processed data were publicly deposited in the NCBI Gene Expression Omnibus database (<https://www.ncbi.nlm.nih.gov/geo/>) under accession number GSE209665 and GSE155078.

#### **Quality control, alignment, and gene counting**

Raw reads were uploaded to the Massachusetts Green High-Performance Computing Cluster used for all bioinformatic and statistical analyses unless specifically noted. The raw read quality was checked by FastQC<sup>36</sup> (v1.0.0). For libraries prepared by NEBNext Ultra II Directional RNA library Prep Kit, the sequencing adapters were removed by Trimmomatic<sup>37</sup> (v0.32). For libraries prepared by TruSeq Stranded mRNA Library Preparation Kit, the sequencing adapters were trimmed with the Illumina FASTQ generation pipeline. The quality-controlled and trimmed reads were then aligned to the *Bifidobacterium longum* subsp. *infantis* ATCC 15697<sup>T</sup> genome (NCBI accession number: NC\_011593.1) using bowtie2<sup>38</sup> (v2.1.0). The aligned reads were sorted by SAMtools<sup>39</sup> (v1.0) and counted against a specific genomic locus using HTSeq<sup>40</sup> (v0.6.1) for differential expression analysis.

#### **Differential gene expression analysis**

Differentially expressed gene analysis was performed using the R Package DESeq2<sup>35</sup> (v3.3.1). The gene count tables generated from HTSeq were used as inputs, and the inbuilt function was applied to identify and quantify the magnitude of differentially expressed genes. Whole transcriptome comparisons were analyzed by hierarchical clustering analysis. The results were visualized by DESeq2 and pheatmap<sup>41</sup> (R package, v1.0.12). Log2 fold-change (Log2FC) of gene expression was

calculated by inbuilt function in R Package DESeq2 and visualized using a heatmap. Statistically significant differences were considered with Benjamini–Hochberg adjusted p-value (FDR) < 0.05.

### **Proteomic analysis of bifidobacterial utilization of nitrogen**

#### **Proteome sample preparation**

Bacterial cultures were grown in mMRS supplemented with NAG or complex nitrogen source or L-cysteine (0.03% w/v). *B. infantis* UMA 272 overnight culture was inoculated (1%) into 2.5 ml of mMRS (0.03% L-cysteine) with a defined nitrogen source. A total of three biological replicates were included for proteomic analysis from each <sup>15</sup>N labeled and unlabeled nitrogen samples. Condition 1: Complex nitrogen control (2%), condition 2: <sup>15</sup>N-L-cysteine (0.03% w/v, HSCH<sub>2</sub>CH(\*NH<sub>2</sub>)COOH, Cambridge isotope laboratories, Inc.), or unlabeled L-cysteine (0.03% w/v), condition 3: unlabeled L-cysteine (0.03%) + unlabeled NAG (2%), or <sup>15</sup>N-L-cysteine (0.03%) + unlabeled NAG (2%) or unlabeled L-cysteine (0.03%) + <sup>15</sup>NAG (2%, C8H15\*NO6, Cambridge isotope laboratories, Inc.).

Cultures were harvested at mid-exponential phase and centrifuged at 12,000 × g for 3 min, washed 2× with PBS buffer and stored at –80°C until further analysis. The cell pellets from storage were suspended in 1X SDS lysis and solubilization buffer (Protifi LLC) then sonicated using Bioruptor pico (Diagenode) at high for 10–15 min (sonication cycle: 30 sec ON, 30 sec OFF). The cell lysate was clarified by centrifugation at 13,000 g for 8 min. Protein concentrations were measured using the Pierce BCA protein Assay kit (Thermo Fisher Scientific) on a Nanodrop spectrophotometer (Thermo Fisher Scientific). Samples were reduced, alkylated, and trypsin digestion, and clean-up was performed according to S-Trap micro universal MS sample prep kit (Protifi LLC). Briefly, each sample was mixed with digestion buffer containing mass spectrometry grade trypsin (Promega) at a 1:50 ratio (wt/wt) and was added to the micro-spin column. Tryptic peptides were then reconstituted and

purified with ZIPTIP® C18 pipette tips (Merck Millipore).<sup>42</sup>

#### **Mass spectrometry proteomics and data analysis**

Tryptic digests were analyzed using a Thermo Easy-nLC 1000 nanoLC system coupled to a Thermo Orbitrap Fusion mass spectrometer (Thermo Fisher Scientific) at the University of Massachusetts Amherst Mass Spectrometry Center.

Mass spectrometry data was analyzed using Peaks studio software (V.10.5, Bioinformatics solutions Inc.). To identify peptides and infer proteins, a search was performed against a series of protein databases using the multi-step database strategy as implemented in Peaks Studio. In step 1, the UniProt protein database ([www.uniprot.org](http://www.uniprot.org)) was searched using a *Homo sapiens* and *Mus musculus* filter along with a common contaminants database. Unmatched *de novo* tags from step 1 were queried against a database constructed from the UniProt database containing *Bifidobacterium longum* subsp. *infantis* ATCC 15697 (<https://www.uniprot.org/proteomes/UP000001360>, Proteome ID: UP000001360).<sup>5</sup> A mass tolerance of 15 ppm was allowed for precursor ion and 0.5 Da for-fragment ions, with the peptide hitting a threshold of ≥25 (–10logP). The Peaks software package was used for qualitative analysis with the following parameters added: a peptide identification filter of 1% FDR for peptide-to-spectrum matches (PSMs) was set, and protein identification was based on at least one unique peptide per protein. For the database search, the following parameters were applied: (i) trypsin as enzyme for digestion, (ii) up to three missed cleavages per peptide, (iii) carbamidomethyl cysteine as fixed, (iv) methionine oxidation/acetylation (N-term), and <sup>15</sup>N labeling on nitrogen as variable modifications. The FDR for peptide and protein identification was assessed using the target-decoy strategy by searching against the reverse database. Proteins and peptides were identified and quantified by at least two of the three replicates.

Multivariate analysis was carried out using MetaboAnalyst 4.0 ([www.metaboanalyst.ca](http://www.metaboanalyst.ca)).<sup>43</sup> In order to generate relative quantifications across samples abundances were normalized to the

central mean based on the total sum of peptide and protein abundances and were log transformed prior to the fold change ( $\text{Log}_2\text{FC}$ ) analysis. Differentially expressed proteins (DEPs) were defined as proteins with a fold change  $>2.0$  and with significance  $p > 0.1$  (FDR corrected) and were visualized using volcano plots. Proteins were annotated with GO terms and KEGG pathways to identify biological functions using UniProt ID mapping.

### **Statistical analysis for growth kinetics, qRT-PCR, and metabolites quantifications**

Bacterial growth biomass, growth rate, and relative gene expression from qRT-PCR were subjected to one-way ANOVA followed by Tukey's HSD Test. Acetate to lactate ratio data was subjected to a two-way ANOVA followed by Tukey's HSD Test. Statistical tests were performed using base R or GraphPad Prism (version 9.0.0 for Mac OS X, GraphPad Software). Plots were generated using GraphPad Prism and R package tidyverse<sup>44</sup>.

### **Acknowledgments**

The authors thank Cindy Kane for technical assistance. We thank Stephen Eyles for technical support at UMass Amherst Mass Spectrometry Core Facility. Ravi Ranjan is acknowledged for support at the UMass Amherst Genomics Resource Laboratory. We thank Liv Dedon (Sela Lab at UMass) for suggestions and critical review of the manuscript. David A. Mills (UC Davis) is acknowledged for providing pooled HMO. Glycom/DSM is acknowledged for providing in-kind support through donating purified HMOs used in this study.

### **Disclosure statement**

No potential conflict of interest was reported by the authors.

### **Funding**

This project was supported by NIH-NICHD grant R01 HD106554 (DAS). Partial fellowship support was provided by the Tan Fellowship to SL and the Stanley Charm Graduate Fellowship and PepsiCo Fellowship to XY.

### **ORCID**

Shuqi Li  <http://orcid.org/0000-0002-7626-4606>  
 Xiaomeng You  <http://orcid.org/0000-0003-2132-1404>  
 Asha Rani  <http://orcid.org/0000-0002-8556-4715>  
 Ezgi Özcan  <http://orcid.org/0000-0001-5365-1550>  
 David A. Sela  <http://orcid.org/0000-0002-3354-7470>

### **Data availability statement**

Transcriptome reads and processed data are deposited in the NCBI Gene Expression Omnibus database (<https://www.ncbi.nlm.nih.gov/geo/>) under accession numbers GSE209665 and GSE155078. Proteome data is deposited in the EMBL-EBI PRoteomics IDentifications Database (PRIDE) under Accession: PXD035809

### **Authors contribution**

DAS conceived of this study. SQL, XY, AR, and DAS designed experiments and analyzed data. SQL, XY, and AR executed the experimental plan. EÖ and XY quantified metabolic end-products. SQL, XY, AR, and DAS wrote, revised, and finalized the manuscript. All authors analyzed data, contributed to the manuscript, and approved the submitted version.

### **References**

1. Cilieborg MS, Boye M, Sangild PT. Bacterial Colonization and Gut Development in Preterm Neonates. *Early Hum Dev.* 2012;88(SUPPL. 1). doi:10.1016/j.earlhumdev.2011.12.027.
2. Chung H, Pamp SJ, Hill JA, Surana NK, Edelman SM, Troy EB, Reading NC, Villablanca EJ, Wang S, Mora JR, et al. Gut Immune Maturation Depends on Colonization with a Host-Specific Microbiota. *Cell.* 2012;149(7). doi:10.1016/j.cell.2012.04.037.
3. Bode L. Human Milk Oligosaccharides: Every Baby Needs a Sugar Mama. *Glycobiology.* 2012;22(9):1147–1162. doi:<https://doi.org/10.1093/glycob/cws074>.
4. German J, Freeman S, Lebrilla C, Mills D Human Milk Oligosaccharides: Evolution, Structures and Bioselectivity as Substrates for Intestinal Bacteria. In *Nestle Nutrition Workshop Series: Pediatric Program;* 2008; Vol. 62. 10.1159/000146322.
5. Sela DA; Chapman J; Adeuya A; Kim JH; Chen F; Whitehead TR; Lapidus A; Rokhsar DS; Lebrilla CB; German JB; Price NP; Richardson PM; Mills DA. The Genome Sequence of *Bifidobacterium Longum* Subsp. *Infantis* Reveals Adaptations for Milk Utilization within the Infant Microbiome. *Proc Natl Acad Sci U S A.* 2008;105(48):18964–18969. doi:10.1073/pnas.0809584105.

6. LoCascio RG; Desai P; Sela DA; Weimer B; Mills DA. Broad Conservation of Milk Utilization Genes in *Bifidobacterium Longum* Subsp. *Infantis* as Revealed by Comparative Genomic Hybridization. *Appl Environ Microb*. 2010;76(22):7373–7381. doi:10.1128/AEM.00675-10.
7. Duar RM, Casaburi G, Mitchell RD, Scofield LNC, Ortega Ramirez CA, Barile D, Henrick BM, Frese SA. Comparative Genome Analysis of *Bifidobacterium Longum* Subsp. *Infantis* Strains Reveals Variation in Human Milk Oligosaccharide Utilization Genes among Commercial Probiotics. *Nutrients*. 2020;12(11). doi:10.3390/nu12113247.
8. Merrick MJ, Edwards RA. Nitrogen Control in Bacteria. *Microbiol Rev*. 1995;59(4):604–622. doi:https://doi.org/10.1128/mr.59.4.604-622.1995.
9. Schell MA; Karmirantzou M; Snel B; Vilanova D; Berger B; Pessi G; Zwahlen MC; Desiere F; Bork P; Delley M; Pridmore RD; Arigoni F. The Genome Sequence of *Bifidobacterium Longum* Reflects Its Adaptation to the Human Gastrointestinal Tract. *Proc Natl Acad Sci U S A*. 2002;99(22):14422–14427. doi:10.1073/pnas.212527599.
10. Huergo LF; Dixon R. The Emergence of 2-Oxoglutarate as a Master Regulator Metabolite. *Microbiol Mol Biol R*. 2015;79(4):419–435. doi:10.1128/membr.00038-15.
11. Yan D, Lenz P, Hwa T. Overcoming Fluctuation and Leakage Problems in the Quantification of Intracellular 2-Oxoglutarate Levels in *Escherichia Coli*. *Appl Environ Microb*. 2011;77:19. doi:10.1128/AEM.05257-11.
12. Yuan J, Doucette CD, Fowler WU, Feng XJ, Piazza M, Rabitz HA, Wingreen NS, Rabinowitz JD. Metabolomics-Driven Quantitative Analysis of Ammonia Assimilation in *E. Coli*. *Molecular Systems Biology*. 2009;5:5. doi:10.1038/msb.2009.60.
13. Oliveira MAS, Gerhardt ECM, Huergo LF, Souza EM, Pedrosa FO, Chubatsu LS. 2-Oxoglutarate Levels Control Adenosine Nucleotide Binding by *Herbaspirillum Seropedicae* PII Proteins. *FEBS J*. 2015;282:24. doi:10.1111/febs.13542.
14. Zhang CC, Zhou CZ, Burnap RL, Peng L. Carbon/Nitrogen Metabolic Balance: Lessons from Cyanobacteria. *Trends In Plant Science*. 2018;23:1116–1130. doi:10.1016/j.tplants.2018.09.008.
15. Brauer MJ, Yuan J, Bennett BD, Lu W, Kimball E, Botstein D, Rabinowitz JD. Conservation of the Metabolomic Response to Starvation across Two Divergent Microbes. *Proc Natl Acad Sci U S A*. 2006;103(51). doi:10.1073/pnas.0609508103.
16. Chubukov V, Sauer U. Environmental Dependence of Stationary-Phase Metabolism in *Bacillus Subtilis* and *Escherichia Coli*. *Appl Environ Microbiol*. 2014;80(9):2901–2909. doi:10.1128/AEM.00061-14.
17. Thum C, Wall CR, Weiss GA, Wang W, Szeto IMY, Day L. Changes in Hmo Concentrations throughout Lactation: Influencing Factors, Health Effects and Opportunities. *Nutrients*. 2021;13(7):2272. doi:https://doi.org/10.3390/nu13072272.
18. Korem Kohanim Y, Levi D, Jona G, Towbin BD, Bren A, Alon U. A Bacterial Growth Law out of Steady State. *Cell Rep*. 2018;23(10):2891–2900. doi:10.1016/j.celrep.2018.05.007.
19. Özcan E, Sela DA. Inefficient Metabolism of the Human Milk Oligosaccharides Lacto-N-Tetraose and Lacto-N-Neotetraose Shifts *Bifidobacterium Longum* Subsp. *Infantis* Physiology. *Front Nutr*. 2018;5:5. doi:https://doi.org/10.3389/fnut.2018.00046.
20. Garrido D, Ruiz-Moyano S, Mills DA. Release and Utilization of N-Acetyl-d-Glucosamine from Human Milk Oligosaccharides by *Bifidobacterium Longum* Subsp. *Infantis*. *Anaerobe*. 2012;18(4):430–435. doi:10.1016/j.anaerobe.2012.04.012.
21. Amon J, Titgemeyer F, Burkovski A. Common patterns – unique features: nitrogen metabolism and regulation in Gram-positive bacteria. *FEMS Microbiol Rev*. 2010 July 1;34(4):588–605. doi:https://doi.org/10.1111/j.1574-6976.2010.00216.x. Blackwell Publishing Ltd.
22. Schimmel P, Kleinjans L, Bongers RS, Knol J, Belzer C. Breast Milk Urea as a Nitrogen Source for Urease Positive *Bifidobacterium Infantis*. *FEMS Microbiol Ecol*. 2021;97:3. doi:10.1093/femsec/fiab019.
23. You X, Rani A, Özcan E, Lyu Y, Sela DA. *Bifidobacterium longum* subsp. *infantis* utilizes human milk urea to recycle nitrogen within the infant gut microbiome. *Gut Microbes*. 2023 Jan-Dec;15(1):2192546. doi: 10.1080/19490976.2023.2192546. PMID: 36967532; PMCID: PMC10054289.
24. Sela DA, Mills DA. Nursing Our Microbiota: Molecular Linkages between *Bifidobacteria* and Milk Oligosaccharides. *Trends In Microbiology*. 2010;18(7):298–307. doi:https://doi.org/10.1016/j.tim.2010.03.008.
25. Ferrario C, Duranti S, Milani C, Mancabelli L, Lugli GA, Turrone F, Mangifesta M, Viappiani A, Ossiprandi MC, van Sinderen D, et al. Exploring Amino Acid Auxotrophy in *Bifidobacterium Bifidum* PRL2010. *Front Microbiol*. 2015;6(NOV). doi:10.3389/fmicb.2015.01331.
26. Wada M, Fukiya S, Suzuki A, Matsumoto N, Matsuo M, Yokota A. Methionine Utilization by *Bifidobacteria*: Possible Existence of a Reverse Transsulfuration Pathway. *Biosci Microbiota Food Health*. 2021;40(1):80–83. doi:https://doi.org/10.12938/BMFH.2020-031.
27. Schöpping M, Gaspar P, Neves AR, Franzén CJ, Zeidan AA. Identifying the Essential Nutritional Requirements of the Probiotic Bacteria *Bifidobacterium Animalis* and *Bifidobacterium Longum* through Genome-Scale Modeling. *Npj Syst Biol Appl*. 2021;7(1). doi:10.1038/s41540-021-00207-4.
28. Chubukov V, Gerosa L, Kochanowski K, Sauer U. Coordination of Microbial Metabolism. *Nat Rev Microbiol*. 2014;12(5):327–340. doi:https://doi.org/10.1038/nrmicro3238.

29. Wray LV, Fisher SH. The *Streptomyces Coelicolor* GlnR Gene Encodes a Protein Similar to Other Bacterial Response Regulators. *Gene*. 1993;130(1):145–150. doi:[https://doi.org/10.1016/0378-1119\(93\)90359-B](https://doi.org/10.1016/0378-1119(93)90359-B).
30. Walter B, Hänssler E, Kalinowski J, Burkovski A. Nitrogen Metabolism and Nitrogen Control in *Corynebacteria*: Variations of a Common Theme. *J Mol Microbiol Biotechnol*. 2006;12:1–2. doi:[10.1159/000096468](https://doi.org/10.1159/000096468).
31. Sprouffske K, Wagner A. Growthcurver: An R Package for Obtaining Interpretable Metrics from Microbial Growth Curves. *BMC Bioinform*. 2016;17(1):172. doi:[10.1186/s12859-016-1016-7](https://doi.org/10.1186/s12859-016-1016-7).
32. R Core Team. R: A Language and Environment for Statistical Computing. R Foundation For Statistical Computing, Vienna, Austria. 2020. <https://ropensci.org/blog/2021/11/16/how-to-cite-r-and-r-packages/>.
33. Bajad SU, Lu W, Kimball EH, Yuan J, Peterson C, Rabinowitz JD. Separation and Quantitation of Water Soluble Cellular Metabolites by Hydrophilic Interaction Chromatography-Tandem Mass Spectrometry. *J Chromatogr A*. 2006;1125(1):76–88. doi:[10.1016/j.chroma.2006.05.019](https://doi.org/10.1016/j.chroma.2006.05.019).
34. Garrido D, Ruiz-Moyano S, Lemay DG, Sela DA, German JB, Mills DA. Comparative Transcriptomics Reveals Key Differences in the Response to Milk Oligosaccharides of Infant Gut-Associated Bifidobacteria. *Sci Rep*. 2015;5:5. doi:[10.1038/srep13517](https://doi.org/10.1038/srep13517).
35. Love MI, Huber W, Anders S. Moderated Estimation of Fold Change and Dispersion for RNA-Seq Data with DESeq2. *Genome Biol*. 2014;15:12. doi:[10.1186/s13059-014-0550-8](https://doi.org/10.1186/s13059-014-0550-8).
36. Andrews S. FastQC: A Quality Control Tool for High Throughput Sequence Data. 2010.
37. Bolger AM, Lohse M, Usadel B. Trimmomatic: A Flexible Trimmer for Illumina Sequence Data. *Bioinformatics*. 2014;30:15. doi:[10.1093/bioinformatics/btu170](https://doi.org/10.1093/bioinformatics/btu170).
38. Langmead B, Salzberg SL. Fast Gapped-Read Alignment with Bowtie 2. *Nat Methods*. 2012;9(4):357–359. doi:[10.1038/nmeth.1923](https://doi.org/10.1038/nmeth.1923).
39. Danecek P, Bonfield JK, Liddle J, Marshall J, Ohan V, Pollard MO, Whitwham A, Keane T, McCarthy SA, Davies RM, et al. Twelve Years of SAMtools and BCFtools. *GigaScience*. 2021;10(2). doi:[10.1093/giga-science/giab008](https://doi.org/10.1093/giga-science/giab008).
40. Anders S, Pyl PT, Huber W. HTSeq-A Python Framework to Work with High-Throughput Sequencing Data. *Bioinformatics*. 2015;31(2):166–169. doi:[10.1093/bioinformatics/btu638](https://doi.org/10.1093/bioinformatics/btu638).
41. Kolde R, Kolde MR. Package ‘pheatmap’. R package. 2015;1(7): 790.
42. Dedon LR, Özcan E, Rani A, Sela DA. Bifidobacterium *Infantis* Metabolizes 2' Fucosyllactose-Derived and Free Fucose Through a Common Catabolic Pathway Resulting in 1,2-Propanediol Secretion. *Front nutrition*. 2020;7:7. doi:[10.3389/fnut.2020.583397](https://doi.org/10.3389/fnut.2020.583397).
43. Chong J, Soufan O, Li C, Caraus I, Li S, Bourque G, Wishart DS, Xia J. MetaboAnalyst 4.0: Towards More Transparent and Integrative Metabolomics Analysis. *Nucleic Acids Res*. 2018;46:W1. doi:[10.1093/nar/gky310](https://doi.org/10.1093/nar/gky310).
44. Wickham H, Averick M, Bryan J, Chang W, McGowan L, François R, Grolemund G, Hayes A, Henry L, Hester J, et al. Welcome to the Tidyverse. *J Open Source Softw*. 2019;4(43). doi:[10.21105/joss.01686](https://doi.org/10.21105/joss.01686).