The reductive glycine pathway allows autotrophic growth of Desulfovibrio desulfuricans

Irene Sánchez-Andrea1✉, Iame Alves Guedes1,7, Bastian Hornung2,7, Sjef Boeren3, Christopher E. Lawson4, Diana Z. Sousa1, Arren Bar-Even5,8, Nico J. Claassens1,5✉ & Alfons J. M. Stams1,6✉

Six CO₂ fixation pathways are known to operate in photoautotrophic and chemoautotrophic microorganisms. Here, we describe chemolithoautotrophic growth of the sulphate-reducing bacterium Desulfovibrio desulfuricans (strain G11) with hydrogen and sulphate as energy substrates. Genomic, transcriptomic, proteomic and metabolomic analyses reveal that D. desulfuricans assimilates CO₂ via the reductive glycine pathway, a seventh CO₂ fixation pathway. In this pathway, CO₂ is first reduced to formate, which is reduced and condensed with a second CO₂ to generate glycine. Glycine is further reduced in D. desulfuricans by glycine reductase to acetyl-P, and then to acetyl-CoA, which is condensed with another CO₂ to form pyruvate. Ammonia is involved in the operation of the pathway, which is reflected in the dependence of the autotrophic growth rate on the ammonia concentration. Our study demonstrates microbial autotrophic growth fully supported by this highly ATP-efficient CO₂ fixation pathway.
Desulfovibrio is a genus of sulphate-reducing bacteria that is ubiquitous in oligotrophic and eutrophic environments. These bacteria couple the reduction of sulphate with the oxidation of a variety of electron donors such as lactate, ethanol, formate and hydrogen, but not acetate. Growth of Desulfovibrio sp. on formate and hydrogen was thought to be dependent on the presence of an organic carbon source (e.g., acetate)\(^2\)\(^-\)\(^7\). However, *D. desulfuricans* strain FI\(^1\) and the *Desulfovibrio* strains HRM1 and P23\(^8\)\(^-\)\(^1\(^0\) were reported to grow autotrophically, though insight into the CO\(_2\) fixation pathway was not obtained.

In a previous study, *D. desulfuricans* strain G11 (hereafter *D. desulfuricans*) grew syntrophically with the methanogen *Methanobrevibacter arborophilus* strain AZ (DSM 744) on formate as carbon and energy source\(^1\(^1\). Based on this co-culture experiment, it could be hypothesised that *D. desulfuricans* either is able to grow autotrophically (or formatotrophically), or alternatively that it uses an organic compound excreted by the co-cultured partner. Autotrophic growth had never been demonstrated for a pure culture of *D. desulfuricans* strain G11.

Here, we were able to grow *D. desulfuricans* autotrophically, and investigated the CO\(_2\) fixation pathway involved in autotrophic growth. Genomic, transcriptomic, proteomic and metabolomic analyses reveal that *D. desulfuricans* assimilates CO\(_2\) via the reductive glycine (rGly) pathway, a previously proposed yet unconfirmed seventh CO\(_2\) fixation pathway.

**Results**

*D. desulfuricans* is able to grow autotrophically. We transferred *D. desulfuricans* to chemolithoautotrophic conditions in a basic anaerobic mineral medium\(^1\(^2\) with sulphate and a gas phase consisting of 80% H\(_2\) and 20% CO\(_2\) (further referred to as autotrophic conditions). The inoculum was a culture growing in chemolithoheterotrophic conditions (acetate/H\(_2\)/CO\(_2\)/sulphate, further referred to as heterotrophic). After the initial transfer, growth was observed, but only after a long lag phase (>6 days) and with a low yield (OD\(_{600}\) < 0.12). However, after several transfers, the growth characteristics in autotrophic conditions improved and reached those observed in heterotrophic conditions with a doubling time of ~24 h (Fig. 1 and Supplementary Data 1). Periodical transfer of the autotrophic cultures in the late exponential phase (with 10% of inoculum) successfully avoided long lag phases. We investigated if laboratory evolution took place along long-term cultivation. Hence, we again transferred a cell population, which was transferred heterotrophically for 3 years and not experienced autotrophic conditions before, to autotrophic conditions. During the first transfer it showed again a long lag phase, but after a second autotrophic transfer in the late exponential phase, a short lag phase and fast growth were observed, similar to the growth of the autotrophic culture that was transferred over 3 years (Supplementary Data 1). This indicates that the long-term transfer is not needed to reach fast autotrophic growth and that this phenotype is not based on genetic mutations. Indeed, no genetic mutations could be identified in the genome sequence of the autotrophic culture after 35 transfers when compared with the heterotrophic culture (both after 3 years of subcultivation), confirming that no laboratory evolution took place. We also transferred the autotrophic culture to formatotrophic conditions (formate/CO\(_2\)/sulphate) and observed good growth (Supplementary Data 1). The purity of the cultures was routinely confirmed using microscopy and sequencing of 16S rRNA gene.

No known CO\(_2\) fixation route is fully encoded in the genome. We studied the genome of *D. desulfuricans* to identify genes involved in autotrophy, including possible metabolic routes for CO\(_2\) fixation. Analysis of the genome of *D. desulfuricans* confirmed the presence of most of the common biosynthesis pathways, including those for synthesis of amino acids and vitamins (Supplementary Data 2 and Supplementary Table 1). However, the bacterium lacks key enzymes of the known CO\(_2\) fixation pathways\(^1\(^3\)\(^-\)\(^7\), including ribulose-1,5-bisphosphate carboxylase/oxygenase (Rubisco) of the Calvin cycle and acetyl-CoA synthase/monoxide dehydrogenase of the reductive acetyl-CoA pathway (Supplementary Table 2). As the *D. desulfuricans* genome lacks malate dehydrogenase and succinyl-CoA synthetase, it cannot operate the reductive TCA cycle\(^1\(^3\)\) or any of its variants\(^1\(^6\)\(^-\)\(^1\(^7\). The lack of genes for known pathways for CO\(_2\) fixation is also observed in other *Desulfovibrio* members (Supplementary Data 3). Indeed, most *Desulfovibrio* members do not grow autotrophically with hydrogen, with a few exceptions mentioned above\(^6\)\(^-\)\(^1\(^8\). The observed autotrophic growth of *D. desulfuricans* indicates the presence of an unknown CO\(_2\) fixation pathway in this microorganism.

Comparative omics suggest CO\(_2\) fixation via glycine. To elucidate the identity of the CO\(_2\) fixation pathway without solely relying on genome annotation, we used a comparative multi-omics approach. We first compared the transcriptome and proteome of cells grown autotrophically (H\(_2\)/CO\(_2\)/sulphate) and heterotrophically (acetate/H\(_2\)/CO\(_2\)/sulphate). We consistently found several genes and proteins that were highly upregulated under autotrophic conditions (Figs. 2, 3, Supplementary Fig. 1 and Supplementary Data 4, 5).

Highly overexpressed genes include those encoding the glycine cleavage system (GCS, DsvG11_0325-0328), the glycine reductase (GR) complex (DsvG11_1441-1448) as well as formate–tetrahydrofolate (THF) ligase (FTL, DsvG11_3068), acetyl-CoA synthetase (ACS, DsvG11_2043) and genes encoding proteins associated with ammonia limitation (GS-GOGAT system). The same genes were also upregulated in autotrophically grown cells when compared with cells grown heterotrophically with lactate as sole energy source (without H\(_2\)) (Supplementary Fig. 1). Hence, we propose that these upregulated genes are involved in autotrophic CO\(_2\) fixation as depicted in Fig. 3 and Supplementary Fig. 2.

The GCS is a ubiquitous complex that catalyses the reversible cleavage of glycine to CO\(_2\), methylene-THF and ammonia (\(\Delta G^\circ = +4.6 \text{kJ} \text{per mol at pH 7.5, ionic strength 0.25 and all reactants at 1 mM}\))\(^1\(^8\). The GCS is known to operate in the reductive carboxylation direction in various anaerobic microorganisms as a sink of electrons\(^2\)\(^-\)\(^2\(^2\), thus producing glycine from CO\(_2\) and methylene-THF. FTL catalyses the condensation of formate with THF and, together with the bifunctional methylene-THF dehydrogenase/cyclohydrolase (MTDC), can convert formate to methylene-THF, the direct substrate of GCS for reductive carboxylation. Hence, the combined reductive activity of FTL, MTDC and GCS can result in the conversion of formate, ammonia and CO\(_2\) to glycine.

For CO\(_2\) to be converted to glycine, it should be first reduced to formate. The molybdenum-containing formate dehydrogenase (FDH) of *D. desulfuricans* (DsvG11_0566-0569) was described to catalyse CO\(_2\) reduction to formate with very good kinetic parameters \((k_{\text{cat}} \text{ of } 46.6 \text{s}^{-1}, \text{K}_{\text{m}} \text{ of } 15.7 \mu\text{M})\)\(^2\(^3\). FDH is highly expressed at all conditions tested (Supplementary Data 4), while some of its subunits are upregulated during autotrophic growth. Moreover, the observed accumulation of formate in the medium under autotrophic conditions indicates the reductive activity of FDH (Fig. 1a). Overall, these findings suggest that *D. desulfuricans* converts CO\(_2\) to glycine via the combined operation of FDH, the THF-dependent enzymes and the enzymes of the GCS.
Assimilation of glycine into the central metabolism could proceed via two main routes: (i) glycine conversion to acetyl-phosphate, catalysed by GR (DSVG11_1441-1448), which can be further converted to acetyl-CoA and pyruvate; (ii) glycine conversion to serine via serine hydroxymethyltransferase (SHMT, DsvG11_2276). Serine could be further assimilated into the central metabolism via several routes, the most direct one employing serine deaminase (SDA, DsvG11_1577) to form pyruvate. Alternatively, serine could be assimilated via conversion to phospho-glycerate, which could proceed via serine transamination to hydroxypyruvate (e.g., by pyruvate-serine transaminase DSVG11_0309 or other transaminases), subsequent reduction of hydroxypyruvate to glycerate (potentially by 2-hydroxyacid dehydrogenases DSVG11_0256 or DSVG11_0961) and finally the generation of phospho-glycerate by glycerate kinase (DSVG11_0656 or DSVG11_1884). During autotrophic growth, assimilation probably proceeds primarily via the glycine reduction route rather than the serine route, as GR was highly upregulated, while SHMT, SDA and alternative serine assimilation enzymes were not upregulated and expressed at a lower level than GR.

Acetyl-phosphate generated from glycine can be converted to acetyl-CoA either directly, via phosphoacetyltransferase (PTA, DsvG11_0941), or indirectly, via the combined activity of acetate kinase (ACK, DsvG11_0942) and acetyl-CoA synthetase (ACS, DsvG11_2043). The high up-regulation of ACS under autotrophic conditions, and the accumulation of acetate in the growth medium, indicate that the indirect route is the major pathway of acetyl-CoA biosynthesis. Finally, acetyl-CoA can be converted to pyruvate via the activity of pyruvate:ferredoxin oxidoreductase (PFOR, DsvG11_0940), which was highly expressed under all growth conditions in the transcriptome, as well as abundant in the proteome. Pyruvate can also be synthesised by condensation of acetyl-CoA and formate\textsuperscript{24}. D. desulfuricans possesses two putative pyruvate-formate lyases (PFL, DSVG11_0854 and DSVG11_2562), of which DSVG11_0854 is slightly upregulated at transcript level in autotrophic conditions (Supplementary Data 4 and 5). However, as both PFL enzymes are overall having low expression levels at transcript level and not detected in the proteome, we assume they play a minor or no role in pyruvate synthesis.

**Fig. 1** Comparison of growth of D. desulfuricans under autotrophic and heterotrophic conditions. Autotrophic growth (H₂/CO₂/sulphate) after some adaptation transfers (a) is similar to heterotrophic growth (acetate/H₂/CO₂/sulphate (b). Increasing ammonium concentrations do have an effect on the autotrophic growth (c), but not on heterotrophic growth (d). All growth experiments were performed at 30 °C and 175 rpm in triplicates in 250 ml glass bottles containing 100 ml anoxic minimal medium. Error bars represent the standard deviation. Source data are provided with this paper, and raw data can be found in Supplementary Data 1.
The observation that besides formate about 1 mM of acetate is excreted during autotrophic growth is remarkable (Fig. 1), as the addition of 2 mM acetate results in heterotrophic growth. Currently, it is not clear at which concentration the switch from heterotrophic to autotrophic growth and vice versa occurs. Acetate excretion was observed previously also in an autotrophic co-culture of *D. desulfuricans* strain F1 with the non-autotrophic dechlorinating bacterium *Dehalococcoides mccartyi*. The CO₂ assimilation pathway in *Desulfovibrio* strain F1 was not investigated, but it produced formate and acetate when grown autotrophically⁸, suggesting that it may use the same pathway as described here for *D. desulfuricans* strain G11. Further research is needed to shed light on the metabolic and ecologic significance of acetate secretion.

Autotrophic growth improves at higher ammonia concentration. Remarkably, genes related to ammonia limitation²⁵ were also strongly upregulated when grown autotrophically, although the same ammonia concentration was provided in both heterotrophic and autotrophic growth conditions. These include an ammonia transporter (AT, DsvG11_2780), the nitrogen-regulatory protein (DsvG11_2781), glutamine synthetase (GS, DsvG11_2782) and glutamate-oxoglutarate aminotransferase (GOGAT, DsvG11_2965-67) (Supplementary Fig. 2 and Supplementary Data 4, 5). The up-regulation of the AT may be related to the requirement of the GCS for a high intracellular concentration of ammonia: the kinetically and thermodynamically less favourable, reductive, glycine-producing direction of the GCS could benefit from higher intracellular ammonia concentrations¹⁸. However, it is not completely clear why the GS-GOGAT ammonia assimilation route would be upregulated and the glutamate dehydrogenase (GDH) down-regulated under autotrophic conditions. The ammonia concentrations used in both autotrophic and heterotrophic experiments are likely sufficient to operate ammonia assimilation via the more energy-efficient GDH route. Possibly, GS and GOGAT are upregulated together with the AT in autotrophic conditions as they are in the same operon.
Based on these observations, we hypothesised that autotrophic growth of *D. desulfuricans* is strongly dependent on the concentration of ammonia provided in the medium. We tested the effect of ammonia on growth and found that, while heterotrophic growth was not affected by ammonia concentrations (Fig. 1d and Supplementary Data 1), the growth rate under autotrophic conditions increased with increasing ammonia (Fig. 1c). Importantly, the final biomass yield under autotrophic conditions did not change with increasing ammonia concentrations. This is to be expected as excess ammonia was not assimilated to biomass, but ammonia is temporarily fixed by the GCS and then released by GR.
Fig. 3 Comparative omics for genes and proteins involved in the reductive glycine pathway. Plots per enzyme represent the log10-fold change in autotrophic condition (H2/CO2/sulphate) versus heterotrophic condition (acetate/H2/CO2/sulphate) and versus heterotrophic growth on lactate as sole energy source (lactate/CO2/sulphate), for both proteome and transcriptome analysis. Long-dashed lines are the transporters involved, short dashes indicate the alternative variant route via serine. Cultures were performed in four biological replicates. Growth conditions were 30 °C and 175 rpm in 250 mL glass bottles containing 100 mL anoxic minimal medium. Abbreviations are described in the legend of Fig. 2, with the addition of: ATP/ADP (adenosine triphosphate); FDHap (formate dehydrogenase, accessory protein); FT (formate transporter); GR_Bb (glycine reductase complex; component B, subunit alpha); NAD(P)H (nicotinamide adenine dinucleotide (phosphate)); SDA (serine dehydratase-like); THF: tetrahydrofolate; ND: not detected. Source data are provided with this paper, and complete transcriptomics and proteomics data can be found in Supplementary Data 4 and 5, respectively.

Metabolomics confirms the proposed rGly pathway. All the enzymatic steps in the proposed pathway from CO2 to glycine are reversible, and were previously proven in vitro to catalyse the reactions in the direction required for CO2 fixation. To demonstrate the operation of proposed CO2 fixation pathway in vivo, we performed 13C-labelling experiments. We cultivated D. desulfuricans on H2, sulphate and unlabelled CO2, and added the 13C-isotope form of the pathway intermediate formate. This strategy was intended to generate partially labelled proteinogenic amino acids. As some of the amino acid carbons originate from formate while others from CO2, the labelling pattern of different metabolites depends on their metabolic origin (if labelled CO2 would be used instead, all metabolites would be fully labelled, thus providing no information on their metabolic origin). A key signature of this pathway would be the generation of single-labelled glycine and double-labelled serine, as indeed was observed (Fig. 4c) confirms that the GR route (together with PFOR) is the primary one for carbon assimilation in our strain. Furthermore, we analysed the labelling pattern of threonine and proline, which reflect the labelling of their precursors oxaloacetate and oxoglutarate, respectively. The labelling pattern of these two amino acids also supports the activity of GR (Fig. 4a, b). The labelling of proline and threonine also confirms the absence of a functional (reductive) TCA cycle, which would have shuffled the labelling of oxaloacetate and oxoglutarate. In fact, the labelling pattern of the amino acids shows that CO2 is not fixed via the reductive TCA cycle, in which case we would expect to observe only partial labelling of all amino acids, as formate is not directly assimilated and labelling could originate only from the assimilation of CO2 derived from formate oxidation.

Tracer experiments were also performed by adding 15 mM of 13C-formate to cells at the mid-exponential phase of autotrophic growth, and determining dynamic label incorporation in metabolites. Consistent with formate incorporation into D. desulfuricans' metabolome, intracellular formate pools were ~90% 13C-enriched after 0.5 h of 13C-label introduction (Supplementary Fig. 3A and Supplementary Data 6). In agreement with the proposed pathway, the concentrations of single-labelled glycine (Supplementary Fig. 3B) and of double-labelled serine (Supplementary Fig. 3C) increased over time. Only single-labelled
isotopomers of phosphoenolpyruvate, aspartate and 3-phosphoglycerate (Supplementary Fig. 4E–G) were observed to increase during the experiment, confirming that pyruvate was not formed by serine deamination, but rather via GR. This is consistent with the single-labelled intracellular acid pool (Supplementary Fig. 3D) also increasing over time, as well as the low level of SDA in the transcriptomics and proteomics data and the high abundance of the GR. Remarkably, glycine and acetate, two metabolites expected to label first based on the proposed pathway (Fig. 3), labelled slower than several central carbon metabolites (Supplementary Fig. 3). One potential explanation could be substrate channelling through the pathway enzymes, where pathway intermediates such as glycine might be immediately passed to downstream enzymes without equilibrating with the bulk cytoplasmic solvent, as recently reviewed38. Direct channelling of glycine, keeping its concentrations low, would be thermodynamically beneficial for the operation of the pathway.

**Discussion**

The pathway used for CO2 fixation by *D. desulfuricans* corresponds to the rGly pathway, which was shown before to serve as a sink for reducing power, but not for autotrophic growth20,22,29–31. The rGly pathway was proposed to support autotrophic growth32, but its operation for this purpose was never confirmed. Recently, it was suggested that the anaerobic phosphate oxidiser *Candidatus Phosphitivorax anaerolimis* uses the rGly pathway for CO2 fixation in an enrichment culture33. However, this was only inferred from metagenomic data, i.e., the absence of genes from any known CO2 fixation pathway and the presence of almost all components of the rGly pathway besides GR (which could be replaced by SDA). Unfortunately, in that study autotrophic growth was not demonstrated for the enrichment culture as a medium with organic carbon compounds (cysteine and rumen fluid) was used and studies with labelled CO2 were not done, neither were any functional omics experiments performed. As *Candidatus Phosphitivorax anaerolimis* has yet to be isolated and no physiological evidence was provided for the activity of the suggested route, it remains unclear if this bacterium indeed grows autotrophically and uses the rGly pathway. Very recently, it was observed that the acetogen *Clostridium drakei* combines the GR variant of the rGly pathway with the reductive acetyl-CoA pathway during autotrophic growth on H2/CO2 in the presence of yeast extract34. As the reductive acetyl-CoA pathway is not complete in *D. desulfuricans*, the rGly pathway is the sole carbon fixation pathway in this bacterium. In summary, our and other recent findings suggest that the rGly pathway is phylogenetically widespread allowing autotrophic growth in Gram-negative bacteria as shown here for *D. desulfuricans*, and in Gram-positive bacteria, such as anaerobic clostridia (in cooperation with the reductive acetyl-CoA pathway)33,34.

Most of the enzymatic components of the rGly pathway are present in many microorganisms29. While GR is quite a unique enzyme, it could be replaced by the serine route. As oxygen-tolerant variants of FDH are known to catalyse CO2 reduction35,36, the rGly pathway could also be active in aerobic microorganisms. To investigate if the enzymes of the different rGly variants are naturally present in other microorganisms, a tblastx search was performed for all relevant genes and operons in all genomes in the NCBI NT database. We found that the complete set of genes of the rGly pathway for the serine route was present in genomes of 203 microorganisms, and for the route via GR in 30 (Supplementary Data 7). Genes to support both routes are present simultaneously in 21 genomes, which included a few members of the *Desulfovibrio* genus but also *Desulfomonile, Desulfosporosinus* and some *Clostridium* strains, including the aforementioned *C. drakei*. Given the diverse types of FDHs in nature, some may be missed by our tblastx analyses; when FDH is excluded from the analyses, up to 557 and 174 genomes harbour all other genes of the serine and glycine variants of the rGly pathway, respectively (Supplementary Data 7) with 157 harbouring both. Many microorganisms listed are known as heterotrophs, or acetogens assumed to use the reductive acetyl-CoA pathway. Our genome analysis suggests that these microorganisms can potentially utilise the rGly pathway for autotrophic growth, but this cannot be concluded only based on genome annotations, and requires further study.

The rGly pathway is one of the most energy-efficient routes for CO2 fixation, as it consumes only 1-2 ATP molecules for the biosynthesis of pyruvate, similarly to the reductive acetyl-CoA pathway that consumes 1 ATP per pyruvate (we note that the latter pathway can additionally generate ATP via bifurcation mechanisms, see also Supplementary Table 3). Acetate can be synthesised without net ATP investment when the rGly pathway proceeds via GR and PTA, and this variant can even generate net ATP if it operates in concert with bifurcating enzymes, as observed for CO2 fixation via the reductive acetyl-CoA pathway in acetogenic bacteria37.

Due to this efficiency and to the thermodynamic feasibility of all its steps, the rGly pathway was suggested as a synthetic route for CO2 fixation and formate assimilation, also in aerobic conditions. Successful modular engineering of the rGly pathway via serine has been demonstrated recently in several biotechnological hosts: *Escherichia coli*, *Capriavidus necator* and yeasts39–42. This shows the feasibility of the operation of this pathway and its potential for biotechnological applications.

To conclude, we demonstrate here that *D. desulfuricans* can grow autotrophically and formatotrophically, as previously hypothesised11. CO2 fixation is performed via the rGly pathway, which was not yet proven to operate in nature to support full autotrophic growth. Further meta-omics and physiological studies are required to elucidate the role of this CO2 fixation pathway and its different variants in the microbial world and its impact on the global biogeochemical carbon cycle.

**Methods**

**Growth conditions and cultivation conditions.** *Desulfovibrio desulfuricans* strain G11 (DSM 7057) was retrieved from our own culture collection at the Laboratory of Microbiology (Wageningen University & Research). Unless otherwise indicated, it was grown in 250 ml glass bottles containing 100 ml anoxic medium. The medium was composed of basal bicarbonate-buffered medium containing the following components (in gram per liter): Na2HPO4·2H2O, 0.53; KH2PO4, 0.41; NH4Cl, 0.3; CaCl2·2H2O, 0.11; MgCl2·6H2O, 0.10; NaCl, 0.3; NaHCO3, 4.0 and Na2S·9H2O, 0.48. Furthermore, acid and alkaline trace element solutions (both 1 ml per liter) and a vitamins solution (0.2 ml per liter) were added. The acid trace element solution contained (in mM): FeCl3, 7.5; H3BO3, 1; ZnCl2, 0.5; CuCl2, 0.1; MnCl2·4H2O, 0.5; CoCl2, 0.5; NiCl2, 0.1 and HCl, 50. The alkaline trace element solution contained (in mM): Na2SeO3, 0.1; Na2WO4, 0.1; Na2MoO4, 0.1; and NaOH, 10. The vitamin solution contained (gram per liter): biotin, 0.02; nicacin, 0.2; pyridoxine, 0.5; riboflavin, 0.1; thiamine, 0.2; cyanocobalamin, 0.1; p-aminobenzoic acid, 0.1 and pantothenic acid, 0.115. The basal medium was supplemented with 20 mM sulphate, and the bottles’ headspace filled with H2/CO2 (1.5 atm, 80:20 v/v) to provide H2 as electron donor and CO2 as carbon source for autotrophic growth. For heterotrophic (acetate/H2/CO2/sulphate) conditions, cultures were additionally provided with 2 mM of acetate. Heterotrophic cultures on lactate were supplemented with 20 mM of lactate as sole energy source with Na2CO3 (1.5 atm, 80:20 v/v) as gas phase. Formotrophic cultures were supplemented with 20 mM of formate as carbon and energy source with Na2CO3 (1.5 atm, 80:20, v/v) as gas phase. Cultures were incubated at 30 °C in the dark.

**Physiological studies and analytical methods.** Growth under autotrophic and heterotrophic conditions was monitored in cultures grown as mentioned above and shaking at 175 rpm for good gas transfer. An inoculum of 10% was used (at 70%...
of exponential phase) from cultures grown in the corresponding autotrophic and heterotrophic conditions. Gas and liquid samples were taken to determine optical density at 600 nm (OD600), and concentrations of sulphide, sulphate, organic acids and H₂. For sulphate and organic acids determination, aliquots of 1 ml of liquid sample were centrifuged for 5 min at 10000 g. Sulphate concentrations were determined using an ion chromatograph Dionex ICS 2100 (Thermo Scientific, Sunnyvale, CA) equipped with a Dionex IonPac AS16 column at 30 °C and a conductivity detector operated at 25 °C. Helium was used as a gas carrier and 5% krypton was used as an internal standard. Growth was monitored by optical density at 600 nm (OD600) using the spectrophotometer Shimadzu UV-1800.

Cultures were routinely checked for purity. For that, 1 ml of fresh culture was centrifuged, the pellet washed twice with sterile TE buffer. The pellet was then resuspended in 100 μl of TE and 1 μl used as a template for PCR. PCR was performed in a 25 μl reaction volume of 0.2 mM dNTPs, 0.02 μM 5′-GoTaq DNA polymerase (Promega, Madison, USA) and 0.2 μM of forward primer 27 T (5′-AGAGTTTGATCCTTGATCTAAG-3′) and 1492r (5′-TACGGYTACCTTGTTACGACTT-3′) targeting the 16S rRNA gene (M = A or C, Y = C or T). A negative control PCR without template DNA was included. The amplification programme consisted of an initial denaturation step at 95 °C for 15 min, 30 cycles of denaturation at 95 °C for 1 min, 40 °C for 52 s for annealing and elongation at 72 °C for 1 min, followed by a final extension step at 72 °C for 7 min. PCR products were purified with the Zymo DNA Clean & Concentrator kit (Zymo Research, Irvine, CA) and sequenced by Sanger sequencing at GATC Biotech (Konstanz, Germany) according to manufacturer’s instructions. Sequences was identified if homology of 25% between the translation products of the nucleotide sequences was identified. To check for the presence of this pathway in combination with other FDH genes, the DHOR pathway was used for the blastx search, and presence of an operon was evaluated with a custom script by inspecting the co-occurrence of the genes within a range of 20,000 bp to each other. For all genes, the coding sequence of the gene was used for the blastx search. A gene was assumed to be present if homology of 25% between the translation products of the nucleotide sequences was identified.

Exploration of the presence of Rgly pathway in other genomes. Identities of the relevant genes of the Rgly pathway in other organisms were performed with blastx against the NCBI NT database (downloaded on 29.04.2020), with an e value of 9.0001 and the option —max_target_seqs 100000. For the operons of GR, GCS and FDH, the whole operon was used for the blastx search, and presence of an operon was evaluated with a custom script by inspecting the co-occurrence of the genes within a range of 20,000 bp to each other. For all genes, the coding sequence of the gene was used for the blastx search. A gene was assumed to be present if homology of 25% between the translation products of the nucleotide sequences was identified.

To check for the presence of this pathway in combination with other FDH genes, the FDH gene clusters were identified in the following way. For all other FDHs (EC numbers 1.17.1.9, 1.17.1.10, 1.17.1.11, 1.17.5.3) the KEGG entries were used as starting point to identify the organisms, which harbour the first-identified instance of the corresponding genes. If genes were listed for more than one species, these were also used for the identification. For EC 1.17.1.9, the genome of Methylotristrus trichosporium OB3b (NZ_ADY6000001) was used. The Refseq annotation of this genome was used. For EC 1.17.1.10, the genome of Moorella thermoacetica ATCC 39073 (CP000322) was used. The genes moth_2312-moth_231467, including intergenic regions (genomic location 243248-2437304) were used for a blastx search. For EC 1.17.1.11, the genome of obligatron strain 9a (CP000326) was used. The genes moth_2307-moth_231237, including intergenic regions (genomic location 167855-1684753) were used for a blastx search. For EC 1.17.1.11, the genome of Moorella thermaacetica ATCC 39073 (CP000322) was used. The genes moth_2312-moth 2314 (CP000326) were used for a blastx search. For EC 1.17.5.3, the genome of Escherichia coli str. K- 12 substr. MG1655 (NC_000913) was used, with the genes b1474-b1476 (genomic location 307424-3085751). For EC 1.17.1.9, the genome of Methylophilus methylotrophus DSM 1367 (BG507960) was used, with the genes b1474-b1476 (genomic location 12515988-12526887). The detection of the whole pathway was based on the presence of a minimum amount of genes. The pathway was considered to be present if any of the four FDHs highlighted in red.
was present, together with the GCS, the DsvG11_3068 formate-THF ligase, the DsvG11_1518 methylthio-THF cyclohydrolase and each of the two optional routes: (A) GR and at least one of phosphate acetyltransferase (DsvG11_0941) or acetyl kinase (DsvG11_0942), or (B) glycine hydroxymethyltransferase (DsvG11_2276) and SDA (DsvG11_1577). The GCS was assumed to be present if three out of four genes could be detected, for the GR five out of seven genes (under the assumption that the two thioareodoxin genes, or any other two genes, could be present at another genomic location), and for the FDH EC 1.17.2.3 four out of six genes had to be detected (under the assumption that the two genes for the molybdenum cofactor, or any other three genes could be present at a different genomic location). The FDH EC 1.17.1.9 was assumed to be present if five out of four genes could be detected (under the assumption that the accessory protein was the same subunit as one of the other FDHs. The FDH EC 1.17.1.11 was assumed to be present if five out of eight genes could be detected (under the assumption that the accessory protein and the two proteins for the production of the molybdenum cofactor, or any other three genes, could be present at a different genomic location). The FDH EC 1.17.1.10 was not evaluated, since only two genes were identified in D. desulfuricans, which seems to have homology with the other FDHs. The FDH EC 1.17.1.11 was assumed to be present if five out of eight genes could be detected (under the assumption that the accessory protein and the two proteins for the production of the molybdenum cofactor, or any other three genes, could be present at a different genomic location). The FDH EC 1.17.3.5 was assumed to be present if all three genes could be detected. The relevant data can be found under https://doi.org/10.26898/mf.figshare.8970689.

Transcriptomic analysis. For transcriptome sequencing, cells were grown under autotrophic, and two heterotrophic conditions (acetate/H2/CO2/sulphate and lactate/H2/CO2/sulphate). The FDH EC 1.17.1.11 was assumed to be present if five out of eight genes could be detected (under the assumption that the two genes for the molybdenum cofactor, or any other three genes, could be present at a different genomic location). The FDH EC 1.17.1.9 was assumed to be present if five out of four genes could be detected (under the assumption that the two genes for the molybdenum cofactor, or any other three genes, could be present at a different genomic location). The FDH EC 1.17.1.10 was not evaluated, since only two genes were identified in D. desulfuricans, which seems to have homology with the other FDHs. The FDH EC 1.17.1.11 was assumed to be present if five out of eight genes could be detected (under the assumption that the accessory protein and the two proteins for the production of the molybdenum cofactor, or any other three genes, could be present at a different genomic location). The FDH EC 1.17.3.5 was assumed to be present if all three genes could be detected. The relevant data can be found under https://doi.org/10.26898/mf.figshare.8970689.

Proteomic analysis. For proteomics analysis, cells were cultivated under autotrophic and heterotrophic (acetate/H2/CO2/sulphate) conditions as described above. Each condition was performed in four biological replicates. Fifty millilitres of culture were harvested at late exponential growth phase (70% and 75% of the maximum OD595 for autotrophic, and heterotrophic (acetate/H2/CO2/sulphate), respectively), centrifuged for 10 min at 10 000 g at 4 °C, the pellet washed with 1 ml PBS and transferred to a protein low-binding eppendorf tube. Samples were then centrifuged (10 min, 1 g, 10 000 g) and the cell pellets stored at −80 °C. Cells were then resuspended in 0.5 ml of 100 mM Tris/HCl pH 7.5 and the suspension was sonicated in four times using a Branson sonifier SFX150 equipped with a 3-mm tip (Branson, Carouge, CH), 20% amplitude in cycles of 30 s pulse and 30 s rest on ice. Unbroken cells and cell debris were removed by centrifugation at 10 000 g for 10 min at 4 °C and the protein concentration in the supernatant was measured using the Bradford assay (Pierce BCA Protein Assay reagent, Thermo Scientific). A total of 10 µg from each of the reduced proteins was performed with 1 µg iodine/mole of 100 mM iodoacetamide in 100 mM Tris (pH 8) at room temperature in dark conditions for 1 h. After a clean-up with 100 µl of ABC buffer, stage-tips were moved to clean 0.5 ml low-binding microcentrifuge tubes, and enzymatic digestion was performed by adding 500 µl of trypsin in 20 µl of ABC (0.5 µg/ml) on top. The acid hydrolysis digestion was incubated at room temperature shaking at 50 rpm for 18 h. To stop the digestion, 10% trifluoroacetic acid in H2O was added to the samples until the pH dropped to 3. The remaining liquid was eluted through into the low-binding microcentrifuge tube followed by elution with 75 µl 0.1% formic acid in water and 5-µl 50% AcN/ 50% 0.1% formic acid. The peptide samples were analysed after injecting (18 µl sample size) on a 1.0 x 32 mm Magic C18AQ 200A 5 µ beads pre-concentration column (Bruker Nederland) at a constant pressure of 270 bar (resulting in a flow of ~7 µl per min). Peptides were then transferred onto a 0.10 x 250 mm Magic C18AQ 200A (3 µm bead size) analytical column with an acetonitrile gradient (flow rate 0.5 µl per min) with a Pressure EASY System. Gradient was applied at 8–33% acetonitrile in water with 23.6 mM formic acid over 50 min, followed by a fast increase in 3 min with a percentage of acetonitrile to 80% (with 20% water and 23.6 mM formic acid in both the acetonitrile and the water) as a column cleaning step. In between pre-concentration and analytical column, a P-777 Upchurch Microcross column was included. Within the waste line of the Microcross an elution pump (3.5 kV) directly to the eluent via a stainless-steel needle. FTMS spectra were measured in full-scan positive mode within m/ɛ 340 and 1400 on a LTQ-Orbitrap XL (Thermo electron, San José, CA) at high resolution (60,000). CID fragmented MSMS scans of the four most abundant 2+ and 3+ charged peaks in the FTMS scan were recorded in a data-dependent mode in the exclude list exclusion (MS/MS threshold 15,000, MSMS exclusion duration 6 s, n 25 ppm). A D. desulfuricans protein sequence database (NCBI accession number 631220) was used together with a database of contaminants, which e.g., contains sequences of common contaminants: BSA (P02769, bovine serum albumin pre- tryptic) (P00760, bovine) tryptic (P00761, porcine), keratin K2AE (P35398, human), keratin K1C9 (P35327, human), keratin K2C1 (P04264, human) and keratin K1C1 (P35327, human). The label-free quantification (LFQ) as well as the match between runs options were enabled. De-amidated peptides were allowed to be used for protein quantification. Other quantification settings were kept default. Filtering and further analysis of the MaxQuant/Andromeda workflow output and the analysis of the abundances of the identified proteins were performed with the Perseus 1.5.5.3 module (available at the MaxQuant suite). Peptides and proteins were accepted for further analysis when they had a false discovery rate (FDR) of <1% and proteins with at least two identified peptides of which at least one was acetylated and one not modified. Reversed hits were deleted from the MaxQuant output. The normal logarithm was taken from protein LFQ MS1 intensities as obtained from MaxQuant. Zero Log LFQ values were replaced by a value of 4.70 (a value slightly lower than the lowest measured value) to make sensible ratio calculations possible. Relative protein quantitation of autotrophic to heterotrophic (acetate/H2/CO2/sulphate) condition was done with Perseus by applying a two sample t-test on the LFQ intensity values of the LFQ intensity values of the two conditions to remove solid particles. The nLC-MSMS system quality was checked with PTXQC74 using the MaxQuant result files. The mass spectrometry proteomics data have been deposited to the PeptideXchange Consortium via the PRIDE75-77,79 partner repository with the dataset identifier PXD003114.

Stable isotopic labelling with 13C-formate. For the proteinogenic amino acids labelling studies, cells were grown in biological triplicates in 30 ml bottles containing 10 ml anoxic medium prepared as described above, with 30 mM sulphate, H2/CO2 in the gas phase, and supplemented with 75 mM sodium formate (59, 61-63). During the stationary phase, the cells were harvested by centrifugation of 9 ml of culture for 5 min at 10 000 g. The cells were washed once with distilled water and then resuspended in 6 M HCl for hydrolysis at 95 °C for 24 h. Then samples were completely dried under an air stream at 95 °C. Hydrolysed samples were resuspended in 1 ml of distilled water, and centrifuged for 5 min at 10 000 g. Supernatants were collected and hydrolysed amino acids were analysed with UPLC-electrospray ionisation (ESI)-MS/MS. Chromatography was performed with a Waters Acquity UPLC system (Waters, Etten-Leur, The Netherlands), using an HSS T3 C18 reversed phase column (100 × 2.1 mm, 1.8 µm; Waters, Etten-Leur, The Netherlands). The mobile phases were 0.1% formic acid in H2O/A (95/5) and 0.1% formic acid in acetonitrile (B). The flow rate was 0.4 ml min−1 and the following gradient was used: 0-1 min 99% A; 1-5 min linear gradient from 99% A to 82% A; 5-6 min linear gradient from 82% A to 1% A; 6-8 min 1% A; 8-8.5 min linear gradient from 1% A to 99% A; 8.5-11 min re-equilibrate with 99% A. Mass spectra were acquired using an ion spray voltage of 3.5 kV and a source temperature of 350 °C. The decay of the isotope ion was monitored in positive ionisation mode, with a scan range of 50.0-300.0 m/z. The spectra were recorded during the first 5 min of the LC gradients. Data analysis was performed using
Xcalibur (Thermo Scientific, Sunnyvale, CA). The identification amino acids was based on retention times and m/z values, which were determined by analysing amino acid standards (Sigma-Aldrich, St. Louis, MI) under the same conditions.

**Dynamic isotopic labelling with 13C-formate.** For the time series metabolomics on 13C isotopic tracer experiments, cells were grown in biological triplicates in 1 L glass bottles containing 500 ml anoxic medium prepared as described above, with 30 mM sulphate and shaking at 175 rpm. The headspace was flushed daily with H2/CO2 (80/20% v/v) to avoid limitation of H2 and CO2. When cells were at the late exponential phase (73% of maximum OD600) 20 ml sample was taken at time 0 and 15 mM of 13C-formate was added. After that, 20 ml samples were taken at 30 s, 1, 2.5, 4, 5, 5, 7, 9 and 24 h. The samples were immediately filtered through 0.22 μm 47 mm Nylor filters (Millipore, HAWP04700) inserted in sterile syringe filtering holders (Swinnex, Millipore, MX004700).

**Postgate, J. R.** Recent advances in the study of the sulfate-reducing bacteria. *Arch. Microbiol.* **138**, 257–262 (1984).

**Mcinerney, M. J., Bryant, M. P. & Pfennig, N.** Anaerobic bacterium that degrades fatty acids in syntrophic association with methanogens. *Arch. Microbiol.* **122**, 129–135 (1979).

**Jansen, K., Thauer, R. K., Widdel, F.** & Fuchs, G. Carbon assimilation pathways in sulfate-reducing bacteria. *FEMS Microbiol. Lett.* **210**, 487–498 (1998).

**Brysch, K., Schneider, C., Fuchs, G. & Widdel, F.** Lithoautotrophic growth of sulfate-reducing bacteria, and description of *Desulfobacter autotrophicus* gen. nov., sp. nov. *Arch. Microbiol.* **148**, 264–274 (1987).

**Fichtel, K., Mathes, F., Könneke, M., Cypenka, H. & Engelen, B.** Isolation of sulfate-reducing bacteria from sediment above the deep-subseafloor aquifer. *Front. Microbiol.* 3, https://doi.org/10.3389/fmicb.2012.00065 (2012).

**Dolling, J., Jiang, B., Henstra, A. M., Stams, A. J. M. & Plugge, C. M.** Syntrophic growth on formate: a new microbial niche in anoxic environments. *Appl. Environ. Microbiol.* **74**, 6126–6131 (2008).

**Stams, A. J. M., Van Dijk, J. B., Dijkema, C. & Plugge, C. M.** Growth of syntrophic propionate-oxidizing bacteria with fumarate in the absence of methanogenic bacteria. *Appl. Environ. Microbiol.* **59**, 1114–1119 (1993).

**Berg, I. A.** Ecological aspects of the distribution of different autotrophic CO2 fixation pathways. *Appl. Environ. Microbiol.* **77**, 1925–1936 (2011).

**Berg, I. A. et al.** Autotrophic carbon fixation in archaea. *Nat. Rev. Microbiol.* **8**, 364–380 (2010).

**Buchanan, B. B. & Arnon, D. I.** A reverse KREBS cycle in photosynthesis. *Proc. Natl. Acad. Sci. U.S.A.* 8, 246–263 (2008).

**Flamholz, A., Noor, E., Bar-Even, A. & Milo, R.** EQuilibrator—the biochemical thermodynamics calculator. *Nucleic Acids Res.* **40**, D770–D775 (2012).

**Durrrre, P. & Andreessen, J. R.** Pathway of carbon dioxide reduction to acetate without a net energy requirement in *Clostridium putrificum*. *FEBS Microbiol. Lett.* **15**, 51–56 (1982).

**Schneeburger, A., Frings, J. & Schink, B.** Net synthesis of acetate from CO2 by *Eubacterium acidaminophilum* through the glyoxate reductase pathway. *FEBS Microbiol. Lett.* **177**, 2–7 (1999).

**Waber, L. J. & Wood, H. G.** Mechanism of acetate synthesis from CO2 by *Clostridium acidiurici*. *J. Bacteriol.* **140**, 468–478 (1979).

**Maia, L. B., Fonseca, L., Moura, I. & Moura, J. J. G.** Reduction of carbon dioxide by a Molybdenum-containing formate dehydrogenase: a kinetic and mechanistic study. *J. Am. Chem. Soc.* **138**, 8834–8846 (2016).

**Zelebuch, L. et al.** Pyruvate formate-lyase enables efficient growth of *Escherichia coli* on acetate and formate. *Biochemistry* **55**, 2423–2426 (2016).

**Kikuchi, G.** The glycol cleavage system: composition, reaction mechanism, and physiological significance. *Plant Mol. Biol.* **20**, 169–187 (1993).

**Bay, E. & Maden, H.** Tetrahydrofolate and tetrahydromethanopterin co-factor: functionally distinct carriers in C1 metabolism. *Biochem. J.* **350**, 609–629 (2000).

**Sweetlove, L. J. & Fernie, A. R.** The role of dynamic enzyme assemblies and substrate channeling in metabolic regulation. *Nat. Commun.* **9**, https://doi.org/10.1038/s41467-018-0433-8 (2018).

**Braakman, R. & Smith, E.** The emergence and early evolution of biological carbon-fixation. *PLoS Comput. Biol.* **8**, https://doi.org/10.1371/journal.pcbi.1002455 (2012).

**Fonknechten, N. et al.** *Clostridium sticklandii*, a specialist in amino acid degradation: revisiting its metabolism through its genome sequence. *BMC Genom.* **11**, 555 (2010).

**Postgate, J. R. & Campbell, I. L.** Classification of *Desulfovibrio* species, the nonsporulating sulfate-reducing bacteria. *Bacteriol. Rev.* **30**, 732–738 (1966).

**Badziong, W., Thauer, R. K. & Zeikus, J. G.** Isolation and characterization of *Desulfovibrio* growing on hydrogen plus sulfate as the sole energy source. *Arch. Microbiol.* **116**, 41–49 (1978).

**Jansen, K., Thauer, R. K., Widdel, F. & Fuchs, G.** Carbon assimilation pathways in sulfate-reducing bacteria. *FEMS Microbiol. Lett.* **210**, 487–498 (1998).

**Mcinerney, M. J., Bryant, M. P. & Pfennig, N.** Anaerobic bacterium that degrades fatty acids in syntrophic association with methanogens. *Arch. Microbiol.* **122**, 129–135 (1979).

**Postgate, J. R.** Recent advances in the study of the sulfate-reducing bacteria. *Arch. Microbiol.* **138**, 257–262 (1984).

**Sorokin, Y. I.** Role of carbon dioxide and acetate in biosynthesis by sulfate-reducing bacteria. *Nature* **310**, 551–552 (1986).

**Ding, C., Alvarez-Cohen, L. & He, J.** Growth of *Dehalococcoides mccartyi* species in an autotrophic consortium producing limited acetate. *Biodegradation* **29**, 487–498 (2018).
31. Fuchs, G. CO2 fixation in aceticogenic bacteria: variations on a theme. FEMS Microbiol. Lett. 39, 181–183 (1981).

32. Cotton, C. A. R. R., Edlich-Muth, C. & Bar-Even, A. Reinforcing carbon fixation: CO2 reduction replacing and supporting carboxylation. Curr. Opin. Biotechnol. 49, 49–56 (2018).

33. Figueroa, I. A. et al. Metagenomics-guided analysis of microbial chemolithoautotrophic phosphate oxidation yields evidence of a seventh form of CO2 fixation in Clostridium. Proc. Natl Acad. Sci. 115, E92–E101, https://doi.org/10.1073/pnas.1715549114 (2017).

34. Song, Y. et al. Functional cooperation of the glycine synthase/reductase and NADH-dependent formate dehydrogenase in Clostridium. Proc. Natl Acad. Sci. USA 117, 7516–7523 (2020).

35. Hartmann, T. & Leimkühler, S. The oxygen-tolerant and NAD+‐dependent formate dehydrogenase from Rhodobacter capsulatus is able to catalyze the reduction of CO2 to formate. FEBS J. 320, 6083–6096 (2019).

36. Yu, X., Niks, D., Mulchandani, A. & Hille, R. Efficient reduction of CO2 by the molybdenum-containing formate dehydrogenase from Cupriavidus necator (Ralstonia eutropha). J. Biol. Chem. 292, 16872–16877 (2017).

37. Schuchmann, K. & Müller, V. Autotrophy at the thermodynamic limit of life: a model for energy dissipation and CO2 assimilation in aceticogenic bacteria. Nat. Rev. Microbiol. 12, 809–821 (2014).

38. Bar-Even, A., Noor, E., Flamholz, A. & Milo, R. Design and analysis of metabolic pathways supporting autotrophic growth for electricity-dependent cultivation of microbes. Biochim. Biophys. Acta - Bioenerg. 1827, 1039–1049 (2013).

39. Claassens, N. J. et al. Replacing the Calvin cycle with the reductive glycine pathway (PPP) metabolites by liquid chromatography-mass spectrometry. FEBS J. 29, 1–13 (2020).

40. Gonzalez De La Cruz, J., Machens, F., Messerschmidt, K. & Bar-Even, A. Core catalysis of the reductive glycine pathway demonstrated in yeast. ACS Synth. Biol. 8, 911–917 (2019).

41. Kim, S. et al. Growth of E. coli on formate and methanol via the reductive glycine pathway. Nat. Chem. Biol. https://doi.org/10.1038/s41589-020-0473-5 (2020).

42. Yishai, O., Bouzon, M., Döring, V. & Bar-Even, A. In vivo assimilation of one-carbon via a synthetic reductive glycine pathway in Escherichia coli. ACS Synth. Biol. 7, 2023–2028 (2018).

43. Clime, J. D. Spectrophotometric determination of hydrogen sulfide in natural waters. Limnol. Oceanogr. 14, 454–458 (1969).

44. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. Basic local alignment search tool. J. Mol. Biol. 215, 403–410 (1990).

45. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing data. EMBnet. J. 17, 10–12 (2011).

46. Schmieder, R. & Edwards, R. Quality control and preprocessing of metagenomic datasets. Bioinformatics 27, 863–864 (2011).

47. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. Nat. Methods 9, 357–359 (2012).

48. Li, H. et al. The Sequence Alignment/Map format and SAMtools. Bioinformatics 25, 2073–2079 (2009).

49. Walker, B. J. et al. Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. PLoS ONE 9, e112963 (2014).

50. Peng, Y., Leung, H. C., Yiu, S. M. & Chin, F. Y. IDBA-UD: a de novo assembler for single-cell and metagenomic sequencing data with highly uneven depth. Bioinformatics 28, 1420–1428 (2012).

51. Hornung, B. V. H., Kuiper, E. J. & Smits, W. K. An in silico survey of Clostridioles difficile extrachromosomal elements. Microb. Genom. 5, https://doi.org/10.1099/mgen.0.000296 (2019).

52. Hyatt, D. et al._prodigal: prokaryotic gene recognition and translation initiation site identification. BMC Bioinform. 11, 119 (2010).

53. Hunter, S. et al. InterPro in 2011: new developments in the family and domain prediction database. Nucleic Acids Res. 40, D306–D312 (2012).

54. Laslett, D. & Canback, B. ARAGORN, a program to detect tRNA genes and tmRNA genes in nucleotide sequences. Nucleic Acids Res. 32 11–16 (2004).

55. Lagergren, K. et al. RNAmmer: consistent and rapid annotation of ribosomal RNA genes. Nucleic Acids Res. 35, 3100–3108 (2007).

56. Bland, C. et al. CRISPR recognition tool (CRT): a tool for automatic detection of clustered regularly interspaced palindromic repeats. BMC Bioinform. 8, 209 (2007).

57. Claudel-Renard, C., Chevalet, C., Faraut, T. & Kahn, D. Enzyme-specific profiles for genome annotation: PRIAM. Nucleic Acids Res. 31, 6633–6639 (2003).

58. Lombard, V., Golconda Ramulu, H., Drula, E., Coutinho, P. M. & Henrissat, B. The carbohydrate-active enzymes database (CAZY) in 2013. Nucleic Acids Res. 42, D490–D495 (2014).

59. Yin, Y. et al. dbCAN: a web resource for automated carbohydrate-active enzyme annotation. Nucleic Acids Res. 40, W445–W451 (2012).
