Genomes of rumen bacteria encode atypical pathways for fermenting hexoses to short-chain fatty acids

Timothy J. Hackmann¹*, David Kamanda Ngugi², Jeffrey L. Firkins³, and Junyi Tao¹

¹Department of Animal Science, University of Florida, P.O. Box 110910, Gainesville, Florida 32611, USA. ²Red Sea Research Center, King Abdullah University of Science and Technology (KAUST), Thuwal 23955–6900, Saudi Arabia. ³Department of Animal Science, The Ohio State University, 2029 Fyffe Rd, Columbus, Ohio 43210, USA.

*For correspondence. E-mail thackmann@ufl.edu; Tel. (352) 775-1703; Fax (352) 392-5595

Running title: Atypical fermentation pathways
Originality-Significance Statement

Fermentation pathways for hexoses have been chiseled in the stone of textbooks for decades, with most sources rendering them as they appear in classic 1986 text by Gottschalk. However, our study finds these textbook pathways often cannot explain fermentation products formed by rumen bacteria. In the most extensive analysis of its kind, we reconstructed pathways for glucose fermentation from genomes of nearly 50 species and subspecies of bacteria. Forty-four percent of bacteria had pathways that were atypical or completely unprecedented, and seventeen percent even lacked a full Embden–Meyerhof–Parnas (EMP) pathway. This study shows that reconstruction of metabolic pathways could be incorrect if textbook pathways are used as a guide, and it calls for renewed efforts to delineate fermentation pathways biochemically.
Summary

Bacteria have been thought to follow only a few well-recognized biochemical pathways when fermenting glucose or other hexoses. These pathways have been chiseled in the stone of textbooks for decades, with most sources rendering them as they appear in the classic 1986 text by Gottschalk. Still, it is unclear how broadly these pathways apply, given that they were established and delineated biochemically with only a few model organisms. Here we show that well-recognized pathways often cannot explain fermentation products formed by bacteria. In the most extensive analysis of its kind, we reconstructed pathways for glucose fermentation from genomes of 48 species and subspecies of bacteria from one environment (the rumen). In total, 44% of these bacteria had atypical pathways, including several that are completely unprecedented for bacteria or any organism. In detail, 8% of bacteria had an atypical pathway for acetate formation; 21% for propionate or succinate formation; 6% for butyrate formation; and 33% had an atypical or incomplete Embden–Meyerhof–Parnas pathway. This study shows that reconstruction of metabolic pathways—a common goal of omics studies—could be incorrect if well-recognized pathways are used for reference. Further, it calls for renewed efforts to delineate fermentation pathways biochemically.
Introduction

Despite their phylogenetic diversity, fermentative bacteria are thought to follow only a few pathways when fermenting glucose or other hexoses to short-chain fatty acids (Gottschalk, 1986; White et al., 2012). For fermentation of glucose to acetate, for example, only two pathways are thought to predominate in heterotrophs. One pathway forms acetyl-CoA [in part through the Embden–Meyerhof–Parnas (EMP) pathway], then forms acetate from acetyl-CoA in two steps (involving phosphate acetyltransferase and acetate kinase). The second pathway forms acetyl-P (through the bifidobacterium pathway), then forms acetate from acetyl-P. These and most pathways were delineated decades ago. Most sources still render pathways as they appear in the classic text by Gottschalk (1986).

Some exceptions to typical pathways have arisen. For example, Selenomonas ruminantium has an atypical pathway for forming acetate from acetyl-CoA. This bacterium is missing activities for acetate kinase and phosphate acetyltransferase as found the typical pathway (Joyner and Baldwin, 1966; Melville et al., 1988). Acetate—CoA ligase (ADP-forming) has been suggested to substitute for the missing activities (Melville et al., 1988; Michel and Macy, 1990). As additional examples, Clostridium thermocellum LQ8 and Butyrivibrio proteoclasticus B316 have incomplete (atypical) EMP pathways because they lack activity for pyruvate kinase (Zhou et al., 2013; Olson et al., 2017) and enolase (phosphopyruvate hydratase) (Kelly et al., 2010), respectively. These exceptions call for more attention in determining if more atypical pathways exist.

Determining if atypical pathways exist is particularly significant for omics studies. These studies frequently reconstruct fermentation and other metabolic pathways, as testified by the
number of databases tailored for this metabolic reconstruction [e.g., (Caspi et al., 2016; Chen et al., 2017; Kanehisa et al., 2017)]. Completeness of reconstruction requires knowledge of all possible pathways (typical and atypical).

To determine if more atypical pathways exist (and their incidence), we reconstructed fermentation pathways from genomes of almost fifty species and subspecies of bacteria from one environment (the rumen). These bacteria are well-characterized—most are type strains—and yet we found several instances of atypical pathways. For example, we document a pathway for acetate formation not previously found in bacteria, and we found >15% of bacteria had an incomplete EMP pathway.

**Results**

**Reconstruction of pathways.** In our reconstruction of pathways for hexose fermentation, we used genomes for 48 cultured species and subspecies of rumen bacteria (Supporting Information Table S1). These bacteria represent all hexose-utilizers from the rumen that have 1) a genome sequence, 2) reported fermentation products, and 3) otherwise met our inclusion criteria (Experimental Procedures). The best-characterized strain (usually the type) was chosen for each species or subspecies. Both finished and unfinished (draft) genomes were included. The estimated completeness of the 48 genomes was 99.5 (0.85 standard deviation) %, with only one genome (Acetitomaculum ruminis 139B) less than 97% complete.

From the genomes of these strains, we reconstructed pathways forming the major fermentation products of rumen bacteria (Fig. 1). These fermentation products include formate, acetate, propionate, butyrate, lactate, succinate, ethanol, and hydrogen (Fig. 2 and Supporting Information Table S2). Pathways for other fermentation products (isobutyrate, valerate,
isovalerate, caproate) were not included in our reconstruction because the respective enzymes are generally not found in databases (e.g., KEGG, COG, pfam). Butanol was not included because it was reported as a product for a single bacterium only (see Supporting Information Table S2).

Pathways in Fig. 1 are written in condensed form, with each black solid arrow representing one of twenty-five subpathways. Subpathways are in turn comprised of several reactions as illustrated in Supporting Information Fig. S1. The enzymes that catalyze the reactions, along with the database IDs we searched, are reported in Supporting Information Table S3.

All 25 subpathways shown in Fig. 1 are encoded by at least one genome. Some subpathways have multiple variants (different sets of component reactions), and a total of 52 variants were encoded by at least one genome (Fig. 3, Supporting Information Table S4). In the main text, we focus on those subpathways and variants that form atypical fermentation pathways. Our reconstruction showed several pathways that are atypical, and some are completely unprecedented for bacteria. A total of 21 bacteria, or 44% of the total, had at least one atypical pathway as defined below (Fig. 3, Supporting Information Table S5).

**Conversion of glucose to pyruvate.** Because all bacteria in our analysis are hexoses utilizers, all forty-eight should encode the EMP pathway or another pathway converting glucose to pyruvate (see Fig. 1). Our reconstruction shows that thirty-five encode a typical EMP pathway. Of those thirty-five, three (Actinobacillus succinogenes 130Z, Basfia succiniciproducens JF4016, Blautia schinkii B) also encode the oxidative pentose phosphate pathways, and one (Lactobacillus ruminis RF1) also encodes the phosphoketolase pathway.
Five genomes (all *Bifidobacterium*) do not encode an EMP pathway. They encode the bifidobacterium pathway, oxidative pentose phosphate pathway, and phosphoketolase pathways instead.

No bacterium encoded any variant of the Entner-Doudoroff pathway (Supporting Information Fig. S1BG to BJ).

Four bacteria (*Anaeroplasma bactoclasticum* JR and most *Prevotella*) encode an atypical EMP pathway, which involves a pyrophosphate (PP$_i$)-dependent phosphofructokinase (diphosphate—fructose-6-phosphate 1-phosphotransferase) (Fig. 4B; reaction 2). A H$^+$-pyrophosphatase (reaction 4) generates the required PP$_i$. An additional four bacteria encode a similar pathway, but they do not encode a H+-pyrophosphatase. These bacteria would need to generate PP$_i$ from another source (e.g., biosynthesis). All eight bacteria also encode an ATP-dependent phosphofructokinase (6-phosphofructokinase) (Fig. 4A; reaction 1), and thus the typical EMP pathway is also present.

Eight bacteria have an incomplete EMP pathway. Of those eight, five (*Acetitomaculum ruminis* 139B, two *Butyrivibrio, Oscillibacter ruminantium* GH1, *Prevotella brevis* GA33) are missing enolase (Fig. 5A). The methylglyoxal pathway (Supporting Information Fig. S1BK to BO) bypasses enolase, and conceptually it could serve as an alternative. However, only one bacterium (*Ba. succiniciproducens* JF4016) would appear to encode it, and this bacterium has a full EMP pathway.

Of the eight bacteria with an incomplete EMP pathway, three (*Fibrobacter succinogenes* subspecies, *Treponema saccharophilum* PB) are missing pyruvate kinase (see Fig. 5B). The malate shunt may be an alternative in these bacteria (see Fig. 1). For *T. saccharophilum* PB,
pyruvate, phosphate dikinase could serve as another route (Supporting Information Fig. S1M), though it would have to operate in reverse of its normal direction. Conceptually, glucose or mannose phosphotransferase (PTS) systems could also substitute for missing pyruvate kinase (Supporting Information Fig. S1N,O), but none of the three bacteria had these systems.

**Acetate formation.** Thirty-six of our analyzed bacteria are observed to form acetate (see Fig. 2 and Supporting Information Table S2). Our reconstruction shows most (n = 27) bacteria form acetate with the typical pathway that 1) forms acetyl-CoA in part via the EMP pathway then 2) converts acetyl-CoA to acetate (Fig. 6A). Some bacteria (*Bifidobacterium; n = 5*) form acetate with another typical pathway that 1) forms acetyl-P via the bifidobacterium pathway then 1) converts acetyl-P to acetate (Fig. 1).

All *Selenomonas* (n = 3) and *Mitsuokella jalaludinii M 9*, however, lack these typical pathways for forming acetate. They form acetyl-CoA, but they do not have typical reactions for converting acetyl-CoA to acetate (Fig. 6A)—i.e., those catalyzed by phosphate acetyltransferase (reaction 1) and acetate kinase (reaction 2). Instead of using these reactions, they appear to use atypical reactions (Fig. 6B)—those catalyzed by succinyl-CoA:acetate CoA-transferase (reaction 3) and succinate—CoA ligase (ADP-forming) (reaction 4). No other bacterium examined encodes both of these atypical reactions.

If operated in reverse of its normal direction, acetate—CoA ligase could serve as another route for forming acetate from acetyl-CoA (Mai and Adams, 1996; James et al., 2016). However, neither ADP (Supporting Information Fig. S1Z) nor AMP-dependent (Fig. S1AA) ones were present in the genomes of *Selenomonas* and *Mi. jalaludinii M 9*. The phosphoketolase pathway also forms acetyl-P (Fig. 1A, Fig. S1I), but *Selenomonas* and *Mi. jalaludinii M 9* do not
have this pathway. Pyruvate dehydrogenase (quinone) forms acetate directly from pyruvate (Supporting Information Fig. S1BF), but no bacterium encodes this enzyme.

Succinate (and propionate) formation. Fourteen of our analyzed bacteria are observed to form succinate, and four form propionate (see Table S2 and Supporting Information Table S2). Only two species (Ac. succinogenes 130Z, Ba. succiniciproducens JF4016) encode one typical pathway (Fig. 7A), in which reduced NAD (NAD$_{\text{red}}$) is produced by pyruvate dehydrogenase (reaction 1) and used to reduce fumarate to succinate (reaction 2).

All Prevotella and Selenomonas ($n = 7$) do not encode the typical pathway above, and they instead encode a pathway shown in Fig. 7B. Instead of a pyruvate dehydrogenase that forms NAD$_{\text{red}}$, they encode a pyruvate:ferredoxin oxidoreductase that forms reduced ferredoxin (Fd$_{\text{red}}$; reaction 4). The Fd$_{\text{red}}$ is oxidized by the ion pump Rnf [ferredoxin-NAD$^+$ oxidoreductase (Na$^+$-transporting); reaction 5]. This generates NAD$_{\text{red}}$ and an electrochemical potential ($\Delta\mu$Na$^+$). The NAD$_{\text{red}}$ in turn is used to reduce fumarate to succinate (reaction 2). This pathway is atypical because Rnf balances fermentation and conserves energy, whereas the typical pathway does not involve Rnf.

An electrochemical potential is formed not only by Rnf, but also by NADH dehydrogenase (Figs. 7 and 8). The specific NADH dehydrogenase (Nqr, Ndh) varies by species, as does the ion (H$^+$ or Na$^+$) it translocates and the number of ions translocated (Fig. 8B,C). Together with ATP synthase, Rnf and NADH dehydrogenase generate between 1/4 to 3/4 ATP per fumarate via electron transport phosphorylation (Fig. 8B,C).

Two more bacteria (Porphyromonas levii LEV, Succinivibrio dextrinosolvens 24) follow the same atypical pathway as Prevotella and Selenomonas, except they lack ATP synthase.
Consequently, the electrochemical potential generated by Rnf and fumarate reductase/succinate dehydrogenase cannot drive ATP synthesis, though it could drive secondary active transport or motility.

*Ba. succiniciproducens* JF4016 also follows the same atypical pathway as *Prevotella* and *Selenomonas*, except acetate formation is not reported for the type strain. For other *Basfia*, acetate (and lactate) formation is reported (Scholten et al., 2009).

In addition to the atypical pathway, *Selenomonas* and other bacteria have typical pathways that use hydrogen or formate, not NAD$_{\text{red}}$, to reduce fumarate to succinate (Supporting Information Fig. S1AR,AS). However, *Prevotella* lack these typical pathways, emphasizing the importance of the atypical pathway.

*Prevotella* (as well as some *Selenomonas* and other bacteria) have another pathway for forming succinate, and it forms formate as an additional end product. It is similar to the pathway in Fig. 7B, except pyruvate formate-lyase (formate C-acetyltransferase) (Supporting Information Fig. S1W) substitutes for pyruvate:ferredoxin oxidoreductase (Supporting Information Fig. S1U), and so formate is formed instead of Fd$_{\text{red}}$. However, formation of only small amounts of formate are observed for most *Prevotella* (Supporting Information Table S2), and thus the pathway in Fig. 7B would be responsible for formation of at least some succinate.

All *F. succinogenes* subspecies ($n = 2$) encode an atypical pathway similar to *Prevotella* and *Selenomonas*, but they do not encode Rnf. In absence of Rnf, it is not clear how these bacteria convert Fd$_{\text{red}}$ to NAD$_{\text{red}}$ for NADH dehydrogenase. They lack typical pathways that use formate (or hydrogen) to reduce fumarate to succinate, and they form too little formate to
completely substitute for generating $\text{Fd}_{\text{red}}$ (see Supporting Information Table S2). Thus, the pathway used by these subspecies would appear atypical but is unclear.

**Butyrate formation.** Eleven of our analyzed bacteria are observed to form butyrate (see Table S2 and Supporting Information Table S2). Most ($n = 6$) form butyrate with a typical pathway (Supporting Information Fig. S2A). However, as in a previous study (Hackmann and Firkins, 2015), we found that some *Butyrivibrio* and *Pseudobutyrivibrio* ($n = 3$) encode an atypical pathway for forming butyrate (Supporting Information Fig. S2B). This atypical pathway involves $\text{Ech}$ hydrogenase, which balances fermentation by oxidizing $\text{Fd}_{\text{red}}$. It also is an ion pump, generates an electrochemical potential, and conserves energy.

For two bacteria (*Eubacterium ruminantium* GA195, *Megasphaera elsdenii* LC-1), the pathway used to form butyrate is not clear. These bacteria lack a complete Rnf, which is needed to balance the fermentation. Rnf is not needed if acetate is also formed and in a ratio of $\geq 1:1$ with butyrate [see Supporting Information Fig. S2, Buckel and Thauer (2013), and Louis and Flint (2017)], but ratios $<1:1$ are observed for these two bacteria (Supporting Information Table S2).

**Prediction of fermentation products.** To determine if our pathways could account for fermentation products formed by all bacteria, we formally compared predicted with observed fermentation products for all 48 bacteria (Fig. 2 and Supporting Information Table S2). A fermentation product was predicted if a bacterium had at least one set of subpathways that would lead to its formation (see Fig. 1). A fermentation product was still predicted if genes for hydrogenase, enolase, or both were missing (see Experimental Procedures).
Some cases existed where products were observed to be formed but not predicted. This applied to butyrate and propionate for *Po. levii* LEV, lactate for *Ruminobacter amylophilus* H18, succinate for *Ruminococcus flavefaciens* C94, and ethanol for *Bu. hungatei* JK615. In other cases where products were observed but not predicted, the amounts observed were small (≤0.05 mol/mol hexose equivalent).

More cases existed where the products were predicted but not observed. This applied especially to formate, acetate, lactate, and ethanol. Some of these predictions may be false positives. Alternatively, products may be formed, but only under different conditions. For example, *Me. elsdenii* LC-1 forms propionate during fermentation of lactate, but not glucose (Marounek et al., 1989). Additionally, some products may be formed, but only as intermediates and not end products. For example, formate may be an intermediate in fermentation (see Supporting Information Fig. S1AS and BP) or anabolic pathways (Thauer et al., 1972; Amador-Noguez et al., 2010; Zhou et al., 2015). Finally, some products may in fact be formed, but studies may not have reported them. For example, many studies do not report whether formate was formed or not (see Supporting Information Table S2), and it is unclear if this product was not formed or simply not reported.

Hydrogen was observed but not predicted for several bacteria. This was an expected result because genes for known hydrogenases were annotated with the wrong KO ID (Supporting Information Text S1 and Supporting Information Table S6), and our search would thus miss these and similar hydrogenases.
As mentioned, *F. succinogenes* subspecies would be able to form some succinate, though not in high proportions observed. Further, *E. ruminantium* GA195 and *Me. elsdenii* LC-1 would be able to form butyrate, but only if acetate is also formed and in a ratio of ≥1:1 with butyrate.

**Discussion**

Traditional methods of delineating fermentation pathways (e.g., enzymatic assays or isotopic labeling studies) are laborious, limiting the number of bacteria that can be screened. The largest effort for delineating these pathways in rumen bacteria, for example, included only nine species (Joyner and Baldwin, 1966). Further, only a few key enzymes in each pathway were assayed.

Reconstructing fermentation pathways from genomes is an approach with higher throughput. Our analysis encompasses nearly 50 hexose-utilizing species and subspecies from a single environment (the rumen). It is complementary to work for bacteria from the human colon, which delineated pathways for a smaller number of bacteria and for propionate and butyrate formation specifically (Louis and Flint, 2009; Reichardt et al., 2014; Louis and Flint, 2017). The present analysis reveals a high incidence of atypical pathways for hexose fermentation, and it shows areas of focus for future biochemical characterization.

**Conversion of glucose to pyruvate.** We found several bacteria with atypical EMP pathways. We found several genomes that encoded both PP$_i$- and ATP-dependent phosphofructokinases. Activity of the PP$_i$-dependent or both enzymes has been demonstrated in strains of four rumen bacteria (*Anaeroplasma intermedium* 5LA, *F. succinogenes* S85, *Pr. brevis* GA33, *Pr. ruminicola* 23) (Roberton and Glucina, 1982; Petzel et al., 1989). However, it remained unclear how PP$_i$ was generated as a substrate for the PP$_i$-dependent phosphofructokinase (Roberton and Glucina, 1982). This study’s discovery of a H$^+$-pyrophosphatase in some genomes could partly...
explain how the PP_i-dependent phosphofructokinase can function. Synthesis of glycogen or other macromolecules could supply some PP_i, but calculations show it is not enough for the EMP pathway (Heinonen, 2001; Zhou et al., 2013).

We found bacteria with incomplete EMP pathways. For example, we found enolase missing in genomes of five bacteria. Kelly et al. (2010) found enolase was missing in one of these genomes (Bu. proteoclasticus B316). These authors subsequently assayed and did not detect activity of the enzyme. Our group has documented certain other butyrivibrios (Butyrivibrio and Pseudobutyrovibrio) that lack the enolase gene (Hackmann and Firkins, 2015). Certain uncultured rumen bacteria (Candidatus Alcium) (Solden et al., 2017) were found to lack enolase genes when their genomes were assembled from metagenomic reads. Curiously, one bacterium (Bu. hungatei JK615) we found to be missing the enolase gene had detectable activity in Kelly et al. (2010). Enolase may be absent from the genome of this bacterium may be due to the incompleteness of the draft sequence, though the estimated completeness of the genome was high (100%; Supporting Information Table S1).

As pointed out by Kelly et al. (2010), the methylglyoxal pathway bypasses enolase and thus could serve as an alternative. However, this pathway leads to net loss of two ATP during the EMP pathway. Further, we did not find a full methylglyoxal pathway for any bacterium lacking enolase. We found nearly complete pathways, but genes for the terminal step (forming lactate) were missing when searching for the KO IDs (see Experimental Procedures). In sum, our study points out that more bacteria than the butyrivibrios may be missing enolase, and the methylglyoxal pathway may not be an alternative. Because no other pathway was present that would substitute, their pathway for forming pyruvate is unclear.
We found other genomes had an incomplete EMP pathway because they lack genes for pyruvate kinase. Activity for pyruvate kinase had been found missing in rumen bacterium *Asteroleplasma anaerobium* 161 (Petzel et al., 1989) and the non-rumen bacterium *C. thermocellum* LQ8 (Olson et al., 2017). Activity of malate shunt (*C. thermocellum* LQ8) or pyruvate phosphate dikinase (both bacteria) serves as a substitute (Petzel et al., 1989; Olson et al., 2017). Pyruvate kinase is missing also in genomes of *Candidatus* Alcium mentioned above (Solden et al., 2017). In our study, all three bacteria missing pyruvate kinase (*F. succinogenes* subspecies, *T. saccharophilum* PB) encode some variant of a malate shunt. *T. saccharophilum* PB alone encodes a pyruvate phosphate dikinase.

**Acetate formation.** The pathway for acetate formation encoded by *Selenomonas* and *Mi. jalaludinii* M 9 has not been documented in bacteria. In this pathway, acetate is formed from acetyl-CoA through the action of 1) succinyl-CoA:acetate CoA-transferase and 2) succinate—CoA ligase (ADP-forming). This pathway had been documented for eukaryotes alone (Tielens et al., 2010; Müller et al., 2012; Caspi et al., 2016).

For *Se. ruminantium*, an atypical pathway has been long-suspected because no phosphate acetyltransferase or acetate kinase activities have been detected experimentally (Joyner and Baldwin, 1966; Melville et al., 1988). With strain HD4, some workers (Melville et al., 1988) suggested that an acetate—CoA ligase (ADP-forming) could substitute for missing activities of the typical pathway. In support, these workers purified an enzyme that formed acetate from acetyl-CoA (Michel and Macy, 1990). However, we found no acetate—CoA ligase (ADP-forming) encoded by the genome. Further, the authors found the purified enzyme also formed succinate from succinyl-CoA (as well as propionate from propionyl-CoA). In retrospect, these authors likely did not purify an acetate—CoA ligase (ADP-forming), but rather a complex
containing succinate—CoA ligase (ADP-forming) and a succinyl-CoA:acetate CoA-transferase (see Supporting Information Table S7). Thus, available biochemical evidence suggests that *Se. ruminantium* forms acetate via our proposed pathway (novel for bacteria).

**Succinate and propionate formation.** The pathway for propionate (or succinate) formation encoded by *Prevotella* and *Selenomonas* has been not documented for bacteria or any other organism. In this pathway, Rnf generates NAD$_{\text{red}}$ from Fd$_{\text{red}}$. This activity is needed to balance fermentation, presuming NAD$_{\text{red}}$ is the only redox cofactor used downstream by NADH dehydrogenase. In support of this idea, authors have shown that NAD$_{\text{red}}$ is used to reduce fumarate in *Prevotella*, *Selenomonas*, and other rumen bacteria (White et al., 1962; Melville et al., 1988; Meinhardt and Glass, 1994; Asanuma and Hino, 2000). Authors have not documented whether or not Fd$_{\text{red}}$ is used, and thus it cannot be ruled out that Fd$_{\text{red}}$ could reduce fumarate directly and bypass Rnf. Because Rnf generates an electrochemical potential, bypassing Rnf would come at an energetic disadvantage to the cell, however.

The pathway for succinate formation in *F. succinogenes* subspecies is unclear because its genome does not encode either Rnf or components of the typical pathways. One study indicates show that *F. succinogenes* subsp. *succinogenes* S85 1) has an atypical pyruvate oxidoreductase that generates reduced FMN, not Fd$_{\text{red}}$ and 2) FMN, not NAD$_{\text{red}}$, is used to reduce fumarate (Miller, 1978). Subsequent work, however, has shown that NAD$_{\text{red}}$ can be used to reduce fumarate (Meinhardt and Glass, 1994; Asanuma and Hino, 2000). Thus, the redox cofactor used and its generation are unclear.

**Butyrate formation.** A previous study from our group found an atypical pathway for butyrate formation in most butyrivibrios (*Butyrivibrio* and *Pseudobutyrovibrio* spp.) (Hackmann and
Firkins, 2015). Specifically, the pathway involves the ion pump Ech hydrogenase, which together with Rnf generates an electrochemical potential and conserves energy. This study confirms that pathway is encoded by the butyrivibrios and few other organisms.

**Summary.** We document genomic evidence that rumen bacteria have several atypical pathways for hexose fermentation. We found a pathway of acetate formation documented previously only in eukaryotes, and pathway for succinate (and propionate) formation not documented in any organism. The atypical pathways for the EMP pathway that we found are not unprecedented, but their high incidence in bacteria of the rumen emphasizes they may not be as atypical as thought.

The pathways we reconstructed were able to predict formation of most products by most bacteria, but some bacteria still had products that could not be accounted for by any pathway (typical or atypical). This may indicate more atypical pathways, if not an artifact of incomplete draft genomes or misannotation.

Our finding that 44% of bacteria had at least one atypical pathway emphasizes the need for similar analyses for other bacterial communities. For bacteria of the human gut, pathways of propionate and butyrate formation have been analyzed (Louis and Flint, 2009; Reichardt et al., 2014; Louis and Flint, 2017), and these and other pathways may warrant reanalysis to determine the incidence of novel pathways proposed here. The high incidence of atypical pathways also cautions omics studies that reconstruct fermentation and other metabolic pathways, as these reconstructions will be complete only if all possible pathways are known. Ultimately, this study 1) emphasizes need for further biochemical characterization of fermentation pathways and 2) highlights the pathway steps that require most attention.

**Experimental Procedures**
Selection of strains. To select rumen bacteria for our analysis, we first made a list of species and subspecies of bacteria from reviews (Stewart et al., 1997; Creevey et al., 2014), texts (Russell, 2002; Dehority, 2003), and individual papers. We identified strains corresponding to species and subspecies on that list. We selected a strain for further analysis if it met the following criteria:

1) The species name is validly published [appears in the List of Prokaryotic names with Standing in Nomenclature (Parte, 2014)]

2) A culture is (or was) available in a public collection (e.g., DSMZ or ATCC)

3) A description is available in at least one peer-reviewed journal article

4) Carbohydrates are fermented

5) Fermentation products are reported for at least one hexose

6) A genome sequence is available in a public database

Strains from 87 species and subspecies met criteria 1 to 3 and are in Supporting Information Table S1. Of those, 48 also met criteria 4 to 6 and were included in our final set for analysis. All genomes were available on the IMG database (Chen et al., 2017). Completeness of genomes was estimated with CheckM (Parks et al., 2015).

If more than one strain per species (or subspecies) met all criteria, we chose the strain best studied. For *Se. ruminantium* subsp. *lactilytica*, a non-type strain (HD4) was chosen over the type strain (PC18) because fermentation pathways have been delineated biochemically for the non-type strain [see Melville et al. (1988) and Michel and Macy (1990)]. Results of metabolic reconstruction were identical for these two strains (not shown). In all other cases, the
type strain was chosen when available. In total, 44 out of 48 bacteria were represented by the type strain.

Reconstruction of fermentation pathways. Fermentation pathways were reconstructed by searching for genes of enzymes catalyzing relevant biochemical reactions (Supporting Information Table S3). These searches were conducted following methods from our previous studies (Hackmann and Firkins, 2015; Tao et al., 2016) and as detailed below. Locus tags of all genes found in our searches are reported in Supporting Information Table S8.

Some enzymes comprised of multiple genes, usually corresponding to different subunits. If all genes for an enzyme were present in a bacterium, the enzyme and its reaction were considered present (functional). If one or more genes of the enzyme were missing, the enzyme and the reaction were considered absent.

Some reactions could be catalyzed by multiple isozymes. Only one isozyme had to be present for the reaction to be considered present.

Sets of reactions were organized into subpathways (Supporting Information Fig. S1) and pathways in turn (Fig. 1). A subpathway was considered present in a bacterium only if all reactions were functional. A pathway for a fermentation product was considered present (the product was predicted) if at least one set of subpathways would lead to formation of that product (Fig. 1).

If genes for hydrogenase or enolase (or both) were missing, the pathway for a fermentation product was still considered present. These exceptions were made because 1) our searches did not produce hits for genes for all known hydrogenases (see Supporting Information...
Text S1 and Supporting Information Table S6) and 2) no fermentation products at all would be predicted for bacteria missing enolase.

All searches for genes were performed with the IMG/M data analysis system (Chen et al., 2017). The KO ID (Kanehisa et al., 2017) for the gene was searched first. The COG ID (Galperin et al., 2015) or pfam ID (Finn et al., 2016) was searched second and only if the KO ID was unreliable, such as when the KO ID did not produce a hit against a well-characterized gene (known to produce a functional enzyme) (see Supporting Information Table S3). The COG and pfam ID were searched second because they are not as specific as KO IDs. For example, no COG ID is specific for butyryl-CoA dehydrogenases, and the ID we searched (COG1960) applies to a broad family of acyl-CoA dehydrogenases. Consequently, the locus tags we report for butyryl-CoA dehydrogenase (see Supporting Information Table S8) likely include false positives. In total, 232 KO IDs, 7 COG IDs, and 3 pfam IDs were searched, limiting the number of false positives.

For genes in the methylglyoxal pathway, we searched for the KO IDs throughout, though in a previous study (Hackmann and Firkins, 2015) we had used the less-restrictive COG ID for one gene. Upon reexamination, we found the KO ID correctly produced hits for well-characterized genes (see Supporting Information Table S3).

Genes (or more precisely, ortholog groups) were defined with symbols. These symbols were taken mostly from KEGG (Kanehisa et al., 2017).

Enzymes were defined with symbols, enzyme commission (EC) numbers, and names. The enzyme symbol was derived from the gene symbol. The EC number was generally that defined by KEGG (Kanehisa et al., 2017) or MetaCyc (Caspi et al., 2016) for each KO, COG, or
pfam ID. The enzyme name was generally that defined by the IUBMB (Moss, 2017) for each EC. Each group of isozymes was defined with only one EC number and enzyme name.

Reactions were defined (drawn) in Supporting Information Fig. S1 according to the reaction definition given by KEGG (Kanehisa et al., 2017) for each KO ID, which closely corresponds to the reaction written by the IUBMB (Moss, 2017) for each EC. When the COG or pfam ID was searched, the reaction for the most-closely corresponding KO ID was used. If a reaction was not completely defined in KEGG, we defined the reaction as described in Supporting Information Text S1.

Some subpathways had variants (different sets of reactions yielding similar products). These variants were defined by appending the subpathway name with a Roman numeral and parenthetical description, giving, for example, “Acetyl-CoA → Acetate I (Acetate kinase)”. The name of these variants was our own or adopted from MetaCyc (Caspi et al., 2016), KEGG (Kanehisa et al., 2017), and standard texts (Gottschalk, 1986; White et al., 2012). Only one variant had to be present for the subpathway to be considered present.

We followed White et al. (2012) and Caspi et al. (2016) to define the pathway for glucose fermentation in Propionibacterium freudenreichii subsp. shermanii (see Figs. 7 and 8). Those sources did not define the NADH dehydrogenase and fumarate reductase/succinate dehydrogenase used, and we defined them by searching the genome of the type strain (CIRM-BIA1 / DSM 4902). A different strain (52W / NCIMB 9885) has been used for experimentally delineating most steps of the pathway (Caspi et al., 2016), but a genome sequence for that strain is not available.
Following the convention of Alberty (2003), we did not write the charges of substrates or products of reactions (e.g., $\text{NAD}^+$ is written as $\text{NAD}_{\text{ox}}$). An exception was made for $\text{Na}^+$, $\text{H}^+$, and Supporting Information Fig. S1AO to AS; this was needed in order to correctly represent the change in the number of positive charges inside the cell.

**Fermentation products.** Data for fermentation products were taken from experimental studies using batch cultures (Supporting Information Table S2). Preference was given towards the study describing initial isolation and characterization of the bacteria. Glucose was the most common substrate, and for uniformity, we report products for glucose fermentation when available ($n = 37$). When data for glucose were not available, we used data for cellulose ($n = 4$), cellobiose ($n = 3$), maltose ($n = 1$), galactose ($n = 1$), or mixtures of hexoses and other carbon sources ($n = 2$).

We report formation of products in terms of molar ratios (mol/mol hexose equivalent fermented). Hexose equivalent was defined as 6 mol of carbon in products. If $\text{CO}_2$ was not reported, its production was calculated using stoichiometry. This calculation assumed production of -1 mole $\text{CO}_2$ per mole formate and succinate; 0 mole $\text{CO}_2$ per mole propionate, caproate, and lactate; 1 mole $\text{CO}_2$ per mole acetate, ethanol, valerate, and isovalerate; and 2 moles $\text{CO}_2$ per mole butyrate, isobutyrate, and butanol.

Some studies reported fermentation products qualitatively (not quantitatively). Molar ratios could not be calculated in those cases, and we report formation of products as positive (+), negative (0), or trace (tr) instead.

**Acknowledgements**

We thank G. Suen (University of Wisconsin-Madison), A. Neumann (University of Wisconsin-Madison), and P. Weimer (USDA-ARS, Madison, WI) for conservations on missing pyruvate...
kinase in *Fibrobacter*. We also thank S. Hackmann (University of Florida) for reviewing the manuscript. This work was supported by U.S. Department of Agriculture (USDA) National Institute of Food and Agriculture (NIFA) Hatch Project FLA-ANS-005307 and USDA-NIFA Hatch/Multi-State Project FLA-ANS-005304.

Conflict of Interest

The authors declare no conflict of interest.

References


Figure legends

**Fig. 1.** Overview of fermentation pathways reconstructed from genomes of rumen bacteria. Each step refers to a subpathway shown in detail in Supporting Information Fig. S1 (indicated by letters next to pathway arrows). (A) Core pathways. (B) Ancillary pathways. Abbreviations: -3P = 3-phosphate, -6P = 6-phosphate, CoA = coenzyme A, EMP = Embden–Meyerhof–Parnas, P_i = inorganic phosphate, PEP = phosphoenolpyruvate, and PP_i = pyrophosphate.

**Fig. 2.** Observed and predicted fermentation products for rumen bacteria. Product is considered observed if 1) it was formed in molar ratio of ≥0.05 mol/mol hexose equivalent fermented or 2) its formation was reported qualitatively as positive (+). Product is predicted if at least one set of subpathways is encoded that would lead to formation of that product (Fig. 1). See Supporting Information Table S2 for additional fermentation products, molar ratios, and references to studies reporting data.

**Fig. 3.** Incidence of subpathways and atypical pathways in rumen bacteria. Grey squares show which subpathways or variants (n = 52) are encoded by each genome (n = 48). Colored heat map to right of grey squares shows overall incidence (%) of subpathways/variants across genomes. Barplot below grey squares shows how many atypical pathways (out of six) are encoded by each genome. As shown with dendogram at top, Euclidean distance matrix of pathway representation was used to group genomes by average-linkage clustering. For full names of...
subpathways/variants, see Fig. S1. For subpathways/variants not encoded by any genome (n = 16), see Supporting Information Table S4.

**Fig. 4.** The Embden–Meyerhof–Parnas (EMP) pathway. (A) Typical pathway. (B) Atypical pathway in rumen bacteria with pyrophosphate (PP$_i$)-dependent phosphofructokinase (*Anaeroplasma bactoclasticum* JR and most *Prevotella*) and H$^+$-pyrophosphatase. Panel A includes reactions shown in Supporting Information Fig. S1A, E, J, and L. Panel B includes reactions shown in Supporting Information Fig. S1A, F, J, L, BD, and BE. Reactions: 1. ATP-dependent phosphofructokinase (6-phosphofructokinase); 2. PP$_i$-dependent phosphofructokinase (diphosphate—fructose-6-phosphate 1-phosphotransferase), 3. ATP synthase, 4. H$^+$-pyrophosphatase. Abbreviations: -3P = 3-phosphate, -6P = 6-phosphate, 1,6P$_2$ = 1,6-bisphosphate, NAD$_{ox}$ = oxidized NAD, NAD$_{red}$ = reduced NAD, P$_i$ = inorganic phosphate, PP$_i$ = pyrophosphate.

**Fig. 5.** Incomplete pathways in rumen bacteria. (A) The Embden–Meyerhof–Parnas (EMP) pathway missing enolase (*Acetitomaculum ruminis* 139B, *Butyrivibrio hungatei* JK615, *Bu. proteoclasticus* B316, *Oscillibacter ruminantium* GH1, *Prevotella brevis* GA33). (B) EMP pathway missing pyruvate kinase (*Fibrobacter succinogenes* subspecies, *Treponema saccharophilum* PB). Panels A and B include reactions shown in Supporting Information Fig. S1A, E, J, and L. Reactions: 1. enolase (phosphopyruvate hydratase); 2. pyruvate kinase. Abbreviations: NAD$_{ox}$ = oxidized NAD, NAD$_{red}$ = reduced NAD, -2P = 2-phosphate, and P$_i$ = inorganic phosphate.
Fig. 6. Subpathway for forming acetate from acetyl-CoA. (A) Typical pathway. (B) Atypical pathway found in some rumen bacteria (all Selenomonas and Mitsuokella jalaludinii M 9). Panel A includes reactions shown in Supporting Information Fig. S1X. Panel B includes reactions shown in Supporting Information Fig. S1Y. Reactions: 1. phosphate acetyltransferase; 2. acetate kinase; 3. succinyl-CoA:acetate CoA-transferase; 4. succinate—CoA ligase (ADP-forming). Abbreviations: CoA = coenzyme A, -P = phosphate, and P_i = inorganic phosphate.

Fig. 7. Pathway for fermentation of glucose to propionate (or succinate) and acetate. (A) Typical pathway, which does not involve Rnf and is best characterized for Propionibacterium freudenreichii subsp. shermanii (non-ruminal strains). (B) Atypical pathway, which involves Rnf and is found in some rumen bacteria (all Prevotella and Selenomonas). For some bacteria, the pathway is truncated at succinate, and propionate is not formed. Coefficients x and y are variable and depend on the NADH dehydrogenase present (see Fig. 8). It is assumed that another reaction (e.g., that catalyzed by an antiporter) balances Na^+ and H^+. The typical EMP pathway is shown, but P. freudenreichii subsp. shermanii and some rumen bacteria encode an atypical EMP pathway, which involves a PPI-dependent phosphofructokinase (see Fig. 4B). Panel A includes reactions shown in Supporting Information Fig. S1A, E, J, L, V, X, AH, AI, AN, AO, AT, and BD. Panel B includes reactions shown in Supporting Information Fig. S1A, E, J, L, U, X, Y, AH, AI, AJ, AO, AP, AT, BC, and BD. Reactions: 1. pyruvate dehydrogenase; 2. NADH dehydrogenase and fumarate reductase/succinate dehydrogenase; 3. ATP synthase; 4. pyruvate:ferredoxin oxidoreductase; 5. Rnf [ferredoxin-NAD^+ oxidoreductase (Na^+-
Abbreviations: CoA = coenzyme A, $\text{Fd}_{\text{ox}}$ = oxidized ferredoxin, $\text{Fd}_{\text{red}}$ = reduced ferredoxin, $\text{NAD}_{\text{ox}}$ = oxidized NAD, $\text{NAD}_{\text{red}}$ = reduced NAD, and $\text{P}_i$ = inorganic phosphate.

**Fig. 8.** Pathway for fermentation of glucose to propionate (or succinate) and acetate, showing reactions involved in electron transport phosphorylation. (A) *Propionibacterium freudenreichii* subsp. *shermanii* DSM 4902 (non-rumen strain). (B) *Prevotella*. (C) *Selenomonas*. It is assumed that another reaction (e.g., that catalyzed by an antiporter) balances $\text{Na}^+$ and $\text{H}^+$. Panel A includes reactions shown in Supporting Information Fig. S1AQ and BD. Panel B includes reactions shown in Supporting Information Fig. S1AO, BC, and BD. Panel C includes reactions shown in Supporting Information Fig. S1AP, BC, and BD. Reactions: 1. Nuo NADH dehydrogenase [NADH:ubiquinone reductase ($\text{H}^+$-translocating)]; 2. fumarate reductase/succinate dehydrogenase; 3. ATP synthase; 4. Rnf [ferredoxin-NAD$^+$ oxidoreductase ($\text{Na}^+$-transporting)]; 5. Nqr NADH dehydrogenase [NADH:ubiquinone reductase ($\text{Na}^+$-transporting)]; 6. Ndh NADH dehydrogenase. Abbreviations: $\text{Fd}_{\text{ox}}$ = oxidized ferredoxin, $\text{Fd}_{\text{red}}$ = reduced ferredoxin, Fum = fumarate, $\text{NAD}_{\text{ox}}$ = oxidized NAD, $\text{NAD}_{\text{red}}$ = reduced NAD, $\text{Q}_{\text{ox}}$ = oxidized quinone, $\text{Q}_{\text{red}}$ = reduced quinone, and Suc = succinate.
**Figure 1**

This article is protected by copyright. All rights reserved.
Figure 4

(A) 1 Glucose

1 ATP
1 ADP
1 Fructose-6P
1 ATP
1 ADP
1 Fructose 1,6P₂
4 ADP + 2 Pᵢ + 2 NADₓ
4 ATP + 2 NADᵣₓ → 2 H₂O
2 Pyruvate

(B) 1 Glucose

1 ATP
1 ADP
1 Fructose-6P
1 Pᵢ
1 Pᵢ
1 Fructose 1,6P₂
4 ADP + 2 Pᵢ + 2 NADₓ
4 ATP + 2 NADᵣₓ → 2 H₂O
2 Pyruvate

0.25 ATP + 0.25 H₂O
0.25 ADP + 0.25 Pᵢ
1 H⁺
0.25 H⁺
1 H⁺
0.25 Pᵢ
1 H⁺
1 H⁺
1 PPᵢ + 1 H₂O

This article is protected by copyright. All rights reserved.
(A) 1 Glucose

2 P_i + 2 NAD_ox

2 NAD_red

2 Glycerate-2P

? 2 H_2O

2 Phosphoenolpyruvate

2 ADP

2 ATP

2 Pyruvate

(B) 1 Glucose

2 P_i + 2 NAD_ox

2 NAD_red

2 Glycerate-2P

? 2 H_2O

2 Phosphoenolpyruvate

2 ADP

2 ATP

2 Pyruvate
This article is protected by copyright. All rights reserved.
This article is protected by copyright. All rights reserved.