

# SCIENTIFIC REPORTS

OPEN

## Transcriptomic profiles of *Clostridium ljungdahlii* during lithotrophic growth with syngas or H<sub>2</sub> and CO<sub>2</sub> compared to organotrophic growth with fructose

Muktak Aklujkar<sup>1,2</sup>, Ching Leang<sup>1,3</sup>, Pravin M. Shrestha<sup>1,4</sup>, Minita Shrestha<sup>1,5</sup> & Derek R. Lovley<sup>1</sup>

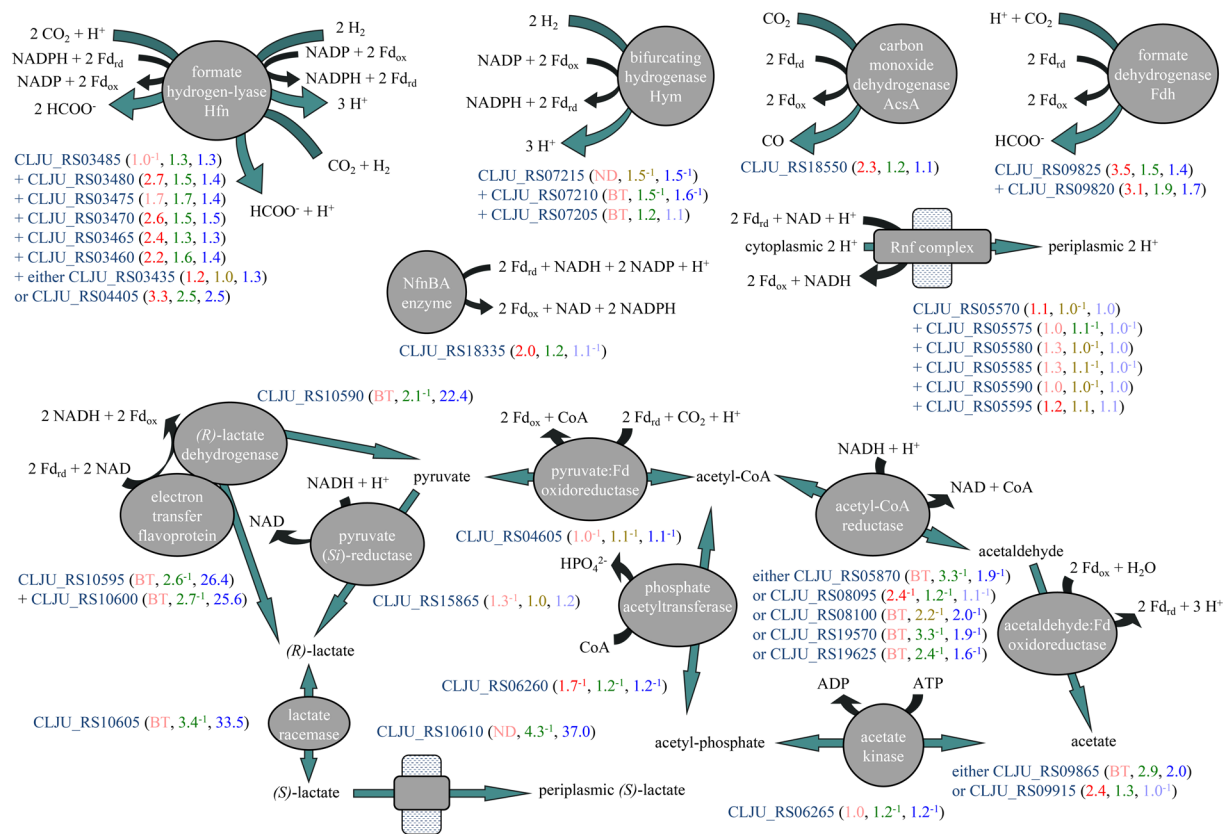
*Clostridium ljungdahlii* derives energy by lithotrophic and organotrophic acetogenesis. *C. ljungdahlii* was grown organotrophically with fructose and also lithotrophically, either with syngas - a gas mixture containing hydrogen (H<sub>2</sub>), carbon dioxide (CO<sub>2</sub>), and carbon monoxide (CO), or with H<sub>2</sub> and CO<sub>2</sub>. Gene expression was compared quantitatively by microarrays using RNA extracted from all three conditions. Gene expression with fructose and with H<sub>2</sub>/CO<sub>2</sub> was compared by RNA-Seq. Upregulated genes with both syngas and H<sub>2</sub>/CO<sub>2</sub> (compared to fructose) point to the urea cycle, uptake and degradation of peptides and amino acids, response to sulfur starvation, potentially NADPH-producing pathways involving (S)-malate and ornithine, quorum sensing, sporulation, and cell wall remodeling, suggesting a global and multicellular response to lithotrophic conditions. With syngas, the upregulated (R)-lactate dehydrogenase gene represents a route of electron transfer from ferredoxin to NAD. With H<sub>2</sub>/CO<sub>2</sub>, flavodoxin and histidine biosynthesis genes were upregulated. Downregulated genes corresponded to an intracytoplasmic microcompartment for disposal of methylglyoxal, a toxic byproduct of glycolysis, as 1-propanol. Several cytoplasmic and membrane-associated redox-active protein genes were differentially regulated. The transcriptomic profiles of *C. ljungdahlii* in lithotrophic and organotrophic growth modes indicate large-scale physiological and metabolic differences, observations that may guide biofuel and commodity chemical production with this species.

*C. ljungdahlii* is capable of lithotrophic growth by the Wood-Ljungdahl pathway with hydrogen as the electron donor, carbon dioxide as the electron acceptor, and acetate as the major end product, with a small amount of acetyl-CoA being assimilated into biomass<sup>1</sup>. *C. ljungdahlii* can also utilize syngas, a gas mixture containing H<sub>2</sub>, CO<sub>2</sub> and CO, as a lithotrophic substrate for the Wood-Ljungdahl pathway<sup>1</sup>. *C. ljungdahlii* also performs electrosynthesis, the reduction of CO<sub>2</sub> to acetate with electrons derived from electrodes<sup>2,3</sup>. The results of genetic modifications to redirect carbon and electron flow to products other than acetate suggest that *C. ljungdahlii* may serve as an effective chassis for converting CO<sub>2</sub> to fuels or other organic commodities<sup>4-8</sup>.

Various carbohydrates and amino acids also support the growth of *C. ljungdahlii*<sup>1</sup>. Under organotrophic conditions, *C. ljungdahlii* continues to operate the Wood-Ljungdahl pathway as an energetically favourable means to dispose of electrons from oxidation of the organic substrate. For example, glycolytic fermentation of fructose to lactate yields only two ATP, whereas oxidation of fructose to acetate plus CO<sub>2</sub> and use of the electrons to fix CO<sub>2</sub> into acetate by the Wood-Ljungdahl pathway yields four ATP<sup>9</sup> minus an estimated 0.14 ATP for reverse electron transport (using reactions described below and assuming that hydrolysis of 3 ATP pumps 11 protons)<sup>10</sup>.

<sup>1</sup>University of Massachusetts, Amherst, MA, 01003, USA. <sup>2</sup>University of Nevada, Reno, NV, 89557-0352, USA.

<sup>3</sup>LanzaTech, Skokie, 60077, IL, USA. <sup>4</sup>Energy Biosciences Institute, University of California, Berkeley, CA, 94720, USA. <sup>5</sup>Earth and Environmental Sciences, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA. Muktak Aklujkar, Ching Leang and Pravin M. Shrestha contributed equally to this work. Correspondence and requests for materials should be addressed to M.A. (email: [aklujkarmuktak@gmail.com](mailto:aklujkarmuktak@gmail.com)) or P.M.S. (email: [pravin@berkeley.edu](mailto:pravin@berkeley.edu))



**Figure 1.** Reactions that distribute electrons to NAD, NADP and ferredoxin (Fd<sub>ox</sub>) in *C. ljungdahliae* growing lithotrophically. Blue-green arrows stand for reactions and black arrows for oxidation/reduction of electron carriers, consumption/production of ATP, and similar effects. Fold changes in the transcript level of each gene are indicated according to RNA-Seq of H<sub>2</sub>/CO<sub>2</sub>-grown *versus* fructose-grown cells (red; statistically insignificant changes in coral; BT = below threshold for detection of differential expression, log<sub>2</sub> RPKM < 5; ND = not detected), microarray analysis of H<sub>2</sub>/CO<sub>2</sub>-grown *versus* fructose-grown cells (green; insignificant changes in olive), and microarray analysis of syngas-grown *versus* fructose-grown cells (blue; insignificant changes in light-blue). Fold changes with the exponent -1 signify downregulation. There are two putative H<sub>2</sub>-oxidizing enzymes: formate hydrogen-lyase Hfn and bifurcating hydrogenase Hym; and two formate-producing enzymes: Hfn and formate dehydrogenase Fdh; the AcsA subunit of acetyl-CoA synthase is carbon monoxide dehydrogenase. The proton-pumping Rnf complex and the NADH-dependent ferredoxin:NADP oxidoreductase Nfn conserve energy as they exchange reduced ferredoxin (Fd<sub>rd</sub>), NADH and NADPH. Inhibition of Rnf by a strong proton gradient would explain the dramatic upregulation of enzymes of (R)-lactate cycling that do not conserve energy from the exchange of two Fd<sub>rd</sub> for NADH during growth with syngas. During growth with H<sub>2</sub>/CO<sub>2</sub>, upregulation of aldehyde:ferredoxin oxidoreductases may maintain a reduced ferredoxin pool.

The discovery of bifurcated electron transfer<sup>11</sup> has led to the realization that metabolic pathways that were thought to be strictly fermentative (lacking electron transport, with substrate-level phosphorylation only) may conserve energy by coupling exergonic redox reactions to electron transfer from NADH to ferredoxin (Fd<sub>ox</sub>). This energy is converted into transmembrane potential with electron transport from ferredoxin (Fd<sub>rd</sub>) to NAD by the Rnf complex, which pumps protons or sodium ions.

The Wood-Ljungdahl pathway consumes one ATP to ligate formate with tetrahydrofolate (THF) and yields one ATP when acetyl-CoA is converted via acetyl-phosphate into acetate. Therefore, its net energy yield depends on the differences in redox potential between the substrates of organotrophy (e.g. fructose) or lithotrophy (e.g. H<sub>2</sub> or CO) and the electron carriers oxidized by four oxidoreductases of the Wood-Ljungdahl pathway. The first oxidoreductase is either formate dehydrogenase (Fig. 1), predicted to accept two electrons from the one-electron carrier Fd<sub>rd</sub> through a ferredoxin-like iron-sulfur cluster-binding subunit, or a formate hydrogen-lyase that accepts one electron from Fd<sub>rd</sub> and one from NADPH<sup>12,13</sup>. The second oxidoreductase is NADPH-dependent methenyl-THF reductase. The third is methylene-THF reductase, for which the electron carrier is unknown. It is hypothesized to perform bifurcated electron transfer by oxidizing two NADH and reducing two Fd<sub>ox</sub> to Fd<sub>rd</sub> as it reduces methylene-THF to methyl-THF<sup>9</sup>. The fourth oxidoreductase is the carbon monoxide dehydrogenase subunit of acetyl-CoA synthase (Fig. 1), which is predicted to require two electrons at the low redox potential of Fd<sub>rd</sub>. Thus, the exchange of electrons among Fd<sub>rd</sub>, NADH, and NADPH is fundamental to the operation of the Wood-Ljungdahl pathway regardless of the mode of growth. Two enzymes central to this exchange are the

proton-translocating Rnf complex<sup>14</sup>, which reduces NAD with two Fd<sub>rd</sub> (Fig. 1), and the Nfn enzyme<sup>15</sup> that transfers four electrons from one NADH and two Fd<sub>rd</sub> to two NADP (Fig. 1).

During organotrophic growth by oxidation of fructose to acetyl-CoA, the glycolytic enzymes glyceraldehyde-3-phosphate dehydrogenase and pyruvate:ferredoxin oxidoreductase donate equal numbers of electrons to NAD and Fd<sub>ox</sub>. This may allow the cell to derive all of its NADPH from the Nfn enzyme. In contrast, lithotrophic substrates such as H<sub>2</sub> and syngas will distribute electrons unevenly to NAD, NADP and Fd<sub>ox</sub>. Nfn-independent pathways may then be required to provide NADH or NADPH. Oxidation of carbon monoxide, the energetically preferable lower-redox-potential electron donor present in syngas, may be attributed to the carbon monoxide dehydrogenase subunit of acetyl-CoA synthase, which reduces only Fd<sub>ox</sub> (Fig. 1). Analysis of the *C. ljungdahlii* genome predicts that H<sub>2</sub> is oxidized by two enzyme complexes that bind NADP at soluble ligand-binding beta-grasp (SLBB) domains<sup>16</sup>: a bifurcating hydrogenase called Hym that couples electron transfer from H<sub>2</sub> to the higher redox potential of NADP and the lower redox potential of Fd<sub>ox</sub> (Fig. 1) and a formate hydrogen-lyase called Hfn - or Hyt in *Clostridium autoethanogenum*<sup>13</sup> - that oxidizes H<sub>2</sub> to reduce CO<sub>2</sub> to formate and is also capable of bifurcating electrons from H<sub>2</sub> to NADP and to Fd<sub>ox</sub> (Fig. 1). Two of the three Fd<sub>rd</sub> produced from 3 H<sub>2</sub> by Hym or Hfn would be consumed by carbon monoxide dehydrogenase (Fig. 1) and one NADPH would be consumed by methenyl-THF reductase; therefore, *C. ljungdahlii* growing under lithotrophic conditions with H<sub>2</sub> would perform electron transport through Rnf to convert the remaining 0.5 NADPH and one Fd<sub>rd</sub> into NADH or another electron carrier for methylene-THF reductase. It is not known how an excess of Fd<sub>rd</sub> or an excess of NADPH may impact the Wood-Ljungdahl pathway.

Previously published transcriptomic studies of *C. ljungdahlii* evaluated responses to growth with CO/CO<sub>2</sub> or fructose<sup>17</sup>, growth with H<sub>2</sub>/CO<sub>2</sub> or fructose and with NH<sub>4</sub><sup>+</sup> or NO<sub>3</sub><sup>-</sup> as the nitrogen source<sup>18</sup>, O<sub>2</sub> stress<sup>19</sup>, and NaCl stress<sup>20</sup> using RNA-Seq. In this study, the transcriptomic profiles of *C. ljungdahlii* grown lithotrophically with either syngas or H<sub>2</sub>/CO<sub>2</sub> and organotrophically with fructose were compared using microarrays, and gene expression with H<sub>2</sub>/CO<sub>2</sub> and with fructose was also compared using RNA-Seq. This approach allowed patterns of differential regulation to be identified as common to the two previously untested lithotrophic growth conditions compared to fructose, such as upregulation of urea cycle genes and downregulation of microcompartment genes, while other genes were upregulated only with syngas, such as (*R*)-lactate dehydrogenase, or only with H<sub>2</sub>/CO<sub>2</sub>, such as histidine biosynthesis genes. Many redox-active protein genes were also identified as differentially regulated.

## Results and Discussion

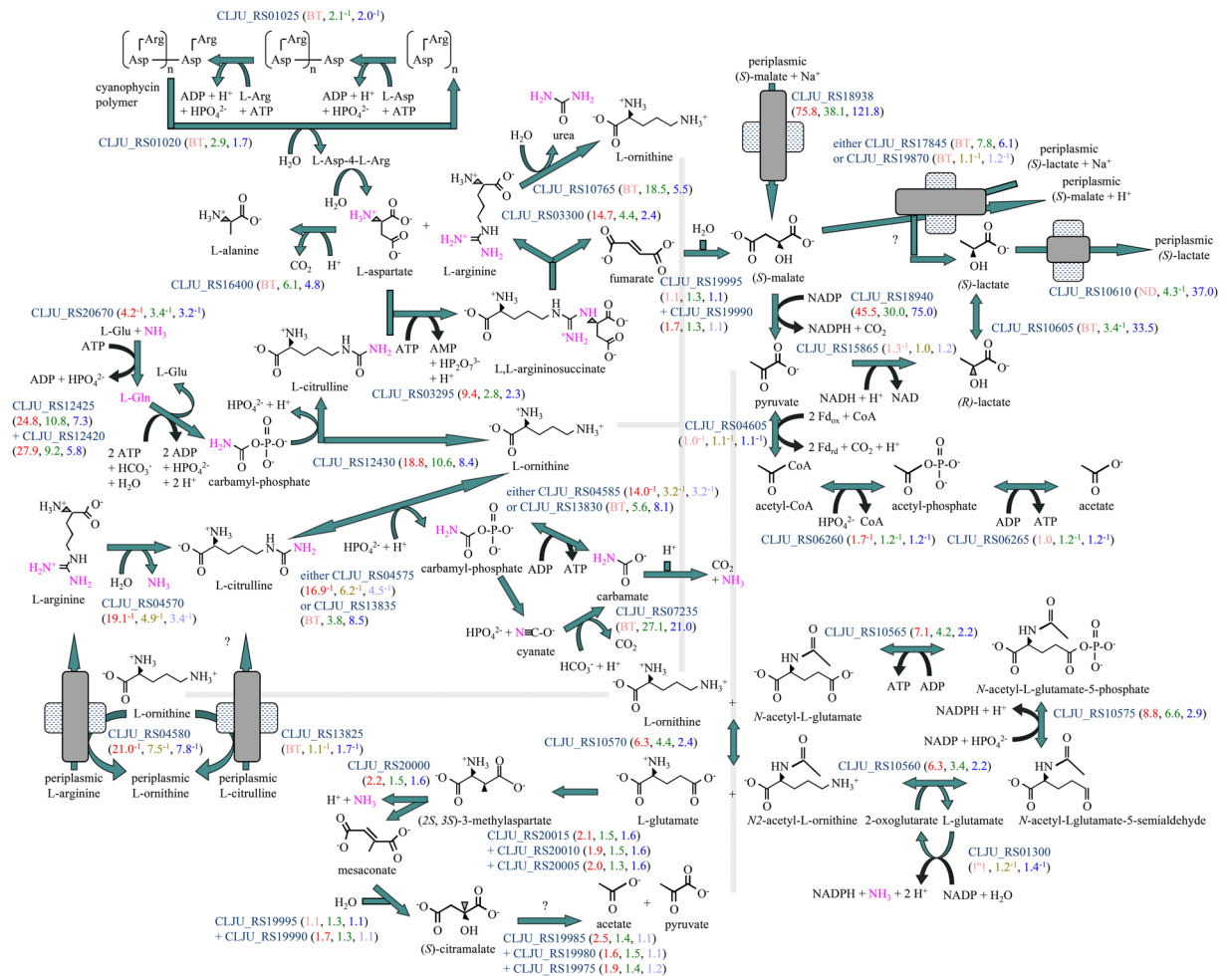
**Comparison of RNA-Seq and microarray datasets.** The total number of protein-coding genes identified in the genome of *C. ljungdahlii* is 4074, including genes that appear frameshifted in homopolymeric tracts, but excluding gene fragments. There are eight sets of two genes and five sets of five genes with ~100% DNA sequence identity, for which gene expression had to be measured collectively, and five sets of two genes with 97–99% DNA sequence identity, which shared some microarray probes. The microarray detected 4019 gene transcripts and RNA-Seq detected 3795 gene transcripts in total. Five gene transcripts without microarray probes were detected by RNA-Seq, but not differentially expressed. Of the 229 gene transcripts that were detected only by microarray, 84 genes showed significant differential expression ( $p \leq 0.05$ ). The number of genes differentially expressed ( $p \leq 0.05$ ) according to at least one method was 3341, of which 66 genes were disqualified because they were below the threshold of significance (median reads assigned per kilobase of target per million mapped reads (RPKM)  $\geq 5$ ) for RNA-Seq and did not meet the  $p$  value criterion for the microarray. Differentially expressed genes were ranked according to their largest significant fold change (Supplementary Table S1).

Many differentially expressed genes encode predicted transcriptional regulators or other genes that are expected to affect the expression of other proteins (Supplementary Table S2).

In the discussion below, genes are described as upregulated if expression was significantly ( $p \leq 0.05$ ) greater for lithotrophic growth with either syngas or H<sub>2</sub>/CO<sub>2</sub> (by microarray analysis and/or RNA-Seq) compared to fructose, and as downregulated if expression was significantly ( $p < 0.05$ ) lesser for a lithotrophic growth condition than for organotrophic growth with fructose. Genes for which the fold change was tenfold or larger are termed “strongly upregulated” or “strongly downregulated”.

**(S)-malate uptake and oxidation during lithotrophic growth.** The two most upregulated genes during lithotrophic growth with syngas (*maeN* CLJU\_RS18938, *ytsJ-2* CLJU\_RS18940) encode a protein with 55% sequence identity to the characterized sodium/(S)-malate symporter of *Bacillus subtilis*<sup>21</sup> and one of the four putative malic enzymes of *C. ljungdahlii* with 64% protein sequence identity to a major NADP-reducing enzyme that oxidatively decarboxylates (S)-malate to pyruvate in *B. subtilis*<sup>22,23</sup>. During lithotrophic growth with H<sub>2</sub>/CO<sub>2</sub>, these two genes were the fourth and sixth most upregulated genes according to the microarray and the first and fourth most upregulated genes according to RNA-Seq (Fig. 2). Genes for the (S)-malate sensor kinase and response regulator (*malK* CLJU\_RS18950, *malR* CLJU\_RS18945) were also upregulated with syngas and with H<sub>2</sub>/CO<sub>2</sub> (Supplementary Table S2). All of this indicates that lithotrophically grown *C. ljungdahlii* attempts to take up and oxidize (S)-malate even though the growth medium contains no organic carbon other than 0.1% yeast extract, vitamins, and nitrilotriacetate. Perhaps *C. ljungdahlii* is accustomed to encounter (S)-malate along with lithotrophic substrates in its natural environment; alternatively, it may be responding gratuitously to a molecule that mimics (S)-malate (one possibility being the tricarboxylate molecule nitrilotriacetate, which is present in the growth medium), or it may be responding to (S)-malate that it has previously excreted.

*C. ljungdahlii* possesses two homologs of the (S)-malic acid/sodium (S)-lactate antiporter of *B. subtilis*<sup>21</sup>. One homolog (CLJU\_RS19870; 32% protein sequence identity) was not differentially regulated with syngas or H<sub>2</sub>/CO<sub>2</sub> versus fructose, although it was the most upregulated gene under NaCl stress<sup>20</sup> and the tenth most upregulated gene with CO/CO<sub>2</sub> versus fructose<sup>17</sup>. The other homolog (CLJU\_RS17845; 22% protein sequence identity)



**Figure 2.** Amino acid degradation pathways that may be operated by lithotrophic *C. ljungdahlia*, including the urea cycle and uptake and oxidation of (S)-malate. Fold changes are indicated as in Fig. 1. Nitrogen being excreted is shown in pink. Arginine, derived either from the storage polymer cyanophycin or from yeast extract in the medium, is split into urea plus ornithine, which is degraded through glutamate, releasing ammonia. To prevent an increase in pH, ammonia is recaptured by glutamine synthetase and converted into carbamyl-phosphate, which reacts with ornithine and the amino group of aspartate in the urea cycle. The fumarate released from aspartate is degraded through (S)-malate, with (S)-lactate possibly excreted as an electron sink and then recaptured in exchange for (S)-malic acid, which is later recaptured; this transient removal of intermediates may ensure that amino acid degradation is sustained and that ammonia is sequestered as urea so that the pH does not rise. Concomitant upregulation of genes for citrulline fermentation, but not uptake of arginine/citrulline or fermentation of arginine through citrulline, suggests that citrulline is diverted from the urea cycle to release carbamyl-phosphate, from which ATP can be recovered and ammonia is released. Upregulation of cyanate lyase implies a high level of carbamyl-phosphate that spontaneously rearranges into cyanate and must be detoxified.

was upregulated with both syngas and  $H_2/CO_2$  (Fig. 2). If *C. ljungdahlia* uses this antiporter to import (S)-lactate or a similar organic acid and export (S)-malate, the sodium/(S)-malate symporter's function could be to recapture (S)-malate. The medium does not contain (S)-lactate, but it may be transiently produced by degradation of the amino acids in yeast extract or cyanophycin, a storage polymer of aspartate and arginine (Fig. 2). This idea is consistent with upregulation of genes for urea cycle enzymes (Fig. 2) by which arginine and aspartate are degraded to ornithine, urea and fumarate, which can be hydrated to (S)-malate. Oxidative decarboxylation of (S)-malate to pyruvate and of pyruvate to acetyl-CoA produces NADPH and two  $Fd_{rd}$  which can be reoxidized by the Wood-Ljungdahl pathway or exchanged for NADH by the Nfn and Rnf enzymes (Fig. 1), generating a proton gradient. Conversion of acetyl-CoA into acetate either generates ATP or exchanges NADH for two  $Fd_{rd}$  (Fig. 1). Therefore, either a high ATP:ADP ratio or a high  $Fd_{rd}:Fd_{ox}$  ratio could cause pyruvate to accumulate, in which case reduction of pyruvate with NADH to (R)-lactate, racemization to (S)-lactate, and excretion would allow sustained amino acid degradation. Subsequent recapture of the electron sink, (S)-lactate, by exchange with (S)-malic acid could serve three purposes: to recover electrons; to acidify the medium to counteract the release of ammonia from amino acids; and to promote sequestration of ammonia by the urea cycle by removing one



product, fumarate (Fig. 2). At a later stage, when amino acids are unavailable to supplement lithotrophic growth, (S)-malate could be recaptured and oxidized.

**Upregulation of cytoplasmic redox-active protein genes.** Three of the twenty-five most upregulated genes with H<sub>2</sub>/CO<sub>2</sub> are predicted to be co-transcribed and encode a transcriptional regulator, a flavodoxin, and a protein of unknown function (Supplementary Table S3). Not one of these genes was differentially regulated with syngas, suggesting that their products function only in lithotrophic growth with H<sub>2</sub>/CO<sub>2</sub>. (Previously, the flavodoxin was not detected in a syngas proteomic sample<sup>24</sup>). Several other genes predicted to encode cytoplasmic redox-active proteins were upregulated with H<sub>2</sub>/CO<sub>2</sub> and/or syngas (Supplementary Table S3). It might be valuable to study whether their functions, such as reduction of disulfide bonds and detoxification of reactive oxygen species, are more important for lithotrophic growth.

**Upregulation of membrane-associated redox protein genes.** *C. ljungdahlii* does not possess genes for biosynthesis of heme or the redox-active portion of any known quinone<sup>9</sup>, and therefore appears to lack electron transport pathways except for the Rnf complex. However, when *C. ljungdahlii* consumes an electrical current to perform electrosynthetic carbon fixation<sup>3</sup>, electrons must traverse the membrane to reach cytoplasmic electron carriers, perhaps through other membrane-associated redox proteins. One such protein complex may be the product of an operon that was upregulated with syngas and with H<sub>2</sub>/CO<sub>2</sub> (Supplementary Table S4), encoding a membrane protein of the DsbD superfamily that may form redox-active disulfide bonds (CLJU\_RS13020), an iron-sulfur cluster-binding membrane protein (CLJU\_RS13025), a part of which resembles the ubiquinol-oxidizing NapH subunit of the periplasmic nitrate reductase of *E. coli*<sup>25</sup>, and a periplasmic protein (CLJU\_RS13030). The substrate specificity of this putative oxidoreductase is an interesting topic for future research.

Several other genes for redox-active proteins with predicted membrane associations were upregulated with syngas and/or H<sub>2</sub>/CO<sub>2</sub> (Supplementary Table S4). Notable among them is the first of two tandem gene sets that each encode homologs of four proteins. The first protein is a (molybdopterin cytosine dinucleotide)-oxothiomolybdenum-binding aldehyde oxidoreductase (CLJU\_RS11830, CLJU\_RS11870). The second protein is a fused FAD-binding pyridine nucleotide-disulfide oxidoreductase and CCG domain pair iron-sulfur cluster-binding oxidoreductase (CLJU\_RS11825, CLJU\_RS11865), which may reduce or oxidize a membrane-bound electron carrier through the CCG domain pair by analogy with quinone/phenazine oxidoreductases<sup>26,27</sup>. The third protein has four conserved cysteine residues (CLJU\_RS11820, CLJU\_RS11860). The fourth protein is of the C\_GCAxxG\_C\_C family (CLJU\_RS11815, CLJU\_RS11850). Both proteins could be redox-active. The nature of the redox reactions carried out by these gene products should be investigated for their possible relevance to electron transport.

**Lactate oxidation as a ferredoxin reoxidation strategy during growth with syngas.** The fourth and fifth most upregulated genes during growth with syngas (*lctP* CLJU\_RS10610 and *larA-3* CLJU\_RS10605, Fig. 1) encode proteins with 49% sequence identity to the predicted (S)-lactate transporter of *Streptococcus iniae*<sup>28</sup> and 63% sequence identity to the characterized lactate racemase of *Lactobacillus plantarum*<sup>29</sup>, respectively. The ninth, tenth and thirteenth most upregulated genes (*lutJ* CLJU\_RS10595, *lutI* CLJU\_RS10600 and *lutK* CLJU\_RS10590, Fig. 1) encode an electron transfer flavoprotein and a candidate (R)-lactate dehydrogenase with 28% protein sequence identity to an (R)-lactate-oxidizing enzyme of *E. coli*<sup>30</sup>. These genes are predicted to be in an operon, possibly including the twelfth most upregulated gene (CLJU\_RS10585, 23.5-fold) that encodes a protein of the hemerythrin-like superfamily. These six genes were only upregulated with syngas (as with CO/CO<sub>2</sub><sup>17</sup>); with H<sub>2</sub>/CO<sub>2</sub>, they were either downregulated or not differentially regulated (Fig. 1). Their strong upregulation with syngas suggests that oxidation of (R)-lactate to pyruvate with NAD, coupled to electron transfer from two Fd<sub>rd</sub> to NAD<sup>31</sup>, may be a particularly important reaction for *C. ljungdahlii* when the electron donor is CO from syngas (Fig. 1).

Carbon monoxide, the energetically preferable electron donor in syngas, is also an intermediate of the Wood-Ljungdahl pathway, produced within acetyl-CoA synthase by the reduction of CO<sub>2</sub> with two Fd<sub>rd</sub>. Of the three genes of *C. ljungdahlii* predicted to encode catalytic subunits of carbon monoxide dehydrogenases, two (*cooS-1* CLJU\_RS04490, *cooS-2* CLJU\_RS08800) were downregulated with syngas (Supplementary Table S1; as with CO/CO<sub>2</sub><sup>17</sup>, although the *CooS-1* protein was abundant in a proteomic sample<sup>24</sup>) and therefore unlikely to contribute to oxidation of CO. The third gene, for which transcripts were most abundant (*acsA* CLJU\_RS18550), is located within the Wood-Ljungdahl pathway gene cluster and presumably functions both to provide CO to acetyl-CoA synthase and to oxidize CO with electron transfer to two Fd<sub>ox</sub> (Fig. 1). Thus, it is likely that when CO is the electron donor, all electrons pass to Fd<sub>ox</sub> and must be transferred to NAD and NADP for use in the methyl branch of the Wood-Ljungdahl pathway. The Rnf complex catalyzes electron transport from two Fd<sub>rd</sub> to NAD (Fig. 1), but only to the extent that it can pump protons against the gradient. When the proton gradient is too strong, another route may be required for electrons from two Fd<sub>rd</sub> to reach NAD. An electron transfer flavoprotein can provide this route by coupling the exergonic oxidation of two Fd<sub>rd</sub> by NAD with an endergonic reaction such as oxidation of (R)-lactate to pyruvate with reduction of NAD<sup>31</sup>.

*C. ljungdahlii* has a putative source for endogenous (R)-lactate: an NADH-dependent pyruvate (Si)-reductase *LdhA*, with 100% protein sequence identity to the characterized enzyme of *Clostridium autoethanogenum*<sup>32</sup>, encoded by a gene that is not differentially expressed (*ldhA* CLJU\_RS15865). With no obvious mechanism to conserve energy from this exergonic reaction, the *LdhA* enzyme may function irreversibly, driving a cycle of pyruvate reduction and (R)-lactate reoxidation (Fig. 1). With *LdhA* oxidizing one NADH and an electron transfer flavoprotein-associated (R)-lactate dehydrogenase oxidizing two Fd<sub>rd</sub> and reducing two NAD, there would be a net irreversible exchange of two Fd<sub>rd</sub> for one NADH independently of the Rnf complex. We speculate that if this

energetically wasteful cycle were abolished, *C. ljungdahlii* might conserve all of the available energy from syngas for bioenergy applications.

**Operation of the urea cycle to limit the release of ammonia.** The arginase gene (CLJU\_RS10765) was upregulated with syngas and with H<sub>2</sub>/CO<sub>2</sub>, suggesting that arginine released by breakdown of the storage polymer cyanophycin may be hydrolyzed to ornithine and urea (Fig. 2). Arginase is the key enzyme of the urea cycle, for which all of the other genes were also upregulated (Fig. 2), suggesting that additional urea is produced from ammonia released by degradation of ornithine and from aspartate released from cyanophycin. The upregulated genes encode one of two isozymes of carbamyl-phosphate synthase and one of three isozymes of ornithine carbamyltransferase, plus argininosuccinate synthetase and argininosuccinate lyase. The argininosuccinate enzymes are encoded by a predicted operon that includes genes for an amino acid uptake ATP-binding cassette (ABC) transporter that were also upregulated with syngas and with H<sub>2</sub>/CO<sub>2</sub> (Supplementary Table S5). The hypothesis that this transporter enables uptake of exogenous argininosuccinate or cyanophycin units is consistent with the 34–37% sequence identity of its periplasmic amino acid-binding protein to *Escherichia coli* and *B. subtilis* proteins for uptake of another side-chain-linked amino acid dimer, cystine<sup>33,34</sup>. The overall gene expression pattern suggests that the enzymes of the urea cycle are abundant in *C. ljungdahlii* growing lithotrophically.

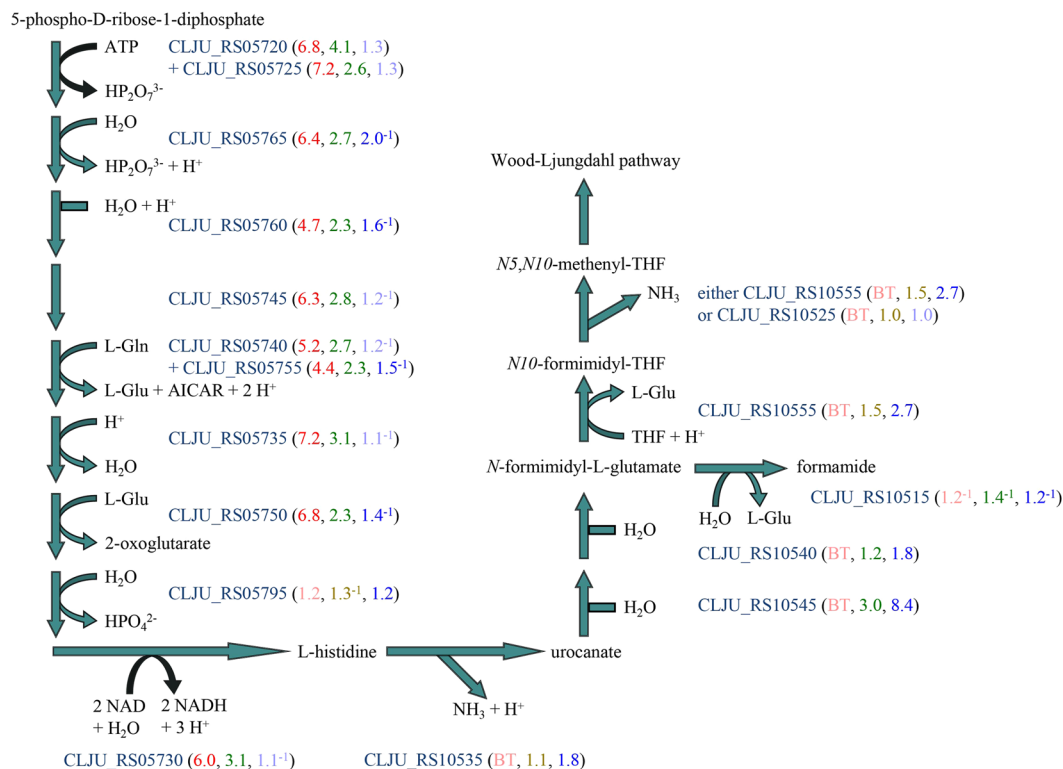
The urea cycle expends five ATP to sequester a molecule of ammonia and the amino group of aspartate (Fig. 2), but the estimated energy yield from degradation of ornithine and fumarate to acetate (Fig. 2) with disposal of electrons through the Wood-Ljungdahl pathway is only 3.38 ATP, assuming that 3 ATP are synthesized by passing 11 protons through ATP synthase as in *Clostridium paradoxum*<sup>10</sup>. Thus, it is energetically prohibitive for *C. ljungdahlii* to degrade cyanophycin solely through the urea cycle. Degradation of arginine and aspartate by non-urea-cycle reactions that release ammonia (Fig. 2) is estimated to yield 4.38 ATP. It is probable that *C. ljungdahlii* expends some of this energy to operate the urea cycle to mitigate the release of ammonia that raises the cytoplasmic pH. This energetic cost should be considered when choosing nitrogenous supplements for media for large-scale lithotrophic growth of *C. ljungdahlii*. Genes for salvaging and detoxification of urea cycle intermediates were also upregulated (Supplementary Discussion S1).

**NADP-reducing degradation of ornithine under lithotrophic conditions.** After arginase converts arginine from either cyanophycin or yeast extract into ornithine, *C. ljungdahlii* degrades ornithine to glutamate by a reversible pathway, for which all of the genes were upregulated with syngas and with H<sub>2</sub>/CO<sub>2</sub> (Fig. 2). The alternative interpretation, that these genes were upregulated for biosynthesis of arginine from glutamate, is implausible because the arginase gene was upregulated and the genes for glutamate biosynthesis were downregulated. Furthermore, gene expression patterns were consistent with degradation of glutamate, for which most of the genes were modestly upregulated with syngas and with H<sub>2</sub>/CO<sub>2</sub> (Fig. 2). Degradation of ornithine to glutamate releases ammonia and produces ATP and two NADPH; degradation of glutamate through pyruvate to acetate releases ammonia and produces ATP and two Fd<sub>rd</sub> (Fig. 2). Thus, ornithine degradation may augment the pool of NADPH in cells growing with either syngas or H<sub>2</sub>/CO<sub>2</sub>, whereas glutamate degradation, as a source of Fd<sub>rd</sub>, may be beneficial for growth with H<sub>2</sub>/CO<sub>2</sub>.

**Upregulation of histidine biosynthesis and histidine degradation.** The histidine biosynthesis operon was upregulated with H<sub>2</sub>/CO<sub>2</sub> (as with NaCl stress<sup>20</sup>), but not with syngas (Fig. 3). Only one enzyme of histidine biosynthesis, the putative histidinol-phosphate phosphatase (*hisK* CLJU\_RS05795), was not differentially regulated. This pattern suggests that *C. ljungdahlii* might need to synthesize proteins with high histidine content specifically to grow lithotrophically with H<sub>2</sub>/CO<sub>2</sub>. To explore this idea, the histidine content of every predicted protein of *C. ljungdahlii* was computed. Of those whose histidine content is at least 3.5%, which is two standard deviations (1.1%) above the mean (1.4%), most genes were not upregulated with H<sub>2</sub>/CO<sub>2</sub> (data not shown). A high-histidine-content protein gene that was upregulated with H<sub>2</sub>/CO<sub>2</sub> (Supplementary Table S6) encodes a transcriptional regulator (CLJU\_RS09505) with a histidine-rich C-terminal extension. Genes upregulated with H<sub>2</sub>/CO<sub>2</sub> and with syngas encode a protein of unknown function (CLJU\_RS11125) between two sets of peptide ABC transporter genes (Supplementary Table S5), for which a paralogous protein (CLJU\_RS18470) is encoded on the 3' side of the Wood-Ljungdahl pathway genes, a repeat-containing protein with two pentahistidine motifs (CLJU\_RS07790), and cyanate lyase (CLJU\_RS07235; Supplementary Discussion S1). These genes are candidates to be investigated for roles in lithotrophic growth.

Surprisingly, the histidine degradation gene cluster was upregulated with syngas; it encodes a transcriptional regulator (CLJU\_RS10550, Supplementary Table S2) and six enzymes that degrade histidine to glutamate (Fig. 3). This pattern suggests that *C. ljungdahlii* supplements its lithotrophic metabolism of syngas by degrading histidine, which it may derive from yeast extract. The one-carbon unit in the side-chain of histidine can be either transferred as a formimidyl group to THF to enter the Wood-Ljungdahl pathway (without the expense of one ATP to ligate formate to THF) or hydrolyzed to formamide and oxidized through formate. The gene for *N*-formimidyl-L-glutamate formimidylhydrolase (*hutE* CLJU\_RS10515) was the only downregulated gene of histidine degradation, indicating that when syngas is available as an electron donor, *C. ljungdahlii* prefers not to oxidize the one-carbon unit and instead uses it to make the methyl group of acetate.

**Upregulation of molybdopterin biosynthesis under lithotrophic conditions.** The genes for ferredoxin-dependent formate dehydrogenase, which were modestly upregulated (Fig. 1), are transcribed divergently from two strongly upregulated genes for incorporation of sulfur into molybdopterin, a precursor of the cofactor of formate dehydrogenase: the sulfur carrier protein gene (*moaD-1* CLJU\_RS09835) was upregulated with H<sub>2</sub>/CO<sub>2</sub>, while the sulfurylase gene (*moeB-1* CLJU\_RS09840) was upregulated with both H<sub>2</sub>/CO<sub>2</sub> and syngas (Supplementary Table S7). Another pair of sulfur carrier protein and sulfurylase genes nearby (*moaD-2*



**Figure 3.** Upregulation of histidine biosynthesis and degradation genes in *C. ljungdahlii* growing lithotrophically. Fold changes are indicated as in Fig. 1. Histidine is degraded to glutamate, which is further degraded as in Fig. 2.

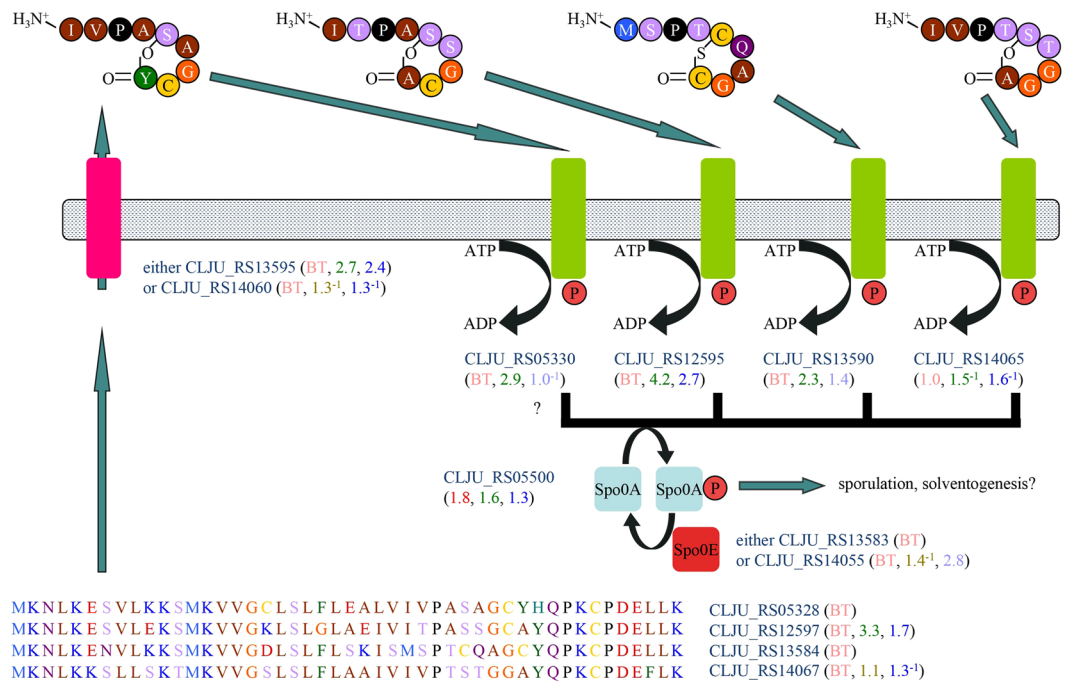
CLJU\_RS09920, *moeB*-2 (CLJU\_RS09925) were upregulated with H<sub>2</sub>/CO<sub>2</sub> only (Supplementary Table S7). Other genes of molybdopterin biosynthesis were less upregulated (Supplementary Table S7). This suggests that when *C. ljungdahlii* grows with H<sub>2</sub>/CO<sub>2</sub>, incorporation of sulfur into molybdopterin may be increased to meet demand for molybdopterin-containing cofactors. Genes for molybdopterin-containing enzymes were also upregulated (Supplementary Discussion S2 and below).

**Indications of sulfur limitation under lithotrophic conditions.** Several genes for uptake and metabolism of sulfur compounds were upregulated under lithotrophic conditions (Supplementary Table S8). One strongly upregulated operon (CLJU\_RS13650-CLJU\_RS13635) encodes components of a transporter with 32–51% identity to a putative *S*-methylcysteine transporter of *B. subtilis*<sup>33</sup> and a putative cystathionine beta-lyase with 37% protein sequence identity to PatB of *B. subtilis*<sup>35</sup>. On the 5' side of this operon is another upregulated operon (CLJU\_RS13670-CLJU\_RS13655); it encodes components of a transporter with 29–41% identity to the D-methionine and high-affinity L-methionine transporter of *E. coli*<sup>36</sup> and 30–37% identity to the D-methionine, L-methionine and methionine-sulfoxide transporter of *B. subtilis*<sup>37</sup>, as well as a racemase/dehydratase. Upregulation of these genes (Supplementary Table S8) suggests that under lithotrophic conditions, *C. ljungdahlii* attempts to degrade thioethers of cysteine and/or homocysteine (e.g. methionine), possibly to access sulfur.

Three genes with 39–42% protein sequence identity to subunits of anaerobic sulfite reductase in *Salmonella enterica* Typhimurium<sup>38</sup> (CLJU\_RS11690-CLJU\_RS11680) were upregulated with syngas but downregulated with H<sub>2</sub>/CO<sub>2</sub> (Supplementary Table S8). These genes are strongly upregulated with NO<sub>3</sub><sup>-</sup> as the nitrogen source, and have been suggested to encode nitrite reductase<sup>18</sup>. On the 3' side of these genes, four genes (CLJU\_RS11675-CLJU\_RS11660) encoding a candidate proton/nitrate symporter and components of a molybdopterin-dependent candidate assimilatory nitrate reductase<sup>18</sup> were upregulated with syngas and downregulated with H<sub>2</sub>/CO<sub>2</sub> (Supplementary Table S8). Thus, *C. ljungdahlii* responds differently to syngas and to H<sub>2</sub>/CO<sub>2</sub> in terms of expression of genes for assimilation of sulfite and/or nitrate. These differences may prove important for industrial-scale applications of *C. ljungdahlii* with different electron donors and sources of sulfur and nitrogen.

Five genes in an operon (CLJU\_RS12535-CLJU\_RS12555) encoding a molybdopterin-dependent oxidoreductase of the dimethylsulfoxide reductase family, including a flavoprotein subunit with 41% protein sequence identity to that of anaerobic sulfite reductase in *Salmonella enterica* Typhimurium<sup>38</sup>, were upregulated with syngas and with H<sub>2</sub>/CO<sub>2</sub> (Supplementary Table S8). The function of this putative complex, possibly in sulfoxide metabolism, is an interesting topic for future research.

Sulfur starvation can be simulated by the presence of chromate, which competes with sulfate uptake and assimilation (although *C. ljungdahlii* cannot assimilate sulfate<sup>24</sup>) and causes oxidative stress that depletes cysteine-based electron carriers<sup>39</sup>. Chromate efflux transporters consist of a pair (or fusion) of paralogous proteins with opposite



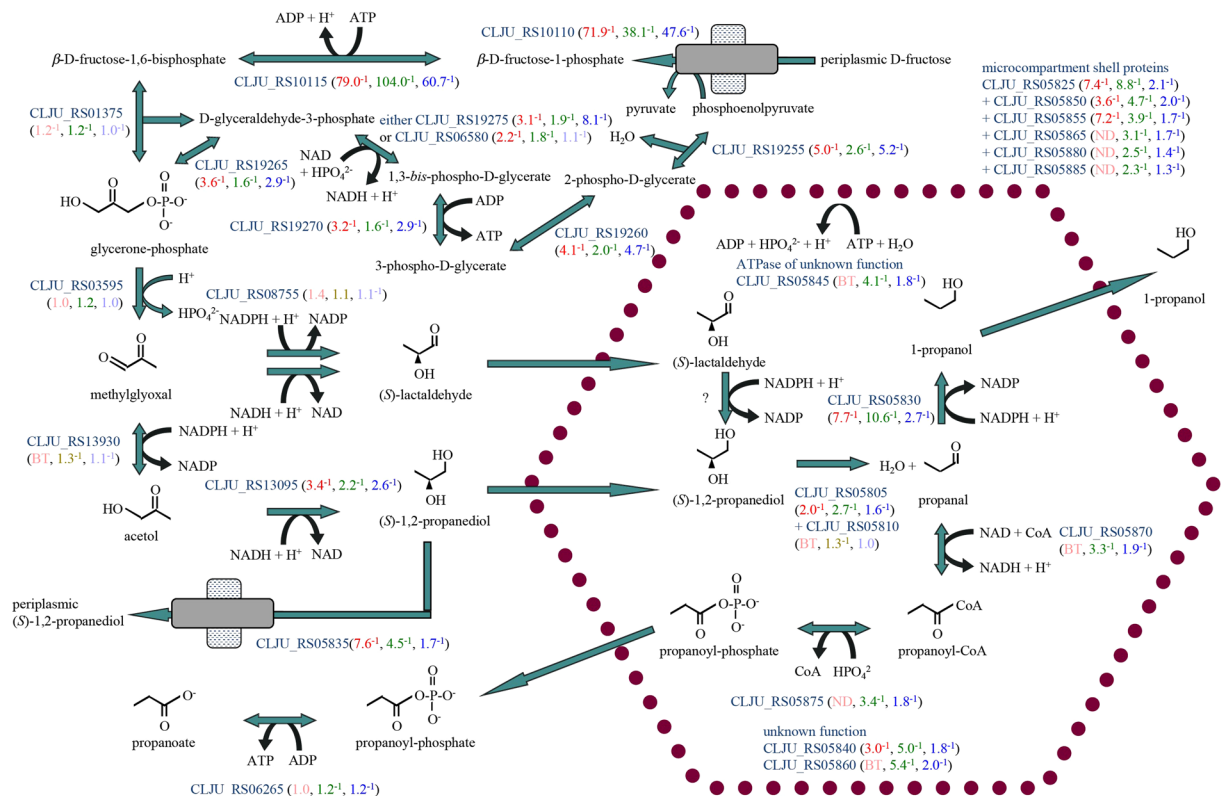
**Figure 4.** Upregulation of quorum sensing genes under lithotrophic conditions. Four autoinducer precursor proteins may be processed into peptide lactones by membrane proteins during secretion. (Lactone structures illustrated are predictions based on LamD of *L. plantarum*; actual structures are unknown). The secreted autoinducer lactones are predicted to be sensed by histidine kinases, which activate Spo0A. Spo0E dephosphorylates Spo0A. Accumulation of phosphorylated Spo0A commits the cell to behaviours such as sporulation, solventogenesis, or biofilm formation.

topology<sup>40,41</sup>. In *C. ljungdahlii*, these two genes were upregulated with syngas and with H<sub>2</sub>/CO<sub>2</sub> (Supplementary Table S8). Both have stronger BLAST hits to subunit 2, but topology prediction by the positive-inside rule<sup>42</sup> indicates that CLJU\_RS17795 encodes subunit 1 and CLJU\_RS17790 encodes subunit 2. This upregulation of chromate efflux, together with the preceding observations, indicates that *C. ljungdahlii* is sulfur-limited under lithotrophic conditions. This may be an important consideration for biofuel production, which might be maximized by control of the levels of enzymes in *C. ljungdahlii* by the availability of sulfur. Notably, sulfur uptake and metabolism proteins that are more abundant in solventogenesis *versus* acidogenesis from syngas<sup>24</sup> are different from these genes upregulated in lithotrophic *versus* organotrophic growth.

**Upregulation of quorum sensing and sporulation genes under lithotrophic conditions.** Biofilm formation by *C. ljungdahlii* is a factor that could improve the efficacy of commodity chemical production from syngas or by electrosynthesis<sup>20</sup>. The genome of *C. ljungdahlii* encodes four sensor histidine kinases (Fig. 4) that are homologs of AgrC of *Clostridium acetobutylicum* (20–22% protein sequence identity), the sensor of a peptide thiolactone autoinducer-dependent signaling system controlling sporulation and synthesis of the storage polymer granulose<sup>43</sup>, and LamC of *Lactobacillus plantarum* (18–21% protein sequence identity), the putative sensor for peptide thiolactone autoinducer-dependent control of biofilm formation<sup>44</sup>. These sensor kinases are 72–80% identical to each other in protein sequence and are each encoded next to genes for precursor proteins of peptide (thio)lactone autoinducers (two of which were discovered through this work). Membrane proteins for autoinducer processing are also encoded near two of the kinases. Several of these quorum sensing genes were upregulated with H<sub>2</sub>/CO<sub>2</sub> and with syngas (Fig. 4). Notably, no response regulators are encoded next to any of the four sensor histidine kinase genes, but genes next to two of them encode homologs of Spo0E, a phosphatase that deactivates the Spo0A protein that regulates both sporulation and solventogenesis in *Clostridium beijerinckii* and *Clostridium cellulolyticum*<sup>45</sup>. These orphan kinases are candidates to activate and deactivate Spo0A, as in *C. acetobutylicum*<sup>46</sup>. Numerous genes that were upregulated under lithotrophic conditions have been implicated in sporulation or cell wall biogenesis (Supplementary Table S9), indicating that the change in growth mode causes global changes in physiology. Investigation of the quorum sensing signaling systems for possible control of sporulation, biofilm formation, and solventogenesis by *C. ljungdahlii* may lead to improved designs for large-scale production of biofuels and commodity chemicals.

In *B. subtilis*, glutamine-hydrolyzing asparagine synthetase is expressed specifically during sporulation, for which it is required<sup>47</sup>. In *C. ljungdahlii*, the homologous gene (*asnO* CLJU\_RS13170) was upregulated with H<sub>2</sub>/CO<sub>2</sub> (Supplementary Table S9), but so was the gene for ammonia-dependent asparagine synthetase (CLJU\_RS03315). Asparaginyl-tRNA (made by transamidation, however) has been proposed to be a signal in the switch from acidogenesis to solventogenesis in *C. acetobutylicum*<sup>48</sup>. Future experiments may determine whether asparagine or asparaginyl-tRNA is also a signal in *C. ljungdahlii*.



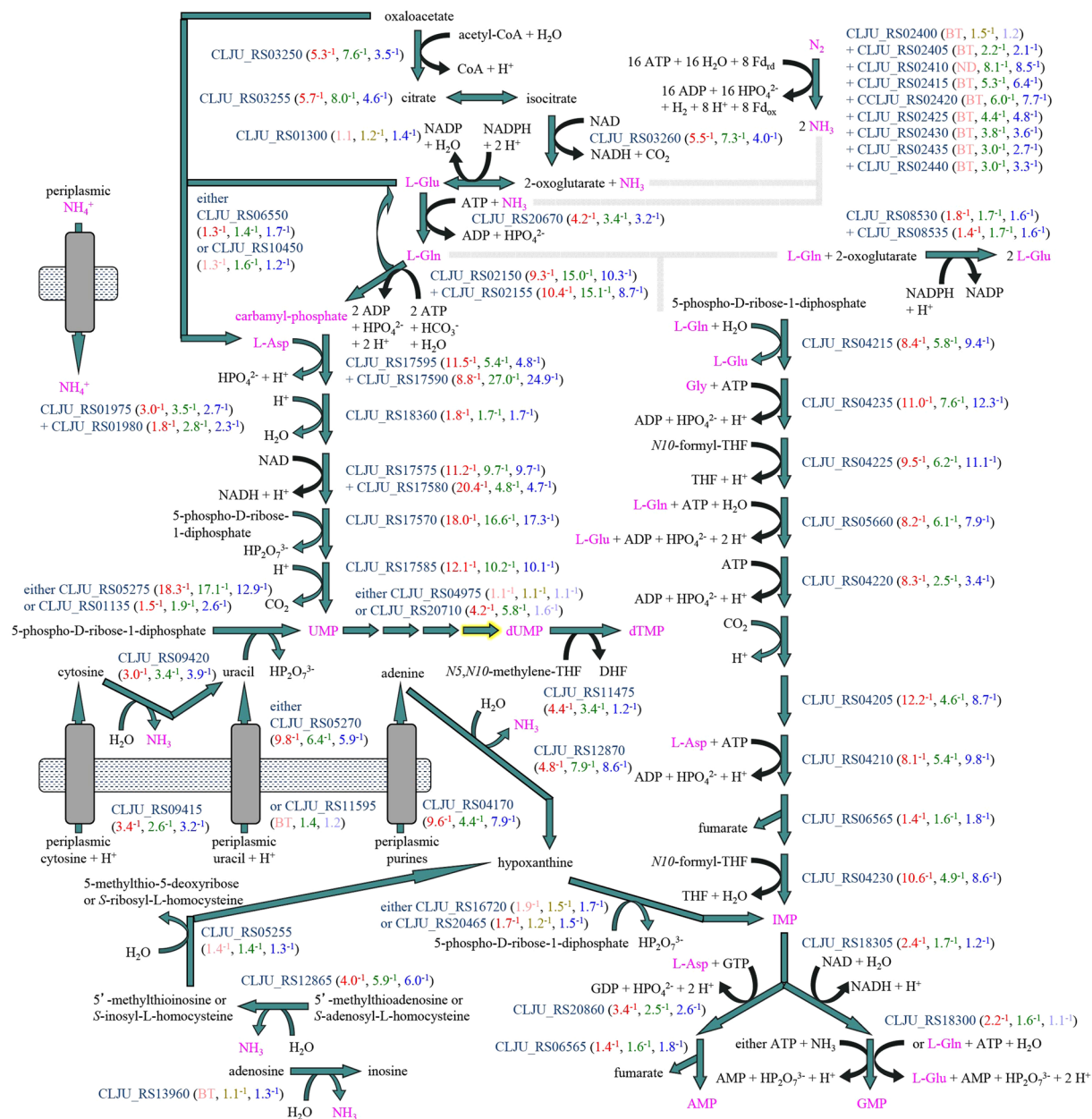


**Figure 5.** Downregulation of enzymes of glycolysis and methylglyoxal detoxification under lithotrophic conditions. Uptake of fructose is driven by conversion of phosphoenolpyruvate into pyruvate at the end of the glycolytic pathway, which otherwise yields one more ATP. Methylglyoxal synthase functions to release phosphate from glycerone-phosphate, producing methylglyoxal for reduction to (S)-1,2-propanediol, which is disproportionated to propanoyl-phosphate and 1-propanol inside a microcompartment within the cytoplasm.

**Upregulation of various membrane-associated proteins.** Genes for transfer of methyl groups from unidentified substrates into the Wood-Ljungdahl pathway were upregulated under lithotrophic conditions (Supplementary Discussion S3, Supplementary Table S10). Several strongly upregulated genes encode membrane proteins (Supplementary Table S11), although their exact functions cannot be predicted yet. A membrane protein of the major facilitator superfamily, possibly a transporter, is encoded by CLJU\_RS14525, which was the third most upregulated gene with  $H_2/CO_2$  according to RNA-Seq and the first most upregulated according to the microarray, and the sixteenth most upregulated with syngas (Supplementary Table S11). Ten adjacent genes were also strongly upregulated (Supplementary Table S11), including a poly-alpha-D-galacturonosidase gene (*pehX* CLJU\_RS14480) with 37% protein sequence identity to that of *Erwinia chrysanthemi*<sup>49</sup>. Poly-alpha-D-galacturonate is a component of plant cell walls; the upregulation of a gene for its degradation suggests that in the presence of  $H_2/CO_2$  or syngas, *C. ljungdahlii* may prepare to encounter plant cell matter. Considering that *C. ljungdahlii* was isolated from chicken yard waste<sup>1</sup>, it is possible that the digestive tract of a chicken is a natural environment where *C. ljungdahlii* degrades plant cell matter while growing lithotrophically with  $H_2$ ,  $CO_2$ , and CO.

**Downregulation of fructose-specific and glycolytic enzyme genes.** Predictably, the three most downregulated genes under lithotrophic conditions compared to organotrophic growth on fructose encode the fructose transport phosphotransferase system-specific membrane protein (*fruA* CLJU\_RS10110), the enzyme fructose-1-phosphate 6-kinase (*fruK* CLJU\_RS10115) that is only required for utilization of exogenous fructose (Fig. 5), and the fructose-1-phosphate-responsive transcriptional regulator (*fruR* CLJU\_RS10120, Supplementary Table S2). Genes for the glycolytic enzymes were also downregulated (Fig. 5; as with  $CO/CO_2$ <sup>17</sup> and  $H_2/CO_2$ <sup>18</sup>).

**Downregulation of glycerol/(S)-1,2-propanediol metabolism genes.** One strongly downregulated gene encodes an NADPH-dependent aldehyde reductase (*pdgQ* CLJU\_RS05830) with 40% protein sequence identity to the acetaldehyde/butanal reductase of *Clostridium saccharobutylicum*<sup>50</sup>. The location of this gene among genes for shell proteins of a microcompartment and for glycerol/(S)-1,2-propanediol dehydratase, most of which were downregulated with syngas and with  $H_2/CO_2$  (Fig. 5), suggests that it functions as 3-hydroxypropanal/propanal reductase. These genes may be expressed during growth on fructose to dispose of methylglyoxal, a toxic byproduct of glycolysis. Methylglyoxal may be reduced to (S)-1,2-propanediol in the cytoplasm and then disproportionated to 1-propanol and propanoyl-phosphate within the microcompartment (Fig. 5). Efforts to produce



**Figure 6.** Downregulation of pyrimidine and purine biosynthesis, salvage and uptake under lithotrophic conditions. Nitrogen being assimilated is shown in pink. Biosynthesis of pyrimidines and purines depends upon ammonium uptake, nitrogen fixation to ammonia and synthesis of oxaloacetate and 2-oxoglutarate, which are the carbon backbones of aspartate and glutamate/glutamine. Pyrimidine biosynthesis begins with one of two isozymes of carbamyl-phosphate synthase, different from the one in the urea cycle (Fig. 2).

solvents such as 1-propanol, 1-butanol, or 1,3-propanediol with *C. ljungdahlii* grown lithotrophically may benefit from genetic engineering to abolish downregulation of these genes.

**Downregulation of pyrimidine and purine biosynthesis, uptake and salvage genes.** Two operons encoding enzymes of pyrimidine biosynthesis and purine biosynthesis were among the most downregulated genes under lithotrophic conditions (Fig. 6). This is consistent with the expectation that growth in the absence of an organic carbon source, being slower, requires less nucleic acid synthesis for genome replication and gene transcription. Downregulation with syngas and with  $\text{H}_2/\text{CO}_2$  was also observed for genes of pyrimidine and purine biosynthesis, salvage and uptake in other locations in the genome (Fig. 6), including one of two isozymes of carbamyl-phosphate synthase. This downregulated isozyme appears to be dedicated to pyrimidine biosynthesis, while the other isozyme that was upregulated (Fig. 2) is dedicated to the urea cycle. Genes for biosynthesis of glutamate, glutamine and aspartate, which are substrates for pyrimidine and purine biosynthesis, were also downregulated, along with genes for ammonium uptake and nitrogen fixation (Fig. 6). Altogether,

these observations indicate a contrast between assimilation of nitrogen into nucleotides under organotrophic conditions and disposal of nitrogen as ammonia and urea under lithotrophic conditions. Genes for phosphate uptake and dephosphorylation of nucleotides were differentially regulated (Supplementary Discussion S4, Supplementary Table S12).

A homolog of the phosphoribosylaminoimidazolecarboxamide formyltransferase enzyme of purine biosynthesis is encoded by a gene (CLJU\_RS10900) that was the fourth most downregulated gene (65.3-fold) with H<sub>2</sub>/CO<sub>2</sub> according to RNA-Seq and the fifth most downregulated (26.2-fold) according to the microarray, consistent with a previous report<sup>18</sup>, but only 1.6-fold downregulated with syngas. The protein sequence divergence of this gene and its location apart from known purine biosynthesis genes suggest that it belongs to another pathway. On its 5' side is a *pfl* RNA motif, hypothesized to be a riboswitch to regulate one-carbon unit metabolism<sup>51</sup>, and on its 3' side is an adenosylcobamide-responsive riboswitch (transcribed with CO/CO<sub>2</sub>)<sup>17</sup> in an antiparallel orientation, suggesting that antisense transcription through CLJU\_RS10900 may occur when vitamin B12 is limiting. It will be interesting to investigate this gene product, which was abundant in a syngas proteomic sample<sup>24</sup>, for a role in one-carbon unit metabolism.

Genes for biosynthesis of threonine, a precursor of vitamin B12, were downregulated and high-threonine-content protein genes that were downregulated included some that are adjacent to adenosylcobamide-responsive riboswitches (Supplementary Discussion S5, Supplementary Tables S13, S14 and S15).

**Downregulation of cytoplasmic redox-active protein genes.** Several downregulated genes are predicted to encode cytoplasmic redox-active proteins (Supplementary Table S3), such as a rubredoxin domain oxidoreductase (*hrb* CLJU\_RS09545; also downregulated with CO/CO<sub>2</sub>)<sup>17</sup> with 47% protein sequence identity to Hrb of *Moorella thermoacetica*<sup>52</sup>. The Hrb protein oxidizes NADH and reduces rubredoxin:oxygen/nitric oxide oxidoreductase (*roo* CLJU\_RS10770), for which the gene was also downregulated. Both genes were previously reported to be upregulated with H<sub>2</sub>/CO<sub>2</sub> versus fructose<sup>18</sup> and under O<sub>2</sub> stress<sup>19</sup>. Downregulation of two genes that encode homologs of the iron-sulfur-oxygen hybrid cluster protein that is thought to detoxify an unidentified reactive compound in response to nitrous oxide stress<sup>53</sup> was more obvious with H<sub>2</sub>/CO<sub>2</sub> than with syngas.

**Downregulation of membrane-associated redox-active proteins with syngas.** Other downregulated genes encode membrane-associated redox-active proteins (Supplementary Table S4). Two genes that were downregulated with syngas (CLJU\_RS14585, CLJU\_RS14610) are notable for their protein sequence similarity to NapH, the ubiquinol-oxidizing subunit of the periplasmic nitrate reductase of *E. coli*<sup>25</sup>, and the predicted locations of iron-sulfur cluster-binding motifs within transmembrane segments or outside the cytoplasmic membrane. It would be interesting to study the roles of these genes in membrane-associated redox reactions, possibly including how electrical current is consumed.

## Conclusions

This study compared the transcriptomic profiles of *C. ljungdahlii* growing lithotrophically with H<sub>2</sub>/CO<sub>2</sub> or with syngas and organotrophically with fructose. The genes that were observed to be differentially regulated offer insights into physiological changes such as quorum sensing and sporulation, the roles of specific metabolic pathways such as the urea cycle, and differential usage of amino acids such as histidine. Numerous redox-active proteins of the cytoplasm and membrane were identified that may be investigated for roles in electron transfer across the cell wall of *C. ljungdahlii*.

## Materials and Methods

**Growth of *C. ljungdahlii*.** *C. ljungdahlii* DSM 13528 (ATCC 55383) was purchased from the German Collection of Microorganisms and Cell Cultures (DSMZ). For general propagation, *C. ljungdahlii* cells were grown anaerobically at 37 °C in PETC 1754 medium supplemented with 1 mM L-cysteine (pH 7) and fructose at 5 g/L. For transcriptomic studies, *C. ljungdahlii* cells were grown in DSMZ 879 medium supplemented with 0.04% L-cysteine, 0.04% sodium sulfide and 0.1% sodium bicarbonate, inoculated to OD<sub>600</sub>~0.05 with precultures grown to OD<sub>600</sub>~0.3–0.5 in DSMZ 879 medium under the same conditions (with fructose, H<sub>2</sub>/CO<sub>2</sub>, or syngas). For organotrophic growth, fructose at 5 g/L was added to DSMZ 879 medium. For lithotrophic growth, pressure tubes containing 5 ml instead of 10 ml of DSMZ 879 medium were used. For lithotrophic growth on H<sub>2</sub>/CO<sub>2</sub>, the headspace of culture tubes was replaced with H<sub>2</sub>/CO<sub>2</sub> (80/20) and pressurized to 20 psi after inoculation. For growth on syngas (H<sub>2</sub>:CO<sub>2</sub>:CO ratio 45:5:50), the headspace was replaced with syngas and pressurized to 20 psi. Cultures that grew lithotrophically were laid flat while shaking at 100 rpm during incubation at 37 °C. In experiments to determine the OD<sub>600</sub> range of mid-log growth, the headspace was pressurized with the appropriate gas mix every 24 hours after cell turbidity reached OD<sub>600</sub>~0.3.

**RNA isolation.** Total RNA was isolated from mid-log DSMZ 879-grown cultures of OD<sub>600</sub>~0.4–0.5 with fructose, OD<sub>600</sub>~0.2 with H<sub>2</sub>/CO<sub>2</sub>, and OD<sub>600</sub>~0.2 with syngas. When cultures reached the appropriate cell density, two volumes of RNeasy Protect Bacteria Reagent (Qiagen, USA) were anaerobically added to the cultures, mixed by vortex for 30 seconds, and incubated at room temperature for 10 minutes before collecting cells at 10,000 × *g* for 5 minutes. Cell pellets were snap-frozen in liquid nitrogen immediately after centrifugation and kept at –80 °C until RNA extraction. Total RNA was isolated with the RiboPure bacteria kit (Ambion, USA) following the manufacturer's instructions. The integrity and quantity of isolated total RNA were checked using a Bio-Rad Experion system (Bio-Rad, USA).

**cDNA synthesis.** cDNA was synthesized from total RNA with the TransPlex Complete Whole Transcriptome Amplification kit (Sigma-Adrich, USA) following the manufacturer's instructions.

**Microarray analysis and chip design.** Triplicate total RNA samples from each growth condition were collected for DNA microarray analysis. The microarray was designed using the published genome sequence (NC\_014328.1) and included 4184 CDS from the *C. ljungdahlii* ATCC55383 chromosome. In addition, probes for genes for the butanol biosynthesis pathway from *C. acetobutylicum* and other bacteria were also included in the microarray design (GeneID:1119056, 1118891, 1118895, 1118894, 1118892, 1118893, 1116040, 1116167, 1119481, 1119480, 1119259, 1119258, 2741560, 946727, 4413431, 5294993, 5291557, 124221917, 5292938, and 9265716). The probe sets were designed by NimbleGen (USA) according to the company's protocols and algorithms, and all probes were manufactured directly onto the slides via photolithographic synthesis of 45-to-60-mer oligonucleotides. For every gene sequence, up to 8 individual probes were designed to cover the gene sequence of interest (for detailed information refer to Supplementary Table S16). Each probe was synthesized in two replicates located in different array regions (BLOCK1 and BLOCK2), resulting in 66,800 unique probe sets covering 4204 genes. Duplicated genes in the *C. ljungdahlii* genome that were omitted from probe sets are listed in Supplementary Table S16.

**mRNA enrichment and Illumina library preparation for RNA-Seq.** Total RNA isolated from triplicate cultures ( $H_2/CO_2$  and fructose, different from the cultures used for microarray analysis) was used for an enrichment of mRNA using the *MICROBExpress* kit (Ambion, USA), following the manufacturer's protocol. Quality and quantity of the enriched mRNA were checked by analyzing aliquots of mRNA using the Experion RNA HiSens kit (Bio-Rad, USA), following the manufacturer's protocol (Supplementary Fig. S1). The enriched mRNA samples thus obtained were used to construct Illumina libraries ( $3 \times 2$  samples) using the ScriptSeq™ v2 RNA-Seq Library Preparation Kit (Epicentre, USA), which enables directional sequencing, following the manufacturer's protocol. Briefly, 200 ng of the total mRNA were chemically fragmented and converted into single-stranded cDNA with random hexamer priming. Next, the single-stranded cDNA was 3'-terminal-tagged. Di-tagged cDNA products were then amplified and enriched by using individual adapters containing unique hexameric indices/barcodes with a final 10-cycle PCR reaction. Altogether, six enriched and purified libraries with unique barcodes (representing three replicates with either  $H_2/CO_2$  or fructose) were obtained, mixed in an equimolar concentration, and used for hybridization in a HiSeq 2000 flow cell for single-end sequencing.

**Assembly of Illumina reads.** All of the raw data generated by Illumina sequencing were quality checked by visualization of base quality scores and nucleotide distributions. Then the sequences were sorted out by trimming of reads and read filtering based on base quality score and sequence properties such as primer contaminations, N content and GC bias using PRINSEQ<sup>54</sup>. All good quality mRNA sequence reads were assembled and mapped against the published genome of *C. ljungdahlii*, DSM 13528 (NC\_014328) using ARRAY STAR (DNA star, USA). Reads belonging to 16S/23S rRNA, reads that matched more than one segment of the genome and reads with more than two mismatches were discarded. The remaining mRNA reads were reanalyzed and normalized with the RPKM (reads assigned per kilobase of target per million mapped reads) method<sup>55,56</sup> using ARRAY STAR. Reads from biological replicates were first compared with each other graphically after mapping onto the template genome. Biological replicates were highly reproducible (Supplementary Fig. S1). Therefore, reads from biological replicates were merged and averaged for each experimental condition. Expression levels were considered significant only when the  $\log_2$  RPKM value was  $\geq 5$  (median  $\log_2$  RPKM value) (Supplementary Fig. S2; Supplementary Table S1) in one of the experiments and were used to calculate fold change (Supplementary Fig. S2). The significance level for downregulation or upregulation was calculated at  $p = 0.05$ .

**Data Availability.** Microarray data and RNA-Seq sequence reads have been submitted to the EMBL databases under accession nos. E-MTAB-3806 and PRJEB9771, respectively.

## References

- Tanner, R. S., Miller, L. M. & Yang, D. *Clostridium ljungdahlii* sp. nov., an acetogenic species in clostridial rRNA homology group I. *Int. J. Syst. Bacteriol.* **43**, 232–236 (1993).
- Lovley, D. R. & Nevin, K. P. Electrobiocommodities: powering microbial production of fuels and commodity chemicals from carbon dioxide with electricity. *Curr. Opin. Biotechnol.* **24**, 385–90 (2013).
- Nevin, K. P. *et al.* Electrosynthesis of organic compounds from carbon dioxide is catalyzed by a diversity of acetogenic microorganisms. *Appl. Environ. Microbiol.* **77**, 2882–2886 (2011).
- Banerjee, A., Leang, C., Ueki, T., Nevin, K. P. & Lovley, D. R. Lactose-inducible system for metabolic engineering of *Clostridium ljungdahlii*. *Appl. Environ. Microbiol.* **80**, 2410–2416 (2014).
- Dürre, P. & Eikmanns, B. J. C1-carbon sources for chemical and fuel production by microbial gas fermentation. *Curr. Opin. Biotechnol.* **35**, 63–72 (2015).
- Leang, C., Ueki, T., Nevin, K. P. & Lovley, D. R. A genetic system for *Clostridium ljungdahlii*: a chassis for autotrophic production of biocommodities and a model homoacetogen. *Appl. Environ. Microbiol.* **79**, 1102–1109 (2013).
- Molitor, B., Kirchner, K., Henrich, A. W., Schmitz, S. & Rosenbaum, M. A. Expanding the molecular toolkit for the homoacetogen *Clostridium ljungdahlii*. *Sci. Rep.* **6**, 31518 (2016).
- Ueki, T., Nevin, K. P., Woodard, T. L. & Lovley, D. R. Converting carbon dioxide to butyrate with an engineered strain of *Clostridium ljungdahlii*. *MBio* **5**, e01636-14–e01636-14 (2014).
- Köpke, M. *et al.* *Clostridium ljungdahlii* represents a microbial production platform based on syngas. *Proc. Natl. Acad. Sci. USA.* **107**, 13087–13092 (2010).
- Meier, T., Ferguson, S. A., Cook, G. M., Dimroth, P. & Vonck, J. Structural investigations of the membrane-embedded rotor ring of the F-ATPase from *Clostridium paradoxum*. *J. Bacteriol.* **188**, 7759–7764 (2006).
- Herrmann, G., Jayamani, E., Mai, G. & Buckel, W. Energy conservation via electron-transferring flavoprotein in anaerobic bacteria. *J. Bacteriol.* **190**, 784–791 (2008).
- Mock, J. *et al.* Energy conservation associated with ethanol formation from  $H_2$  and  $CO_2$  in *Clostridium autoethanogenum* involving electron bifurcation. *J. Bacteriol.* **197**, 2965–2980 (2015).
- Wang, S. *et al.* NADP-specific electron-bifurcating [FeFe]-hydrogenase in a functional complex with formate dehydrogenase in *Clostridium autoethanogenum* grown on  $CO$ . *J. Bacteriol.* **195**, 4373–4386 (2013).



14. Tremblay, P. L., Zhang, T., Dar, S. A., Leang, C. & Lovley, D. R. The Rnf complex of *Clostridium ljungdahlii* is a proton-translocating ferredoxin:NAD<sup>+</sup> oxidoreductase essential for autotrophic growth. *MBio* 4, (e00406–12) (2012).
15. Wang, S., Huang, H., Moll, J. & Thauer, R. K. NADP<sup>+</sup> reduction with reduced ferredoxin and NADP<sup>+</sup> reduction with NADH are coupled via an electron-bifurcating enzyme complex in *Clostridium kluyveri*. *J Bacteriol* 192, 5115–5123 (2010).
16. Burroughs, A. M., Balaji, S., Iyer, L. M. & Aravind, L. A novel superfamily containing the beta-grasp fold involved in binding diverse soluble ligands. *Biol. Direct* 2, 4 (2007).
17. Tan, Y., Liu, J., Chen, X., Zheng, H. & Li, F. RNA-seq-based comparative transcriptome analysis of the syngas-utilizing bacterium *Clostridium ljungdahlii* DSM 13528 grown autotrophically and heterotrophically. *Mol. Biosyst.* 9, 2775–84 (2013).
18. Nagarajan, H. *et al.* Characterizing acetogenic metabolism using a genome-scale metabolic reconstruction of *Clostridium ljungdahlii*. *Microb. Cell Fact.* 12, 118 (2013).
19. Whitham, J. M., Tirado-Acevedo, O., Chinn, M. S., Pawlak, J. J. & Grunden, A. M. Metabolic response of *Clostridium ljungdahlii* to oxygen exposure. *Appl. Environ. Microbiol.* 81, 8379–8391 (2015).
20. Philips, J., Rabaey, K., Lovley, D. R. & Vargas, M. Biofilm formation by *Clostridium ljungdahlii* is induced by sodium chloride stress: experimental evaluation and transcriptome analysis. *PLoS One* 12, e0170406 (2017).
21. Wei, Y., Guffanti, A. A., Ito, M. & Krulwich, T. A. *Bacillus subtilis* YqkI is a novel malic/Na<sup>+</sup>-lactate antiporter that enhances growth on malate at low protonmotive force. *J Biol Chem* 275, 30287–30292 (2000).
22. Meyer, F. M. & Stulke, J. Malate metabolism in *Bacillus subtilis*: Distinct roles for three classes of malate-oxidizing enzymes. *FEMS Microbiol Lett* (2012).
23. Lerondel, G., Doan, T., Zamboni, N., Sauer, U. & Aymerich, S. YtsJ has the major physiological role of the four paralogous malic enzyme isoforms in *Bacillus subtilis*. *J Bacteriol* 188, 4727–4736 (2006).
24. Richter, H. *et al.* Ethanol production in syngas-fermenting *Clostridium ljungdahlii* is controlled by thermodynamics rather than by enzyme expression. *Energy Environ. Sci.* 9, 2392–2399 (2016).
25. Brondijk, T. H., Nilavongse, A., Filenko, N., Richardson, D. J. & Cole, J. A. NapGH components of the periplasmic nitrate reductase of *Escherichia coli* K-12: location, topology and physiological roles in quinol oxidation and redox balancing. *Biochem J* 379, 47–55 (2004).
26. Hamann, N. *et al.* The CCG-domain-containing subunit SdhE of succinate:quinone oxidoreductase from *Sulfolobus solfataricus* P2 binds a [4Fe-4S] cluster. *J Biol Inorg Chem* 14, 457–470 (2009).
27. Hamann, N. *et al.* A cysteine-rich CCG domain contains a novel [4Fe-4S] cluster binding motif as deduced from studies with subunit B of heterodisulfide reductase from *Methanothermobacter marburgensis*. *Biochemistry* 46, 12875–12885 (2007).
28. Gibello, A., Collins, M. D., Dominguez, L., Fernandez-Garayzabal, J. F. & Richardson, P. T. Cloning and analysis of the L-lactate utilization genes from *Streptococcus iniae*. *Appl. Environ. Microbiol.* 65, 4346–4350 (1999).
29. Desguin, B. *et al.* Lactate racemase is a nickel-dependent enzyme activated by a widespread maturation system. *Nat. Commun.* 5, 3615 (2014).
30. Lord, J. M. Glycolate oxidoreductase in *Escherichia coli*. *Biochim Biophys Acta* 267, 227–237 (1972).
31. Weghoff, M. C., Bertsch, J. & Müller, V. A novel mode of lactate metabolism in strictly anaerobic bacteria. *Environ. Microbiol.* <https://doi.org/10.1111/1462-2920.12493> (2014).
32. Köpke, M. *et al.* Reconstruction of an acetogenic 2,3-butanediol pathway involving a novel NADPH-dependent primary-secondary alcohol dehydrogenase. *Appl. Environ. Microbiol.* 80, 3394–3403 (2014).
33. Burguière, P., Auger, S., Hullo, M. F., Danchin, A. & Martin-Verstraete, I. Three different systems participate in L-cystine uptake in *Bacillus subtilis*. *J Bacteriol* 186, 4875–4884 (2004).
34. Butler, J. D., Levin, S. W., Facchiano, A., Miele, L. & Mukherjee, A. B. Amino acid composition and N-terminal sequence of purified cystine binding protein of *Escherichia coli*. *Life Sci* 52, 1209–1215 (1993).
35. Auger, S., Gomez, M. P., Danchin, A. & Martin-Verstraete, I. The PatB protein of *Bacillus subtilis* is a C-S-lyase. *Biochimie* 87, 231–238 (2005).
36. Gal, J., Szvetnik, A., Schnell, R. & Kalman, M. The *metD* D-methionine transporter locus of *Escherichia coli* is an ABC transporter gene cluster. *J Bacteriol* 184, 4930–4932 (2002).
37. Hullo, M. F., Auger, S., Dassa, E., Danchin, A. & Martin-Verstraete, I. The *metNPQ* operon of *Bacillus subtilis* encodes an ABC permease transporting methionine sulfoxide, D- and L-methionine. *Res Microbiol* 155, 80–86 (2004).
38. Huang, C. J. & Barrett, E. L. Sequence analysis and expression of the *Salmonella typhimurium* *asr* operon encoding production of hydrogen sulfide from sulfite. *J Bacteriol* 173, 1544–1553 (1991).
39. Holland, S. L. & Avery, S. V. Chromate toxicity and the role of sulfur. *Metallomics* 3, 1119–1123 (2011).
40. Diaz-Magaña, A. *et al.* Short-chain chromate ion transporter proteins from *Bacillus subtilis* confer chromate resistance in *Escherichia coli*. *J Bacteriol* 191, 5441–5445 (2009).
41. Martinez-Valencia, R., Reyes-Cortes, G., Ramirez-Diaz, M. I., Riveros-Rosas, H. & Cervantes, C. Antiparallel membrane topology of paired short-chain chromate transport proteins in *Bacillus subtilis*. *FEMS Microbiol Lett* 336, 113–121 (2012).
42. von Heijne, G. Membrane protein structure prediction, hydrophobicity analysis and the positive-inside rule. *J. Mol. Biol.* 225, 487–494 (1992).
43. Steiner, E., Scott, J., Minton, N. P. & Winzer, K. An *agr* quorum sensing system that regulates granule formation and sporulation in *Clostridium acetobutylicum*. *Appl. Environ. Microbiol.* 78, 1113–1122 (2011).
44. Sturme, M. H. *et al.* An *agr*-like two-component regulatory system in *Lactobacillus plantarum* is involved in production of a novel cyclic peptide and regulation of adherence. *J Bacteriol* 187, 5224–5235 (2005).
45. Ravagnani, A. *et al.* Spo0A directly controls the switch from acid to solvent production in solvent-forming clostridia. *Mol. Microbiol.* 37, 1172–85 (2000).
46. Steiner, E. *et al.* Multiple orphan histidine kinases interact directly with Spo0A to control the initiation of endospore formation in *Clostridium acetobutylicum*. *Mol Microbiol* 80, 641–654 (2011).
47. Yoshida, K., Fujita, Y. & Ehrlich, S. D. Three asparagine synthetase genes of *Bacillus subtilis*. *J Bacteriol* 181, 6081–6091 (1999).
48. Saad, N. Y. *et al.* Riboswitch (T-box)-mediated control of tRNA-dependent amidation in *Clostridium acetobutylicum* rationalizes gene and pathway redundancy for asparagine and asparaginyl-tRNA<sup>Asn</sup> synthesis. *J Biol Chem* 287, 20382–20394 (2012).
49. He, S. Y. & Collmer, A. Molecular cloning, nucleotide sequence, and marker exchange mutagenesis of the exo-poly-alpha-D-galacturonidase-encoding *pehX* gene of *Erwinia chrysanthemi* EC16. *J Bacteriol* 172, 4988–4995 (1990).
50. Youngleson, J. S., Jones, W. A., Jones, D. T. & Woods, D. R. Molecular analysis and nucleotide sequence of the *adh1* gene encoding an NADPH-dependent butanol dehydrogenase in the Gram-positive anaerobe *Clostridium acetobutylicum*. *Gene* 78, 355–364 (1989).
51. Meyer, M. M. *et al.* Challenges of ligand identification for riboswitch candidates. *RNA Biol* 8, 5–10 (2011).
52. Silaghi-Dumitrescu, R. *et al.* A flavodiiron protein and high molecular weight rubredoxin from *Moorella thermoacetica* with nitric oxide reductase activity. *Biochemistry* 42, 2806–2815 (2003).
53. Filenko, N. A., Browning, D. F. & Cole, J. A. Transcriptional regulation of a hybrid cluster (prismane) protein. *Biochem Soc Trans* 33, 195–197 (2005).
54. Schmieder, R. & Edwards, R. Quality control and preprocessing of metagenomic datasets. *Bioinformatics* 27, 863–4 (2011).
55. Klevebring, D., Bjursell, M., Emanuelsson, O. & Lundeberg, J. In-depth transcriptome analysis reveals novel TARs and prevalent antisense transcription in human cell lines. *PLoS One* 5, e9762 (2010).
56. Mortazavi, A., Williams, B. A., McCue, K., Schaeffer, L. & Wold, B. Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat. Methods* 5, 621–8 (2008).

## Acknowledgements

We thank Cory Gillis, Manju Sharma and Joy Ward for excellent technical support. This research was supported by the Advanced Research Projects Agency-Energy (ARPA-E), U.S. Department of Energy, under award no. DE-AR0000159 and grant N000141310549 from the Office of Naval Research.

## Author Contributions

M.A. performed manual curation of the genome annotation, noted observations, and wrote the manuscript. C.L. designed the experiments, grew cultures for RNA isolation, carried out total RNA extraction, set up microarray analysis, and participated in data analysis and writing. P.M.S. did mRNA. enrichment, RNA-Seq library preparation, and RNA-Seq data analysis. M.S. assisted with illustrations. The project was conceived and guided by D.R.L. All authors read, assisted with editing, and approved the final manuscript.

## Additional Information

**Supplementary information** accompanies this paper at <https://doi.org/10.1038/s41598-017-12712-w>.

**Competing Interests:** The authors declare that they have no competing interests.

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2017