



# Carbon-negative production of acetone and isopropanol by gas fermentation at industrial pilot scale

Fungmin Eric Liew<sup>1,9</sup>, Robert Nogle<sup>1,9</sup>, Tanus Abdalla<sup>1,9</sup>, Blake J. Rasor<sup>2,3,4</sup>, Christina Canter<sup>1</sup>, Rasmus O. Jensen<sup>1</sup>, Lan Wang<sup>1</sup>, Jonathan Strutz<sup>1</sup>, Payal Chirania<sup>5,6</sup>, Sashini De Tissera<sup>1</sup>, Alexander P. Mueller<sup>1</sup>, Zhenhua Ruan<sup>1</sup>, Allan Gao<sup>1</sup>, Loan Tran<sup>1</sup>, Nancy L. Engle<sup>5</sup>, Jason C. Bromley<sup>1</sup>, James Daniell<sup>1</sup>, Robert Conrado<sup>1</sup>, Timothy J. Tschaplinski<sup>5</sup>, Richard J. Giannone<sup>5</sup>, Robert L. Hettich<sup>5</sup>, Ashty S. Karim<sup>2,3,4</sup>, Séan D. Simpson<sup>1</sup>, Steven D. Brown<sup>1</sup>, Ching Leang<sup>1</sup>✉, Michael C. Jewett<sup>2,3,4,7,8</sup>✉ and Michael Köpke<sup>1</sup>✉

**Many industrial chemicals that are produced from fossil resources could be manufactured more sustainably through fermentation. Here we describe the development of a carbon-negative fermentation route to producing the industrially important chemicals acetone and isopropanol from abundant, low-cost waste gas feedstocks, such as industrial emissions and syngas. Using a combinatorial pathway library approach, we first mined a historical industrial strain collection for superior enzymes that we used to engineer the autotrophic acetogen *Clostridium autoethanogenum*. Next, we used omics analysis, kinetic modeling and cell-free prototyping to optimize flux. Finally, we scaled-up our optimized strains for continuous production at rates of up to ~3 g/L/h and ~90% selectivity. Life cycle analysis confirmed a negative carbon footprint for the products. Unlike traditional production processes, which result in release of greenhouse gases, our process fixes carbon. These results show that engineered acetogens enable sustainable, high-efficiency, high-selectivity chemicals production. We expect that our approach can be readily adapted to a wide range of commodity chemicals.**

The accelerating climate crisis has intensified demand for environmentally sustainable, carbon-negative manufacturing of chemicals and materials. Most commodity chemicals are derived exclusively from fresh fossil resources, such as oil, natural gas and coal. Acetone and isopropanol (IPA) are two such examples, with a combined global market of more than US\$10 billion. Both molecules are industrial solvents as well as platform chemicals for the production of materials such as acrylic glass (polymethyl methacrylate) and polypropylene. Additionally, acetone has potential in the fuel sector to reduce emissions and enhance the performance of existing fuels<sup>1,2</sup> and as a precursor for green diesel or jet fuel substitutes<sup>3,4</sup>. IPA is also a disinfectant; its use in World Health Organization-recommended sanitizer formulations against severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2)<sup>5</sup> led to supply shortages and surging prices during the Coronavirus Disease 2019 (COVID-19) pandemic. Acetone is mostly produced as a byproduct of phenol production using the cumene process. Manufacture of both acetone and IPA is dependent on propene cracking or reforming, which are energy-intensive processes that create hazardous waste and greenhouse gas (GHG) emissions<sup>6</sup>. No viable green chemistry alternatives are available.

Biomufacturing by fermentation enables sustainable, purposeful production of desired chemicals using ‘above-ground’ carbon sources. The acetone-butanol-ethanol (ABE) process developed by Chaim Weizmann in the first half of the 20th century was one of the first industrial fermentation processes for chemicals production<sup>7</sup>. In this process, acetone, butanol and ethanol are produced in a mixed fermentation (typically in a ratio of 3:6:1) by the solventogenic *Clostridium* species *Clostridium acetobutylicum*, *Clostridium beijerinckii*, *Clostridium saccharobutylicum* or *Clostridium saccharoperbutylacetonicum* from sugar or starch feedstocks<sup>7</sup>. A few ABE strains that contain a primary–secondary alcohol dehydrogenase (sAdh) also produce low levels of IPA<sup>8</sup>. At its peak, global ABE production reached more than 500 metric tons per annum, but eventually the ABE process was phased out owing to limited selectivity to one desired product and high substrate costs and was continued in only a few countries, such as South Africa and Taiwan, until the 1980s<sup>7</sup>. To develop alternative fermentation production systems, many studies have sought to decouple acetone and IPA production from butanol synthesis and transfer the respective pathways to heterologous hosts or to develop novel synthetic pathways<sup>9,10</sup>. Substantial strain<sup>11,12</sup> and process development<sup>13</sup> has been carried out in *Escherichia coli*, and, although improvements

<sup>1</sup>LanzaTech Inc., Skokie, IL, USA. <sup>2</sup>Department of Chemical and Biological Engineering, Northwestern University, Evanston, IL, USA. <sup>3</sup>Chemistry of Life Processes Institute, Northwestern University, Evanston, IL, USA. <sup>4</sup>Center for Synthetic Biology, Northwestern University, Evanston, IL, USA. <sup>5</sup>Biosciences Division, Oak Ridge National Laboratory, Oak Ridge, TN, USA. <sup>6</sup>Graduate School of Genome Science and Technology, University of Tennessee, Knoxville, TN, USA. <sup>7</sup>Robert H. Lurie Comprehensive Cancer Center, Northwestern University, Chicago, IL, USA. <sup>8</sup>Simpson Querrey Institute, Northwestern University, Chicago, IL, USA. <sup>9</sup>These authors contributed equally: Fungmin Eric Liew, Robert Nogle, Tanus Abdalla. ✉e-mail: [ching.leang@lanzatech.com](mailto:ching.leang@lanzatech.com); [m-jewett@northwestern.edu](mailto:m-jewett@northwestern.edu); [michael.koepke@lanzatech.com](mailto:michael.koepke@lanzatech.com)

in selectivity have been made, low yields remain a major hurdle for commercialization. In the case of acetone and IPA, the maximum theoretical yield for fermentative production from sugars is capped at 50% owing to formation of CO<sub>2</sub> during the breakdown of sugars and acetone synthesis (pyruvate and acetoacetate decarboxylase reactions)<sup>14</sup>.

Gas fermentation using autotrophic organisms is an alternative to sugar fermentation. In this approach, instead of breaking down C5 or C6 sugar units, microbes build up products from C1 units, such as carbon oxides (CO or CO<sub>2</sub>), which are found, for example, in waste gases from heavy industry (for example, steel mills, processing plants or refineries) or in syngas generated from any biomass resource (for example, unsorted and non-recyclable municipal solid waste, agricultural waste or organic industrial waste)<sup>15,16</sup>. Capturing and recycling these waste carbon streams before they enter the atmosphere or environment offers routes to carbon-negative manufacturing and a circular economy. To our knowledge, there is no autotroph that produces acetone or IPA natively, but low-level production of acetone and IPA (up to a rate of 3.8 mg/L/h in autotrophic mode) has been demonstrated in several autotrophic hosts through genetic engineering<sup>14,17–24</sup>.

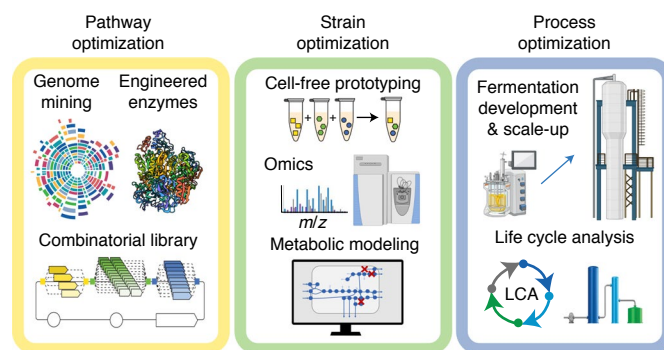
Among the range of autotrophic organisms, anaerobic acetogens are particularly attractive as they do not require light as their source of energy (providing an easier route for scale-up) and make use of the most efficient of the known CO<sub>2</sub> fixation pathways (the reductive acetyl-CoA/Wood–Ljungdahl pathway)<sup>25–27</sup>. Since the first demonstration of genetic engineering of an acetogen in 2010 (ref. 28), considerable advances in the genetic toolbox of acetogens have been made<sup>29,30</sup>. Genetic engineering of acetogens has enabled small-scale production of more than 50 molecules from C1 gases, including industrially important alcohols (butanol, hexanol, IPA, etc.), diols (monoethylene glycol, propanediol, butanediol, etc.), acids (lactic acid, succinic acid, etc.), ketones (acetone and methyl ethyl ketone) and dienes (butylene, isoprene, etc.)<sup>15,17–19,28–30</sup>. However, efforts to develop efficient production strains have been hampered by the lack of high-throughput strain engineering workflows and a pathway to industrial scale-up<sup>29</sup>. The only commodity chemical that has been produced by gas fermentation at industrial scale, to our knowledge, is ethanol, which was achieved by our lab using a naturally selected, non-engineered strain of *Clostridium autoethanogenum* that synthesizes ethanol as a native product<sup>16</sup>. In previous work, we optimized the process for high productivities at g/L/h levels and selectivity of more than 90% and have implemented it at two industrial plants, with an annual production of more than 90,000 metric tons<sup>15,29,31</sup>.

Building on these advances, here we describe an efficient, carbon-negative fermentation process for production of acetone and IPA by rewiring *C. autoethanogenum* from an ethanol producer to a producer of acetone or IPA using enzymes from a historical industrial strain collection, and we demonstrate scale-up of the process to industrial pilot scale.

## Results

Our strategy to achieve efficient acetone and IPA production relied on a three-pronged approach of pathway, strain and process optimization (Fig. 1). Specifically, we integrated high-throughput strain engineering workflows, omics analysis, cell-free systems, kinetic modeling, fermentation scale-up and life-cycle analysis (LCA). We first identified optimal sets of heterologous pathway enzymes to carry out the desired molecular transformations. Then, we optimized our strains for enhanced flux to product. Finally, we carried out process optimization, scale-up and LCA.

**Pathway optimization.** Despite the long history of ABE fermentation<sup>7</sup>, the available pool of enzyme variants for heterologous acetone biosynthesis, specifically thiolase (ThlA), CoA transferase (CtfAB) and acetoacetate decarboxylase (Adc), is limited (Fig. 2a).

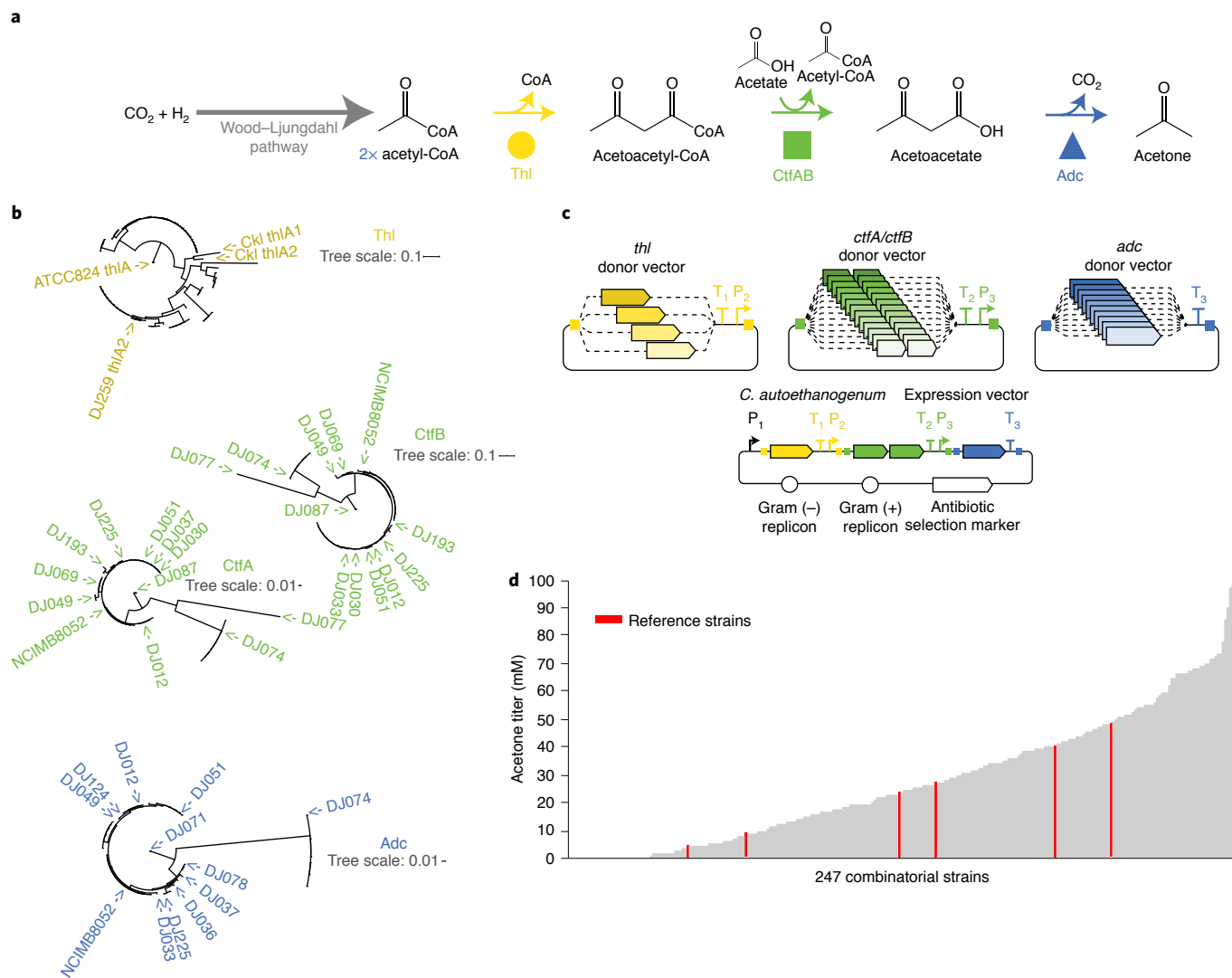


**Fig. 1 | Overview of our three-pronged approach for pathway, strain and process optimization.** Overview of applied tools and strategies to advance acetone and IPA production from waste gases from a proof of concept to industrial level.

In all previous studies<sup>9–14,17–24</sup>, only genes derived from two reference strains—natural isolates *C. acetobutylicum* ATCC824 and *C. beijerinckii* NCIMB8052—have been explored. To expand the sequence space of available enzymes, we mined a recently sequenced collection of 272 ABE strains (assembled by David T. Jones of Otago University and herein referred to as the ‘DJ collection’), which comprises a myriad of optimized industrial strains from four decades of biotechnological development, including production strains from the last industrial ABE operations in South Africa and Taiwan<sup>7,32,33</sup>. We identified a total of 41 unique and not previously described acetone biosynthesis enzymes in the DJ collection (ThlA: 6, CtfA: 11; CtfB: 13 and Adc: 11; Supplementary Table 1). These enzymes had an identity of 72.4–99.8% to the enzymes of reference strains (Supplementary Figs. 1–4). Several alcohol dehydrogenases, but no sAdh enzymes, were found to be present in the DJ collection.

From the identified 41 new enzymes, we selected a representative subset of 30 enzymes (ThlA: 4, CtfA: 6; CtfB: 10 and Adc: 10; Fig. 2b and Supplementary Table 1) to create a combinatorial plasmid library with the aim of optimizing acetone production. This included reference enzymes from ATCC824 and NCIMB8052 as well as two thiolases from *Clostridium kluyveri*, which we had used successfully for butanol production in the past<sup>34</sup>. Because of this previous thiolase prototyping effort<sup>34</sup>, we focused mainly on CtfAB and Adc enzyme variants. Enzyme variants included in the library were combined with three non-repetitive promoters of varying strengths (high, medium and low)<sup>35</sup> to assess acetone production in terms of enzyme abundance via increased expression. To construct the various combinations, we developed a hierarchical approach using three donor vectors and one expression vector based on the modular and universal pMTL80000 vector system<sup>36</sup>, facilitated through Golden Gate assembly and counterselection (Fig. 2c and Supplementary Text). Our approach ensured that each enzyme was driven by a unique promoter forming separate transcription units. Functional transcription units were formed only in the final expression vector to avoid any bias or unwanted recombination in the intermediate host *Escherichia coli*. After assembly in *E. coli*, sequencing of the resulting combinatorial plasmids showed good diversity, confirming that all promoters and genes were present and relatively evenly distributed (within  $\pm 10\%$ ) (Supplementary Fig. 5). In addition to the original sequences, there were point mutations identified for some variants (Supplementary Fig. 5). Unless this resulted in a frameshift, we still carried these mutant variants forward.

Because *C. autoethanogenum* possesses a native primary–secondary alcohol dehydrogenase (sAdh, encoded by CAETHG\_0553) known to reduce acetone to IPA<sup>37,38</sup>, we generated a *C. autoethanogenum* sAdh knockout (KO) strain ( $\Delta 0553$ ) to enable screening of the library in isolation of this reducing



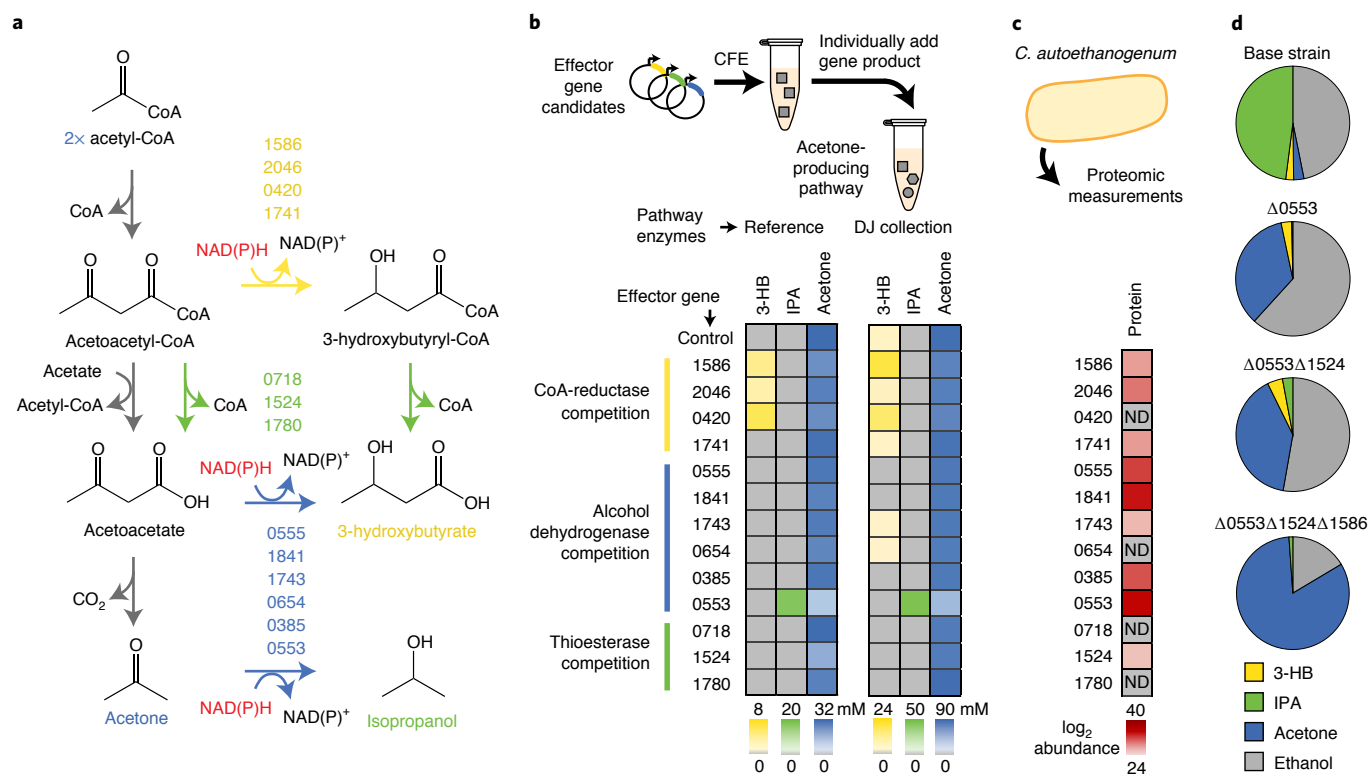
**Fig. 2 | Pathway optimization.** **a**, Acetone and IPA biosynthesis pathway from gases. **b**, Sequence mining of the DJ collection led to identification of a large diversity of acetone biosynthesis enzyme sequences, including 65 unique and not previously described enzymes. Scale is in substitutions per position; expanded trees are found in Supplementary Figs. 1–4. **c**, Combinatorial library assembly strategy to refactor selected acetone biosynthesis genes. **d**, Acetone end-point titers observed in screening 247 strains with unique acetone pathway designs; designs that use genes from wild-type reference strains are highlighted in red.

equivalent consuming reaction. After transformation into  $\Delta 0553$ , a total of 247 strains harboring distinct acetone biosynthesis pathway designs (including all possible promoter–gene combinations for the reference enzymes) were obtained and screened (Supplementary Table 2). We initially investigated screening in a plate-based format, but because acetone is too volatile and we wanted to perform the screening under representative conditions using gases as sole energy and carbon source, we performed the screen in a closed system using pressurized bottles. To our knowledge, this represents the largest library built and screened in an autotroph by at least an order of magnitude.

All of the growth-tested gene variants produced acetone with a wide range of end-point titers (up to 100 mM) observed (Fig. 2d and Supplementary Table 2). Varying promoters alone led to an 11-fold improvement in acetone production, which could be improved another two-fold when further combined with enzyme variants from the DJ collection. The best combination was composed of thiolase from ATCC824, CtfA Var 8 (DJ033; 72.22%

identity to reference), CtfB Var 5 (DJ033; 96.83% identity to reference) and Adc Var 1 (found across a range of DJ strains; 75.72% identity to reference). Notably, the second-best combination included variants that had several point mutations to the original sequence in CtfAB (Supplementary Table 2). The top five designs were moved forward to continuous fermentation testing and genome integration. IPA production was not observed in any of the strains (demonstrating the efficacy of  $\Delta 0553$  for avoiding acetone conversion), but we were surprised to measure production of 3-hydroxybutyrate (3-HB), which we traced back to native enzymes interacting with intermediates of the introduced acetone pathway.

After optimizing the acetone pathway, we next screened a library of *C. autoethanogenum* sAdh enzymes<sup>37</sup> in  $\Delta 0553$  for conversion of acetone to IPA. Wild-type sAdh and engineered variants S199A and S199R that showed near-complete conversion (>97%) of 20 g L<sup>-1</sup> of fed acetone during growth on gas (Supplementary Fig. 6) were moved forward for testing in continuous fermentation and genome integration.



**Fig. 3 | Strain optimization.** **a**, In silico homology search identified 13 enzymes candidates native to *C. autoethanogenum* that were suspected to lead to unwanted 3-HB formation by interacting with the heterologous acetone pathway. **b**, Cell-free prototyping of the identified *C. autoethanogenum* effector candidate enzymes on acetone production using our iPROBE system. Effector gene candidates were produced via CFE and then spiked into *E. coli* cell extracts individually enriched with acetone biosynthesis enzymes. Concentrations of 3-HB, acetone and IPA were measured. Results indicate that *C. autoethanogenum* enzymes 1586, 0420, 0553 and 1524 cause the greatest decrease in acetone titers. Two sets of acetone enzymes were used, one with reference enzymes resulting in moderate acetone production and one with enzymes from the DJ collection resulting in high acetone productivity (mean ± 1 s.d. for  $n=3$  for each set). **c**, Expression levels of identified candidate enzymes (ND, not detected) in *C. autoethanogenum* ( $n=3$ ). Only three of the four identified candidate enzymes are expressed (1586, 0553 and 1524) but not 0420. **d**, Improvements in selectivity of acetone over byproducts 3-HB, IPA and ethanol by serial parallel KO of selected effector candidate enzymes 1586, 0553 and 1524 in *C. autoethanogenum* ( $n=3$ ). A complete selectivity breakdown is presented in Supplementary Fig. 8. Selectivity is the proportion of energy in the consumed feed gas that goes toward the product of interest on an electron basis.

**Strain optimization.** With pathway enzymes identified, we next optimized our production strains for increased titers. Specifically, we used a genome-scale model<sup>16,39,40</sup> and evolutionary algorithm<sup>41,42</sup> to predict gene KOs that would increase flux to acetone and eliminate unwanted byproducts, particularly 2,3-butanediol (2,3-BDO) and 3-HB. These byproducts have high boiling points and would accumulate during continuous fermentation with a recycle loop, adding separation costs. The model identified eight gene KO targets predicted to increase acetone production flux (Supplementary Table 3), including the acetolactate decarboxylase reaction (*budA* gene) in the well-characterized 2,3-BDO pathway<sup>38</sup>. However, the model did not identify genes responsible for 3-HB production in the existing network.

Therefore, we performed an in silico homology search based on enzymes described in the literature that perform similar reactions (Fig. 3a). For example, a 3-HB dehydrogenase (Bdh) has been described in *Cupriavidus*<sup>43,44</sup> capable of reducing acetoacetate to 3-HB. Such activity would be directly competing with Adc, and combinatorial strain designs with low Adc expression were among the highest 3-HB titers (Supplementary Table 2). Several putative Bdh isozymes as well as other promiscuous alcohol dehydrogenases, including sAdh or butanediol dehydrogenases, are present in the *C. autoethanogenum* genome<sup>38,45</sup>. Likewise, a reduction to 3-HB might also manifest at the CoA level, whereby acetoacetyl-CoA is

converted to 3-HB-CoA as described in the butanol pathway of *C. acetobutylicum* (Hbd) or PHB pathway of *Cupriavidus*<sup>46,47</sup>. This reaction is also part of the fatty acid biosynthesis pathway. Although respective FadBJ and Fab enzymes have been shown to be primarily active on acyl carrier protein (ACP)-bound substrates, they have been shown to accept CoA-bound substrates too<sup>48</sup>. The CoA moiety, in turn, can be cleaved off by a thioesterase; several enzymes of this functional class are present in the *C. autoethanogenum* genome. Thioesterases might also directly act on acetyl-CoA or acetoacetyl-CoA, limiting flux into the pathway or competing with the CoA transferase, but do not provide energy conservation. A total of 13 candidate genes were identified in the *C. autoethanogenum* genome that alone or in combination might be responsible for 3-HB formation.

Combinatorial KO of such a large matrix of gene targets was infeasible with current in vivo workflows in *C. autoethanogenum*. Even iterative KO of a down-selected set would take several months and risk missing the correct genes. To address this challenge, we adapted the in vitro Prototyping and Rapid Optimization of Biosynthetic Enzymes (iPROBE) approach that we recently developed to prototype biosynthetic pathways<sup>34</sup> to now prototype KO candidate targets. The concept was to accelerate host strain engineering using cell-free systems by assessing the effect of enriching a single target effector protein on acetone biosynthesis (that is, does a KO candidate decrease acetone yields).

First, we used cell-free gene expression (CFE) to create an array of cell extracts individually enriched with acetone biosynthesis enzymes and effector candidate enzymes.  $^{14}\text{C}$ -leucine incorporation confirmed sufficient soluble enzyme expression (Supplementary Fig. 7). Next, we established in vitro acetone production from glucose using native *E. coli* catabolism coupled with two different combinations of ThlA, CtfAB and Adc at final concentrations of 0.50  $\mu\text{M}$  each. The first combination using reference genes produced around 30 mM acetone, whereas the second combination that incorporates sequences from the DJ collection produced nearly three times more acetone (Supplementary Table 4). The gene KO candidate enzymes were then added to cell-free acetone biosynthesis reactions at 1  $\mu\text{M}$ , mimicking overexpression. A substantial decrease in acetone titer with the addition of enzyme 0553 identified this genomic KO as a priority for deletion in the production strain. Enzymes 1586, 0420 and 1524 were also classified as detrimental to acetone biosynthesis (Fig. 3b). Similar trends were observed for both acetone biosynthesis combinations, although enzyme 1524 did not decrease acetone titer for the second combination, which might be due to more favorable kinetics for acetone production with these enzymes.

3-HB effector candidates identified by iPROBE were then successively knocked-out in cells to eliminate byproducts and increase acetone production. Notably, our proteomics of *C. autoethanogenum* combinatorial strains, as well as previous RNA sequencing and proteomics studies<sup>16,39</sup>, indicated that not all of these gene KO candidates are significantly expressed (Fig. 3c and Supplementary Tables 5 and 6). In particular, 0420 showed negligible expression; therefore, genes 0553, 1586 and 1524 were prioritized for sequential KO. Although  $\Delta 0553\Delta 1524$  only showed a 5% reduction in 3-HB, the triple KO strain  $\Delta 0553\Delta 1524\Delta 1586$  no longer produced measurable levels of 3-HB (Fig. 3d and Supplementary Fig. 8). No obvious negative phenotypic effects or reduction in growth rate were observed from these KOs.

Next, we targeted the 2,3-BDO pathway, which our genome-scale model predicted would be accomplished by elimination of the *budA* gene (CAETHG\_2932). In parallel, we chromosomally integrated the optimized acetone pathway operons identified in this study to increase stability and avoid the requirement for antibiotics. In the resulting strain  $\Delta 0553\Delta 1524\Delta 1586\Delta 2932::P_{fer}\text{-thlA-ctfAB-adc}$ , acetone production increased 27-fold from less than 2 mol% to more than 50 mol%, with acetate and ethanol as the only other products (Supplementary Fig. 8).

Because acetate, which is a precursor for ethanol production in acetogens<sup>49,50</sup>, is required as a co-substrate in the acetone pathway as part of the CoA transferase reaction (Fig. 2a), it cannot be eliminated completely. To further optimize flux to acetone, select strains were analyzed by omics measurements and kinetic modeling for bottlenecks in the pathways and tuning of enzyme levels tested in our iPROBE system.

Proteomic measurements were conducted on ten strains from our combinatorial library. These analyses included four strains across the production range (ranks 1, 14, 26 and 150 of 247) and six strains that had the same set of reference genes but driven by different promoters with varied strengths (Fig. 4a and Supplementary Table 6). A total of 1,914 distinct proteins were quantifiable, with acetone pathway-specific enzymes all among the top 3.2% by median abundance. Statistical analysis showed that acetone production correlated well with the abundance of acetone pathway enzymes, with the lowest abundance found in the lowest-producing strains. Consistently, the two subunits of CtfAB were found to have relatively lower expression levels compared to ThlA and Adc, even if driven by the strongest promoter (Fig. 4a).

We then generated an ensemble of acetone-producing *C. autoethanogenum* models (Supplementary Fig. 9) with different kinetic parameters following a recently described approach<sup>51</sup>. The ensemble was pruned to include only models whose predictions matched

experimental data. By simulating changes to the expression level of each enzyme associated with central metabolism and the acetone pathway, the model predicted that overexpressing CtfAB would increase flux to acetone (Fig. 4b).

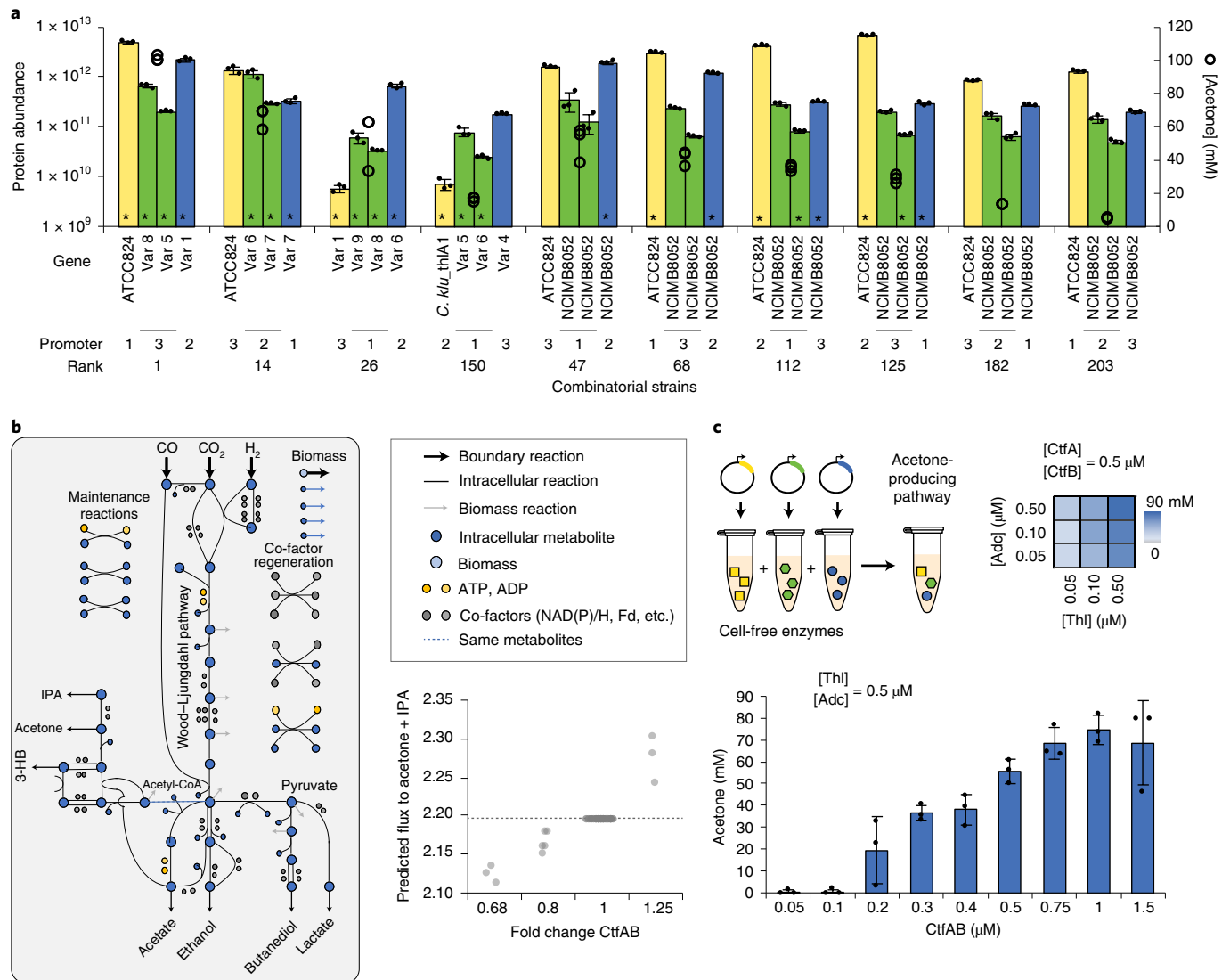
Using our iPROBE system, we next tested this hypothesis by varying levels of the different acetone pathway enzymes and confirmed that increasing CtfAB abundance improved acetone production by the largest amount. ThlA, CtfAB and Adc were expressed by CFE and added to in vitro biosynthesis reactions at a range of concentrations<sup>34</sup>. First, Thl and Adc were co-titrated at low, medium and high concentrations (0.05, 0.1 and 0.5  $\mu\text{M}$ , respectively) with CtfAB constant at 0.5  $\mu\text{M}$ , increasing in vitro acetone titer by three-fold (Fig. 4c). Next, CtfAB was titrated at concentrations ranging from 0.05 to 1.50  $\mu\text{M}$  with Thl and Adc constant at 0.5  $\mu\text{M}$ , increasing acetone titer from the limit of detection to  $74.58 \pm 6.69$  mM (Fig. 4c).

Taken together, our omics measurements, kinetic modeling and cell-free prototyping all suggested CtfAB as the limiting step for acetone production. We, therefore, added a second copy of CtfAB by expressing the CtfAB variant and promoter combination that registered the highest proteomic abundance on a plasmid in strains  $\Delta 0553\Delta 1524\Delta 1586\Delta 2932::P_{fer}\text{-thlA-ctfAB-adc}(-sAdh)$ , which led to an increase of more than 40% in acetone/IPA selectivity (Supplementary Fig. 10).

**Process optimization.** Starting with the best plasmid-based strains from the combinatorial library screen, we next established a continuous fermentation process for acetone in a benchtop continuous stirred-tank reactor (CSTR). The fermentation process showed an average 12 d of stable steady state production (Supplementary Fig. 11) before declining. Ethanol and acetone were the main fermentation products along with acetate, 2,3-BDO and 3-HB as minor products (Supplementary Fig. 8). Whole-genome sequencing showed that no mutations were found related to a decline in performance. Strain  $\Delta 0553\Delta 1524\Delta 1586$  with the same plasmid performed similarly without measurable 3-HB formation (Supplementary Fig. 8). Using a genomically integrated strain  $\Delta 0553\Delta 1524\Delta 1586\Delta 2932::P_{fer}\text{-thlA-ctfAB-adc}$  led to a substantial improvement in acetone selectivity and culture stability, extending steady state production to ~3 weeks before runs were stopped (Fig. 5a and Supplementary Fig. 11). Based on concentrations in the broth, acetone productivity was ~2.5 g/L/h with ~80% gas utilization (Fig. 5d and Supplementary Fig. 11). Actual productivities are likely higher considering that acetone is highly volatile, and we observed acetone in a condenser added to the off-gas line.

Based on the established acetone process, we investigated whether the performance would translate to IPA production. *sAdh* wild-type and variants that had shown best conversion rates of fed acetone were tested in context of the developed acetone strains. Of the variants, S199R exhibited the highest titer and stability for IPA production. Chromosomal integrated strain  $\Delta 0553\Delta 1524\Delta 1586\Delta 2932::P_{fer}\text{-thlA-ctfAB-adc-sAdh}$  performed similarly to our acetone strain after small adjustments in fermentation regimen. As IPA is less volatile, we observed higher productivities of up to ~3 g/L/h with ~85% gas utilization (Fig. 5d and Supplementary Fig. 11) and also slightly increased selectivities (Fig. 5b). For both acetone and IPA strains, expressing an extra copy of CtfAB further increased performance in the CSTR, achieving up to ~90% selectivity (Fig. 5a,b).

Given the successful engineering of *C. autoethanogenum* to heterologously produce acetone and IPA, both processes were scaled up by a factor of 60 $\times$  and from the bench to a field pilot. This included transitioning from a 2-L benchtop CSTR to a 120-L pilot plant using a loop reactor, in which important parameters, such as pressure, can be assessed (Fig. 5c)<sup>15,29</sup>. Several campaigns were carried out with strains  $\Delta 0553\Delta 1524\Delta 1586\Delta 2932::P_{fer}\text{-thlA-ctfAB-adc}(-sAdh)$  (which did not have the extra copy of CtfAB) under similar fermentation conditions and gas composition. Strain performance



**Fig. 4 | Flux optimization.** **a**, Protein abundance (area under the curve) for acetone biosynthesis enzymes (ThIA, yellow; CtfAB, green; Adc, blue) of ten selected combinatorial library strains: four strains across the production range (ranks 1, 14, 26 and 150 of 247) and six strains with fixed reference genes but different promoter combinations (mean  $\pm$  1 s.d. for  $n = 3$ ; all data points shown). Asterisks represent significance abundance differences ( $P \leq 0.05$ ) compared to lowest producer strain (rank 203) as assessed by a two-sided  $t$ -test corrected for multiple comparisons by the Benjamini-Hochberg method. Exact  $P$  values for each comparison are listed in Supplementary Table 6. Proteomics samples were taken from cells harvested during mid-log growth, and overlaid acetone production values are end-point values ( $n = 3$  with all data shown). **b**, Kinetic ensemble modeling ( $n = 50$  trained models) predicted that total flux to acetone and IPA would increase as CtfAB expression increases. No models were able to reach steady state for the 1.5-fold perturbation. An expanded representation of the core metabolic network can be found in Supplementary Fig. 9. **c**, Cell-free prototyping enabled precise modulation of ThI and Adc and of CtfAB to demonstrate the strong dependence of acetone titer on the concentration of CtfAB (mean  $\pm$  1 s.d. for  $n = 3$ ).

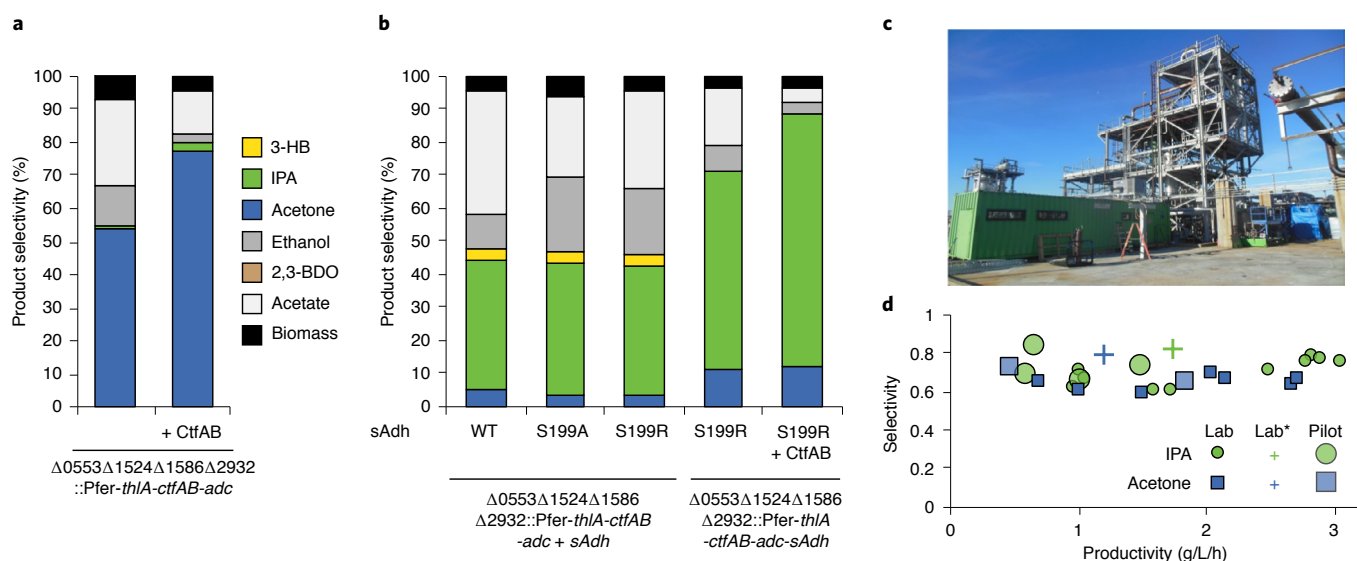
in continuous culture, with respect to product selectivity and productivity as well as stability in the pilot reactor, correlated well with performance in the benchtop CSTR-based fermentations across a range of production rates (Fig. 5d). Productivity varied owing to several factors, most notably strain genotype and bacterial dilution rates due to different reactor configurations. Small variations are expected given the use of lower-grade bulk chemicals and process water and elevated pressure (1.4 bar) at pilot scale.

We used LCA to compare the GHG emissions of acetone and IPA produced using the scaled-up gas fermentation process with current fresh fossil-based production via propene and the cumene process (Fig. 6a). Chemical and energy inputs and product yield were determined using pilot-scale data coupled with design data from industrial-scale ethanol production<sup>15,29</sup>, which uses steel mill off-gas feedstock with utilities (for example, electricity and steam) (Fig. 6b

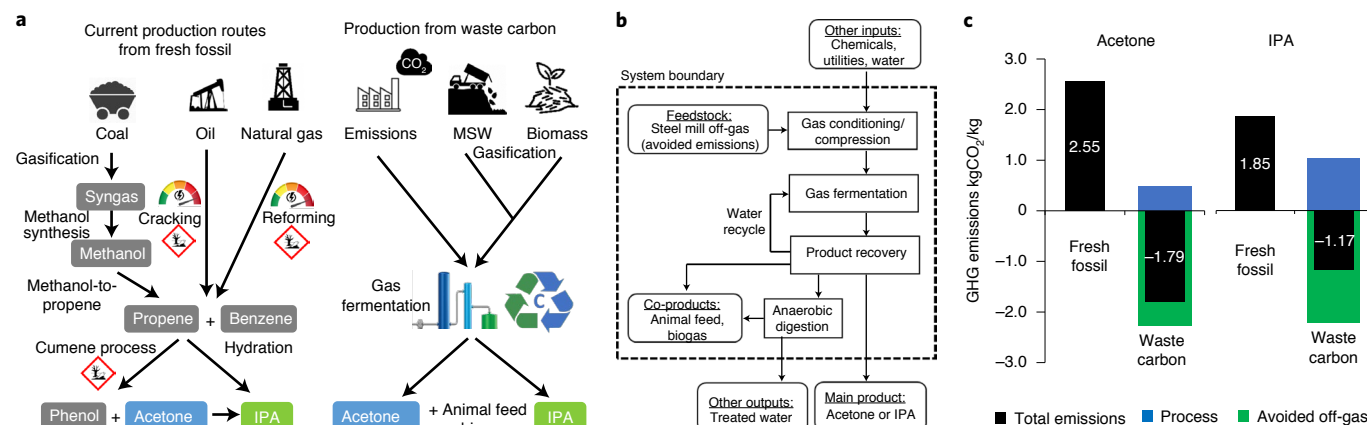
and Supplementary Tables 9 and 10). The GHG emissions of acetone and IPA produced by the gas fermentation process (Fig. 6c) are  $-1.78 \text{ kgCO}_2\text{e/kg}$  acetone and  $-1.17 \text{ kgCO}_2\text{e/kg}$  IPA, respectively. The negative value is due to avoided off-gas emissions combined with low process emissions (Fig. 6c), effectively locking carbon into the product. In contrast, traditional petrochemical production routes result in substantial GHG emissions of  $2.55 \text{ kgCO}_2\text{e/kg}$  for acetone and  $1.85 \text{ kgCO}_2\text{e/kg}$  for IPA.

## Discussion

Despite recent progress in fermentative processes, production of commodity chemicals by sugar fermentation remains difficult to scale<sup>52,53</sup>. Only a few examples have reached industrial level, including 1,3-propanediol, 1,4-butanediol, isobutanol, farnesene, lactic acid and succinic acid—all of which required substantial financial



**Fig. 5 | Process optimization and scale-up.** **a**, Continuous acetone production in lab-scale CSTRs with an integrated strain. Expression of an extra copy of CtfAB led to a considerable improvement in selectivity to acetone (mean  $\pm$  1 s.d. for  $n=3$ ). **b**, Continuous IPA production in lab-scale CSTRs. IPA production was optimized through a combination of different sADH variants and expressing an extra copy of CtfAB and integration of the pathway (mean  $\pm$  1 s.d. for  $n=3$ ). **c**, 120-L scalable pilot reactor at LanzaTech’s Freedom Pines Biorefinery site. **d**, Comparison of key performance metrics selectivity and productivity across a spectrum of individual runs performed in lab-scale CSTRs (small circles, small squares and crosses) and pilot scale (large circles and large squares) for acetone (blue) and IPA (green) strains. Crosses indicate strains with an extra copy of CtfAB. Selectivity is the proportion of energy in the consumed feed gas that goes toward the product of interest on an electron basis. Data points represent maximum steady state conditions where metabolite and gas profiles were stable over multiple days. Respective run profiles are shown in Supplementary Fig. 11. WT, wild-type.



**Fig. 6 | Life cycle analysis.** **a**, Comparison of current routes to acetone and IPA and the developed gas fermentation process (starting from industrial waste gases, municipal solid waste (MSW) or biomass). Steps that are particularly energy intensive or result in toxic waste are indicated by symbols. **b**, Process flow diagram of LanzaTech’s gas fermentation process that shows the various inputs, outputs and process steps as well as the system boundary used for LCA. **c**, LCA analysis compares GHG emissions (shown in black) of traditional production routes with the gas fermentation route. Where traditional routes result in GHG emissions, gas fermentation avoids GHG emissions, effectively fixing carbon into the product. The effects of process (blue) and avoided off-gas (green) are indicated (see Supplementary Tables 9 and 10 for details).

and human investment<sup>54,55</sup>. The economics of turning one commodity (sugar) into another commodity and operating in batch processing mode is challenging; even for high-yielding pathways, it requires extremely high productivities and titers<sup>53</sup>. In the case of acetone and IPA production from sugars, the maximum yield is 50%<sup>14</sup>. Despite substantial development and highly engineered systems that showed improvements over the historical ABE process<sup>7</sup>, only moderate productivities of up to 0.15 g/L/h for acetone and 0.6 g/L/h for IPA have been achieved<sup>11,13</sup>.

Using waste resources instead of sugars as feedstocks decouples production from commodity prices. The first industrial-scale gas fermentation plants are now in operation and are turning waste into ethanol in a continuous fermentation using acetogen *C. autoethanogenum*<sup>15,29,31</sup>. Ethanol is a native product of acetogens<sup>16</sup>, but, through genetic engineering, synthesis of a range of non-native commodity chemicals has been demonstrated at laboratory scale. However, achieved efficiencies (<100 mg/L/h) and selectivities (<10%) remain low, and production is not sustainable over longer periods

in continuous mode. For acetone and IPA specifically, the highest reported autotrophic production rates are 3.8 mg/L/h with a selectivity of 1.2%<sup>17–24</sup>, results that are insufficient for commercialization<sup>53</sup>.

The bioprocess described here achieves industrially relevant productivities of up to ~3 g/L/h and showed continuous production from C1 gases for prolonged periods (~3 weeks) without a decrease in performance. Engineered strains achieved high selectivities of up to 90%, which were previously reported only in acetogens for the native products acetate and ethanol. As acetogens live on the thermodynamic edge of life, it has long been an open question whether they can be engineered to synthesize complex molecules at high selectivity<sup>56,57</sup>. Our results show that this is possible.

To achieve the observed level of performance, we relied on a three-pronged approach comprising innovations in pathway optimization, strain optimization and process development. Our approach has several key features. First, we explored a large sequence space of biosynthetic genes to identify an optimal pathway design. Whereas previous studies aiming to optimize autotrophic acetone or IPA production explored only a handful of designs that all relied on the same gene sequences from two natural isolates<sup>17–24</sup>, we mined a large genome collection of industrial strains<sup>32,33</sup> and screened a library of hundreds of designs, which led to a significant increase in production level. In some instances, even a few mutations resulted in several-fold improvement in performance. The developed combinatorial library framework that enabled us to prototype this design space will be applicable to generation and screening of other pathways in acetogens and other autotrophs.

Second, to achieve high selectivity, we implemented a substantial level of strain engineering. Previous studies relied on plasmid-based pathway expression with a maximum of one additional genome modification. Our final production strain comprised multiple genome modifications, including pathway integration, four gene KOs and overexpression of two genes. To our knowledge, this represents the largest number of combined genome modifications in an autotroph or *Clostridium* strain. However, our original pool of candidates to KO was even larger, and it would have taken many months to years to generate all possible KO combinations. To guide this effort, we applied a recently described cell-free framework (iPROBE)<sup>34</sup> to rapidly prototype candidate genes for KO efficacy, streamlining an often laborious and time-consuming strain generation process<sup>17,29</sup>. A common challenge in metabolic engineering is that annotations of enzymes are often poor, and many enzymes are promiscuous and present as multiple isoenzymes<sup>58–60</sup>. In our case, we observed unwanted byproducts when introducing our heterologous pathway from interactions with native enzymes, lowering overall yield and increasing downstream separation requirements. An in silico search yielded a large number of candidate enzymes, and considering that these may act in combination, iPROBE<sup>34</sup>, which was applied here in a different way, allowed us to narrow down this list to a few key candidate enzymes in days, enabling precision gene KOs to effectively shut down byproduct formation. Our work also highlights the value of recent approaches to systematically expand genome-scale models to incorporate putative reactions arising from enzyme promiscuity<sup>61</sup>.

Working both in vitro (using iPROBE)<sup>34</sup> and in vivo, we interwove omics measurements, kinetic ensemble modelling and systems biology analyses to identify critical bottlenecks in the production pathway that, when alleviated, led to considerably improved flux to the target product. Although we speculated that the thiolase step was the bottleneck in our pathway (given that the Claisen condensation of two molecules of acetyl-CoA to form acetoacetyl-CoA is an endergonic reaction)<sup>62</sup>, our analysis correctly predicted that, in fact, the downstream CoA transferase step is most critical (providing an important pull for the pathway). Careful balancing of the pathway turned out to be critical; for example, acetate production could not be simply eliminated as it is required as co-factor for the

CoA transferase. To enable IPA production, we screened a library of engineered sADH variants. Variant S199R, which has been engineered to have a less rigid active site cavity, was thereby found to provide the best performance, matching the in vitro data for catalytic activity on acetone<sup>37</sup>.

Third, we scaled-up the process in a 120-L field pilot using a loop reactor designed to maximize energy efficiency and resembling industrial gas fermentation operations carried out at 46,000 metric tons per annum scale<sup>15,29</sup>. A key challenge with the scale-up of fermentations is the increasing heterogeneity of the bioreactor environment with scale<sup>31,55</sup>. Achieving high rates of gas conversion in a single pass is a desirable performance metric for a gas fermentation process<sup>31</sup>. However, achieving this means that there will be zones within the bioreactor where the fermentation broth has a very high gas substrate content (for example, where the gas substrate is continuously delivered into the reactor) and other zones with very low gas substrate content (for example, where any unreacted gas substrates and inert gases exit the bioreactor). Microbes within the reactor cycle through these different zones continuously. This heterogenous environment can be challenging to accurately reproduce at a bench scale. It is, therefore, important to understand the limits of metabolic robustness, as was done in detail for ethanol production<sup>50</sup>. Thermodynamic and kinetic models allowed us to explore and optimize pathway balance and metabolic robustness for acetone and IPA production. Developing process monitoring and control strategies that enable product selectivity and yields to be maintained despite these fluctuations within the bioreactor environment is a key challenge when scaling a gas fermentation technology. Specifically, defining the mechanisms and degree to which the gas, chemical and physical environment can influence selectivity is critical for dynamically adjusting one parameter to compensate for imposed fluctuations in another. We performed rigorous testing of performance across a spectrum of conditions, with selectivity maintained over a range of production rates. Even though the field pilot had been designed with the ethanol process in mind, scale-up of acetone and IPA was similar to previous scale-up of the ethanol process, and a good correlation between bench scale and pilot scale was observed across a range of production rates, with small variations due to process setup and elevated pressure levels. The observed performance provides confidence that the developed strains and fermentation process will translate to industrial scale.

The flexibility of fermentation to use the same bioreactor infrastructure for multiple conversions (for example, ethanol, acetone and IPA) is a key benefit over traditional chemical manufacturing (where plants are typically purpose-built for a single conversion process). The other key benefit offered by gas fermentation is the negative carbon footprint. Traditional chemical manufacturing processes are associated with substantial release of GHG emissions into the atmosphere (2.55 kgCO<sub>2</sub>e/kg of produced acetone and 1.85 kgCO<sub>2</sub>e/kg of produced IPA, respectively). In contrast, gas fermentation allows manufacturing of the same products while also avoiding release of GHG emissions, effectively fixing carbon into the product. LCA confirmed that acetone and IPA produced by gas fermentation have a negative carbon footprint of –1.78 kgCO<sub>2</sub>e/kg of produced acetone and –1.17 kgCO<sub>2</sub>e/kg of produced IPA. Our bioprocess provides a green alternative to today's production routes to these important chemicals, which exclusively rely on fresh fossil feedstocks. Thus, our work promotes implementation of a circular economic model in which the carbon from agriculture, industrial and societal waste streams is recycled into a chemical synthesis value chain and displaces manufacture of products from fresh fossil resources.

### Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of



author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41587-021-01195-w>.

Received: 6 October 2020; Accepted: 9 December 2021;  
Published online: 21 February 2022

## References

1. Aguado-Deblas, L. et al. Acetone prospect as an additive to allow the use of castor and sunflower oils as drop-in biofuels in diesel/acetone/vegetable oil triple blends for application in diesel engines. *Molecules* **25**, 2935 (2020).
2. Elfasakhany, A. Performance and emissions analysis on using acetone–gasoline fuel blends in spark-ignition engine. *Eng. Sci. Technol. Int. J.* **19**, 1224–1232 (2016).
3. Anbarasan, P. et al. Integration of chemical catalysis with extractive fermentation to produce fuels. *Nature* **491**, 235–239 (2012).
4. Ryan, C. F. et al. Synthesis of aviation fuel from bio-derived isophorone. *Sustain. Energy Fuels* **4**, 1088–1092 (2020).
5. Kratzel, A. et al. Inactivation of severe acute respiratory syndrome coronavirus 2 by WHO-recommended hand rub formulations and alcohols. *Emerg. Infect. Dis.* **26**, 1592–1595 (2020).
6. Wernet, G. et al. The ecoinvent database version 3 (part I): overview and methodology. *Int. J. Life Cycle Assess.* **21**, 1218–1230 (2016).
7. Jones, D. T. & Woods, D. R. Acetone–butanol fermentation revisited. *Microbiol. Rev.* **50**, 484–524 (1986).
8. Ismaiel, A. A., Zhu, C. X., Colby, G. D. & Chen, J. S. Purification and characterization of a primary–secondary alcohol dehydrogenase from two strains of *Clostridium beijerinckii*. *J. Bacteriol.* **175**, 5097–5105 (1993).
9. Hanai, T., Atsumi, S. & Liao, J. C. Engineered synthetic pathway for isopropanol production in *Escherichia coli*. *Appl. Environ. Microbiol.* **73**, 7814–7818 (2007).
10. May, A. et al. A modified pathway for the production of acetone in *Escherichia coli*. *Metab. Eng.* **15**, 218–225 (2013).
11. Liang, L. et al. CRISPR Enabled Trackable genome Engineering for isopropanol production in *Escherichia coli*. *Metab. Eng.* **41**, 1–10 (2017).
12. Soma, Y., Yamaji, T., Matsuda, F. & Hanai, T. Synthetic metabolic bypass for a metabolic toggle switch enhances acetyl-CoA supply for isopropanol production by *Escherichia coli*. *J. Biosci. Bioeng.* **123**, 625–633 (2017).
13. Jojima, T., Inui, M. & Yukawa, H. Production of isopropanol by metabolically engineered *Escherichia coli*. *Appl. Microbiol. Biotechnol.* **77**, 1219–1224 (2008).
14. Jones, S. W. et al. CO<sub>2</sub> fixation by anaerobic non-photosynthetic mixotrophy for improved carbon conversion. *Nat. Commun.* **7**, 12800 (2016).
15. Simpson, S. D. & Köpke, M. Pollution to products: recycling of ‘above ground’ carbon by gas fermentation. *Curr. Opin. Biotechnol.* **65**, 180–189 (2020).
16. Marcellin, E. et al. Low carbon fuels and commodity chemicals from waste gases—systematic approach to understand energy metabolism in a model acetogen. *Green Chem.* **18**, 3020 (2016).
17. Kato, J. et al. Metabolic engineering of *Moorella thermoacetica* for thermophilic bioconversion of gaseous substrates to a volatile chemical. *AMB Express* **11**, 59 (2021).
18. Hoffmeister, S. et al. Acetone production with metabolically engineered strains of *Acetobacterium woodii*. *Metab. Eng.* **36**, 37–47 (2016).
19. Banerjee, A., Leang, C., Ueki, T., Nevin, K. P. & Lovley, D. R. A lactose-inducible system for metabolic engineering of *Clostridium ljungdahlii*. *Appl. Environ. Microbiol.* **80**, 2410–2416 (2014).
20. Garrigues, L., Maignien, L., Lombard, E., Singh, J. & Guillouet, S. E. Isopropanol production from carbon dioxide in *Cupriavidus necator* in a pressurized bioreactor. *N. Biotechnol.* **56**, 16–20 (2020).
21. Lee, H. J., Son, J., Sim, S. J. & Woo, H. M. Metabolic rewiring of synthetic pyruvate dehydrogenase bypasses for acetone production in cyanobacteria. *Plant Biotechnol. J.* **18**, 1860–1868 (2020).
22. Torella, J. P. et al. Efficient solar-to-fuels production from a hybrid microbial–water-splitting catalyst system. *Proc. Natl Acad. Sci. USA* **112**, 2337–2342 (2015).
23. Hirokawa, Y., Dempo, Y., Fukusaki, E. & Hanai, T. Metabolic engineering for isopropanol production by an engineered cyanobacterium, *Synechococcus elongatus* PCC 7942, under photosynthetic conditions. *J. Biosci. Bioeng.* **123**, 39–45 (2016).
24. Zhou, J., Zhang, H., Zhang, Y., Li, Y. & Ma, Y. Designing and creating a modularized synthetic pathway in cyanobacterium *Synechocystis* enables production of acetone from carbon dioxide. *Metab. Eng.* **14**, 394–400 (2012).
25. Claassens, N. J., Cotton, C. A. R., Kopljar, D. & Bar-Even, A. Making quantitative sense of electromicrobial production. *Nat. Catal.* **2**, 437–447 (2019).
26. Fast, A. G. & Papoutsakis, E. T. Stoichiometric and energetic analyses of non-photosynthetic CO<sub>2</sub>-fixation pathways to support synthetic biology strategies for production of fuels and chemicals. *Curr. Opin. Chem. Eng.* **1**, 380–395 (2012).
27. Bar-Even, A., Noor, E. & Milo, R. A survey of carbon fixation pathways through a quantitative lens. *J. Exp. Bot.* **63**, 2325–2342 (2012).
28. Köpke, M. et al. *Clostridium ljungdahlii* represents a microbial production platform based on syngas. *Proc. Natl Acad. Sci. USA* **107**, 13087–13092 (2010).
29. Fackler, N. et al. Stepping on the gas to a circular economy: accelerating development of carbon-negative chemical production from gas fermentation. *Ann. Rev. Chem. Biomol. Eng.* **12**, 439–470 (2021).
30. Jin, S. et al. Synthetic biology on acetogenic bacteria for highly efficient conversion of C1 gases to biochemicals. *Int. J. Mol. Sci.* **21**, 7639 (2020).
31. Takors, R. et al. Using gas mixtures of CO, CO<sub>2</sub> and H<sub>2</sub> as microbial substrates: the do’s and don’ts of successful technology transfer from laboratory to production scale. *Microb. Biotechnol.* **11**, 606–625 (2018).
32. Jones, D. T. Applied acetone–butanol fermentation. In: *Clostridia: Biotechnology and Medical Applications* (eds Dürre, P. & Bahl, H.) 125–168 (Wiley, 2001).
33. Jones, D. T. & Keis, S. Origins and relationships of industrial solvent-producing clostridial strains. *FEMS Microbiol. Rev.* **17**, 223–232 (1995).
34. Karim, A. S. et al. In vitro prototyping and rapid optimization of biosynthetic enzymes for cellular design. *Nat. Chem. Biol.* **16**, 912–919 (2020).
35. Krüger, A. et al. Development of a clostridia-based cell-free system for prototyping genetic parts and metabolic pathways. *Metab. Eng.* **62**, 95–105 (2020).
36. Heap, J. T., Pennington, O. J., Cartman, S. T. & Minton, N. P. A modular system for *Clostridium* shuttle plasmids. *J. Microbiol. Methods* **78**, 79–85 (2009).
37. Maddock, D. J., Patrick, W. M. & Gerth, M. L. Substitutions at the cofactor phosphate-binding site of a clostridial alcohol dehydrogenase lead to unexpected changes in substrate specificity. *Protein Eng. Des. Sel.* **28**, 251–258 (2015).
38. 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).
39. Valgepea, K. et al. H<sub>2</sub> drives metabolic rearrangements in gas-fermenting *Clostridium autoethanogenum*. *Biotechnol. Biofuels* **11**, 55 (2018).
40. de Souza Pinto Lemgruber, R. et al. Systems-level engineering and characterisation of *Clostridium autoethanogenum* through heterologous production of poly-3-hydroxybutyrate (PHB). *Metab. Eng.* **53**, 14–23 (2019).
41. Maia, P., Rocha, I. & Rocha, M. Identification of robust strain designs via tandem pFBA/LMOMA phenotype prediction. In: *GECCO 2017: Proceedings of the Genetic and Evolutionary Computation Conference Companion 1661–1668* (Association for Computing Machinery, 2017).
42. Zitzler, E., Laumanns, M. & Thiele, L. SPEA2: improving the Strength Pareto Evolutionary Algorithm. In: *Proceedings of the Fifth Conference on Evolutionary Methods for Design 95–100* (Association for Computing Machinery, 2001).
43. Takanashi, M. & Saito, T. Characterization of two 3-hydroxybutyrate dehydrogenases in poly(3-hydroxybutyrate)-degradable bacterium, *Ralstonia pickettii* T1. *J. Biosci. Bioeng.* **101**, 501–507 (2006).
44. Segawa, M., Wen, C., Orita, I., Nakamura, S. & Fukui, T. Two NADH-dependent (S)-3-hydroxyacyl-CoA dehydrogenases from polyhydroxyalkanoate-producing *Ralstonia eutropha*. *J. Biosci. Bioeng.* **127**, 294–300 (2019).
45. Tan, Y., Liu, Z.-Y., Liu, Z. & Li, F.-L. Characterization of an acetoin reductase/2,3-butanediol dehydrogenase from *Clostridium ljungdahlii* DSM 13528. *Enzyme Microb. Technol.* **79–80**, 1–7 (2015).
46. Kim, J., Chang, J. H., Kim, E. J. & Kim, K. J. Crystal structure of (R)-3-hydroxybutyryl-CoA dehydrogenase PhaB from *Ralstonia eutropha*. *Biochem. Biophys. Res. Commun.* **443**, 783–788 (2014).
47. Boynton, Z. L., Bennet, G. N. & Rudolph, F. B. Cloning, sequencing, and expression of clustered genes encoding beta-hydroxybutyryl-coenzyme A (CoA) dehydrogenase, crotonase, and butyryl-CoA dehydrogenase from *Clostridium acetobutylicum* ATCC 824. *J. Bacteriol.* **178**, 3015–3024 (1996).
48. Vick, J. E. et al. *Escherichia coli* enoyl-acyl carrier protein reductase (FabI) supports efficient operation of a functional reversal of the β-oxidation cycle. *Appl. Environ. Microbiol.* **81**, 1406–1416 (2015).
49. Liew, F. et al. Metabolic engineering of *Clostridium autoethanogenum* for selective alcohol production. *Metab. Eng.* **40**, 104–114 (2017).
50. Mahamkali, V. et al. Redox controls metabolic robustness in the gas-fermenting acetogen *Clostridium autoethanogenum*. *Proc. Natl Acad. Sci. USA* **117**, 13168–13175 (2020).
51. Greene, J., Daniell, J., Köpke, M., Broadbelt, L. & Tyo, K. E. J. Kinetic ensemble model of gas fermenting *Clostridium autoethanogenum* for improved ethanol production. *Biochem. Eng. J.* **148**, 46–56 (2019).
52. Wehrs, M. et al. Engineering robust production microbes for large-scale cultivation. *Trends Microbiol.* **27**, 524–537 (2019).
53. Hoff, B. et al. Unlocking nature’s biosynthetic power—metabolic engineering for the fermentative production of chemicals. *Angew. Chem. Int. Ed. Engl.* **60**, 2258–2278 (2021).

54. Nielsen, J. & Keasling, J. D. Engineering cellular metabolism. *Cell* **164**, 1185–1197 (2016).
55. Crater, J. S. & Lievense, J. C. Scale-up of industrial microbial processes. *FEMS Microbiol. Lett.* **365**, 138 (2018).
56. Bertsch, J. & Müller, V. Bioenergetic constraints for conversion of syngas to biofuels in acetogenic bacteria. *Biotechnol. Bioeng.* **8**, 210 (2015).
57. Schuchmann, K. & Müller, V. Autotrophy at the thermodynamic limit of life: a model for energy conservation in acetogenic bacteria. *Nat. Rev. Microbiol.* **12**, 809–821 (2014).
58. Erb, T. J. Back to the future: why we need enzymology to build a synthetic metabolism of the future. *Beilstein J. Org. Chem.* **15**, 551–557 (2019).
59. Diether, M., Nikolaev, Y., Allain, F. H. & Sauer, U. Systematic mapping of protein-metabolite interactions in central metabolism of *Escherichia coli*. *Mol. Syst. Biol.* **15**, e9008 (2019).
60. Kim, H. M., Chae, T. U., Choi, S. Y., Kim, W. J. & Lee, S. Y. Engineering of an oleaginous bacterium for the production of fatty acids and fuels. *Nat. Chem. Biol.* **15**, 721–729 (2019).
61. Amin, S. A., Chavez, E., Porokhin, V., Nair, N. U. & Hassoun, S. Towards creating an extended metabolic model (EMM) for *E. coli* using enzyme promiscuity prediction and metabolomics data. *Microb. Cell Fact.* **18**, 109 (2019).
62. Vögeli, B. et al. Archaeal acetoacetyl-CoA thiolase/HMG-CoA synthase complex channels the intermediate via a fused CoA-binding site. *Proc. Natl Acad. Sci. USA* **115**, 3380–3385 (2018).

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

© The Author(s), under exclusive licence to Springer Nature America, Inc. 2022

## Methods

**Strains and chemicals.** *C. autoethanogenum* DSM19630, a derivative of type strain DSM10061, obtained from the German Collection of Microorganisms and Cell Cultures, was used<sup>40</sup>. Strains were handled using anaerobic techniques and media described earlier<sup>40,50</sup>. Molecular-grade chemicals (Sigma-Aldrich) and RO water were used for laboratory experiments; for piloting, industrial chemicals and process water were used. For cloning and cell-free prototyping, *E. coli* strains NEB Express (NEB), DH5 $\alpha$  (Thermo Fisher Scientific) and BL21 Star (DE3) (Life Technologies) were used.

**Sequence mining.** Sequences for 272 industrial ABE strains are available through GenBank (Supplementary Table 7). Acetone biosynthesis genes (Supplementary Table 1) were extracted from these genomes by E.C. numbers (ThlA: E.C. 2.3.1.9, CtfAB E.C. 2.8.3.9 and Adc E.C. 4.1.1.4) annotated in the Joint Genome Institute Integrated Microbial Genomes & Microbiomes platform<sup>65</sup>. The nucleotide sequences were in silico translated, and the resulting amino acid sequences were aligned using ClustalW (version 2.1)<sup>66</sup> as a part of the Geneious (R9.1.8) software package (Biomatters). Phylogenetic trees were generated using IQ-TREE (version 1.6.12) software<sup>65</sup> with the mod LG + G4 substitution model. This model was chosen by running the first tree using the model finder option<sup>66</sup>. The resulting trees were visualized using the Interactive Tree Of Life (iTOL) (version 5) web service<sup>67</sup>.

**Combinatorial library.** The acetone combinatorial library employed a hierarchical Golden Gate assembly approach that first employed *AarI* to construct the donor vectors, before combining with a recipient vector to perform the final assembly using *BsaI*. The counter-selectable marker *ccdB* (flanked with *BsaI*) was cloned into pMTL82251-P<sub>fer/wl/pfor</sub><sup>36</sup> (Supplementary Table 1) to generate the recipient vectors. Each of these promoters has a built-in ribosomal binding site (Supplementary Table 1). Combinatorial-hierarchical designs were generated using the TeselaGen DESIGN module (version 19) (TeselaGen Biotechnology)<sup>68</sup>. The construction of donor vectors involved *AarI* assembly into pTypeII vector (Invitrogen) using two DNA parts: amplicon containing transcriptional terminator and promoter that was subcloned into plasmid pCR-Blunt (Invitrogen) and amplicon containing *thlA*, *ctfA-ctfB* or *adc*. Primer sequence (Supplementary Table 8) and a detailed step-by-step description can be found in the Supplementary Text and Supplementary Fig. 11. After *BsaI* assembly in *E. coli*, the resulting acetone combinatorial plasmids were sequence verified using MiSeq (Illumina) as described below. After transformation<sup>40</sup> into *C. autoethanogenum*  $\Delta$ 0553 strain, the resulting acetone combinatorial strains were subjected to Schott bottle growth on 1.5-bar synthetic gas blend (50% CO, 10% H<sub>2</sub>, 30% CO<sub>2</sub> and 10% N<sub>2</sub>; Airgas) for 8 d before samples were collected for OD<sub>600nm</sub> measurement and high-performance liquid chromatography (HPLC) analysis as described below.

**Gene expression.** Plasmids for expression of sAdh variants and an extra copy of CtfAB were constructed as follows, before transformation and screening, as described above. The CtfAB overexpression vectors were each built using seamless cloning techniques using the same plasmid backbone as the combinatorial library. Oligonucleotides used for amplifying the backbone and the overexpression insert, also from the combinatorial library, are listed in Supplementary Table 8. Plasmids for overexpressing sAdh variants were obtained from previous work<sup>37,38</sup>. Testing of the sAdh variants was conducted in the same growth conditions as described above with the addition of an acetone spike, 20 g L<sup>-1</sup> final concentration, 24 h after inoculation. Each bottle was immediately sampled for HPLC analysis as described below, followed by daily samples for the next 8 d.

**Chassis strain construction.** Chromosomal gene insertion and deletion mutants were constructed via homologous recombination based on the methods described in Liew et al.<sup>69</sup>. In brief, shuttle plasmids containing homology arms were designed using the TeselaGen DESIGN module (version 19) (TeselaGen Biotechnology)<sup>68</sup>, assembled and then transformed<sup>40</sup> into *C. autoethanogenum*. Screening of correct insertion or deletion mutants were carried out using PCR and confirmed by DNA sequencing as described below. Oligonucleotides used for amplification of homology arms and PCR screening are listed in Supplementary Table 8.

**Genome-scale and kinetic modeling.** Genome-scale modeling was carried out with model iCLAU786 (ref. <sup>70</sup>) expanded by missing secondary alcohol dehydrogenase and acyl-CoA thioesterase reactions as specified in Supplementary Table 3. COBRAPy (version 0.8.2) and OptFlux (version 3.2.10) were used for simulations with Gurobi (version 7.0.2) optimizer<sup>71,72</sup>. Strain optimization targets were identified using Strength Pareto Evolutionary Algorithm 2 (ref. <sup>43</sup>), with mutant strain phenotypes predicted using parsimonious flux balance analysis and minimization of metabolic adjustment<sup>41</sup>. The multi-objective strain optimization evaluated designs with a maximum of 20 gene KOs and excluded essential and transport reactions as targets. The optimization objective functions maximized product flux and minimized the number of KOs, subject to maintaining a minimum of 20% of wild-type growth rate. Kinetic ensemble modeling was carried out as described previously<sup>31</sup>. The core metabolic network used has been expanded with the acetone pathway and is shown in Supplementary Fig. 9.

**Cell-free prototyping.** Crude *E. coli* cell extracts were prepared as previously described<sup>73</sup>, using BL21 Star (DE3) cells (Life Technologies). CFE was carried out based on the Cytomim system<sup>74</sup>, and the expression of each enzyme was quantified using radioactive leucine incorporation assays as previously described using a MicroBeta2 instrument (PerkinElmer)<sup>75</sup>. Acetone biosynthesis genes (Supplementary Table 1) and effector gene candidates (Supplementary Table 4) were codon-optimized for *E. coli*, cloned into pJLI plasmid (Addgene, 69496) for CFE and sequence-confirmed as described below. Cell-free acetone biosynthesis reactions contained 120 mM glucose, 8 mM magnesium glutamate, 10 mM ammonium glutamate, 134 mM potassium glutamate, 100 mM Bis-Tris buffer, 0.5 mM kanamycin, 1 mM adenosine triphosphate, 1 mM nicotinamide adenine dinucleotide, 1 mM coenzyme A and varying concentrations of acetone pathway enzymes produced in CFE. Acetone pathway enzymes were added at a final concentration of 0.5  $\mu$ M for gene KO prototyping. Buffer pH was adjusted with glacial acetic acid to decrease reaction pH to ~6 for increased activity of acetone pathway enzymes<sup>76</sup>. After 20 h, reactions were quenched with 20  $\mu$ l of 10% (w/v) trichloroacetic acid, and precipitated proteins were removed by centrifugation for 10 min at 20,000g. Supernatant was transferred to fresh 1.5-ml tubes and centrifuged again for 10 min at 20,000g to remove residual protein before chromatography.

**DNA sequencing.** Sequence confirmation of PCR amplicons, assembled plasmid constructs and generated strains was carried out using MiSeq (Illumina). Clonal plasmids were isolated using the QIAprep Miniprep Kit (Qiagen) following the protocol. Paired-end libraries for next-generation sequencing were produced using the Nextera XT Kit (Illumina) and quantified by Quant-iT (Thermo Fisher Scientific) before 2 $\times$  300-bp sequencing on a MiSeq Reagent Kit V3 (Illumina). Sequences were processed, aligned and checked for mismatches using the Geneious (R9.1.8) software package (Biomatters).

**Analytcs.** Biomass concentrations were determined using a spectrophotometer (OD<sub>600nm</sub>) and dry cell weight measurements as described previously<sup>34,40</sup>.

Metabolites 2,3-BDO, 3-HB, acetate, acetone, ethanol and IPA from fermentation broth (or cell-free assay supernatants (5- $\mu$ l injection)) were measured using an Agilent 1100 and 1200 series HPLC system as described previously<sup>34</sup>. Samples were separated with 5 mM sulfuric acid flowing at 0.6 ml min<sup>-1</sup> on a Rezex ROA-Organic Acid H+ (8%) LC Column (Aminex) at 20°C. Metabolites were quantified using a refractive index detector based on the retention time of standard solutions for each compound.

Headspace gas composition data were measured by gas chromatography using a 490 Micro GC (Agilent) with two installed channels. Channel 1 was a 10-m Mol-Sieve column running at 70°C, with 200-kPa argon and a backflush time of 4.2 s; channel 2 was a 10-m PPQ column running at 90°C, with 150-kPa helium and no backflush. The injector temperature for both channels was set at 70°C. Run times were set to 120 s, but all peaks of interest usually eluted before 100 s.

In each case, data were collected using Agilent OpenLab CDS ChemStation software.

**Proteomics.** For proteomics samples, cells were grown in bottles to mid-log phase (OD<sub>600nm</sub> = 0.8–0.9) as described above. Cell pellets were harvested by centrifugation, followed by supernatant removal, rapid freezing in liquid nitrogen and storage at –80°C until analysis. Cells were lysed by bead-beating in Tris-HCl (100 mM at pH 8.0), adjusting sample to 4% SDS, and heat treatment (95°C for 10 min). Crude protein was obtained by centrifugation at 21,000g for 10 min followed by quantification with a NanoDrop OneC spectrophotometer (Thermo Fisher Scientific). Samples were adjusted to 10 mM dithiothreitol (10 min at 95°C) to reduce proteins and then 30 mM iodoacetamide (20 min at room temperature in darkness) and cleaned up via protein aggregation capture<sup>77</sup>. Aggregated protein (on magnetic Sera-Mag (GE Healthcare) beads) was then digested with proteomics-grade trypsin (1:75 w/w, Promega) in 100 mM Tris-HCl pH 8.0 overnight at 37°C and again for 3 h at 37°C the next day. Tryptic peptides were then collected, acidified to 0.5% formic acid, filtered through a 10-kDa MWCO spin filter (Vivaspin 2, Sartorius) and quantified by NanoDrop OneC. Three micrograms of peptides were then analyzed by 1D LC-MS/MS using a Vanquish uHPLC coupled directly to an Orbitrap Q Exactive mass spectrometer running Xcalibur (version 4.2.47) data collection software (Thermo Fisher Scientific), as previously described<sup>78</sup>. Peptides were separated by a 180-min organic gradient across an in-house pulled nanospray emitter packed with 15 cm of 1.7- $\mu$ m Kinetex reversed-phase resin (Phenomenex). Peptide fragmentation spectra were analyzed/sequenced by Proteome Discoverer (version 2.3) software (Thermo Fisher Scientific), and peptides were quantified by chromatographic area under the curve. *C. autoethanogenum* reference proteome (UniProt UP000017590) plus acetone biosynthesis enzymes were used as reference. Peptide abundances were summed to their respective proteins, and protein abundances were log<sub>2</sub>-transformed and normalized with InfernoRDN (version 1.1.6687.24196)<sup>79</sup>.

**Fermentation and piloting.** Continuous fermentation was carried out under strictly anaerobic conditions at 37°C in 2 L of BioFlo 115 CSTR reactors (New Brunswick) with consistent media supply and gas flow as described previously<sup>34,50</sup>.

CSTR runs used a synthetic gas blend of 50% CO<sub>2</sub>, 10% H<sub>2</sub>, 30% CO and 10% N<sub>2</sub> (Airgas) at atmospheric pressure and a condenser to minimize loss of volatile products. Pilot runs were performed with a blend of 50% CO<sub>2</sub>, 10% H<sub>2</sub>, 20% CO<sub>2</sub> and 20% N<sub>2</sub> and pressure of 1.4 bar in a custom pilot reactor (Fig. 5c) at LanzaTech's Freedom Pines site. Medium dilution rates used were between 1.0 and 6.5 d<sup>-1</sup> and gas flow rates from 200 to 800 standard cubic centimeters per minute (SCCM) for CSTRs and from 11.5 to 23 normal liters per minute (NLPm) for pilot reactors. After an initial ramp as described earlier, cultures were run as continuous chemostats. Pathway and culture integrity were routinely assessed using DNA sequencing as described above. Gas and metabolite measurements were automatically sampled every 1–3 h and analyzed as described above. Samples for biomass were taken at least three times a day and processed as described above. For productivity and selectivity calculations, only concentrations in the broth were considered (not considering any stripped product). Selectivity is a biocatalyst-specific measure that describes the proportion of energy in the consumed feed gas that goes toward the product of interest on electron basis. Gas utilization is calculated as follows: % of CO consumed = ((% CO<sub>out</sub> × Outlet Gas flow) – (% CO<sub>in</sub> × Inlet Gas flow)) / (% CO<sub>in</sub> × Inlet Gas flow).

**LCA.** The scope of the LCA is cradle to gate, meaning that end-of-life is excluded. Carbon in the steel mill off-gas would end up in the atmosphere without the gas fermentation process, so a credit is taken for carbon sequestered in the acetone (2.27 kgCO<sub>2</sub>e/kg acetone) and IPA (2.20 kgCO<sub>2</sub>e/kg IPA). The steel mill off-gas carries no burden of upstream steel mill emissions as the off-gas is an unintended consequence of the steel-making process. Therefore, the GHG emissions of steel production do not change with the addition of the gas fermentation process. The LCA treats the biogas and *C. autoethanogenum* microbial protein co-products using system expansion by taking credits for the former replacing natural gas and the latter replacing soybean meal based on protein content. Inputs that contributed less than 5% of the total GHG emissions were excluded from the analysis. Emission factors (Supplementary Tables 9 and 10) were taken directly from the GREET 1 model (<https://greet.es.anl.gov/>) and ecoinvent database<sup>6</sup>.

**Visualization.** Figure 1 was created with BioRender.

**Statistical information.** All statistical information provided in this manuscript is derived from  $n = 3$  or more independent experiments unless otherwise noted in the text or figure legends. Error bars represent 1 s.d. of the mean derived from these experiments. All error bars propagated errors. These data do not represent a distribution of measured data but, rather, a calculation with propagated errors. Statistical analyses of proteomics were performed with Perseus (version 1.6.14.0)<sup>80</sup>. Statistical significance ( $P \leq 0.05$ ) was assessed using two-tailed  $t$ -tests against the strain ranked 203 (C1; lowest acetone producer), and  $P$  values were adjusted using Benjamini–Hochberg correction (false discovery rate  $\leq 0.05$ ).

**Reporting Summary.** Further information on research design is available in the Nature Research Reporting Summary linked to this article.

## Data availability

All data presented in this manuscript are available as supplementary data files. Supplementary Tables 1–10 are available on Zenodo: <https://doi.org/10.5281/zenodo.5838304>. All proteomics raw data are available at the ProteomeXchange Consortium via the MassIVE repository (<ftp://massive.ucsd.edu/MSV000085940/>) (MassIVE accession: MSV000085940; ProteomeXchange accession: PXD020853). All genome sequences are available through the Joint Genome Institute Integrated Microbial Genomes & Microbiomes system platform or the National Center of Biotechnology Information GenBank under accession numbers provided in Supplementary Table 7. Any additional data may be available from the authors upon reasonable request. Materials are available upon reasonable request and under a material transfer agreement, but strains may require a license.

## References

- Chen, I. M. A. et al. The IMG/M data management and analysis system v.6.0: new tools and advanced capabilities. *Nucleic Acids Res.* **49**, D751–D763 (2021).
- Larkin, M. A. et al. Clustal W and Clustal X version 2.0. *Bioinformatics* **23**, 2947–2948 (2007).
- Nguyen, L., Schmidt, H. A., Von Haeseler, A. & Minh, B. Q. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol. Biol. Evol.* **32**, 268–274 (2014).
- Kalyaanamoorthy, S., Minh, B. Q., Wong, T. K. F., von Haeseler, A. & Jermini, L. S. ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat. Methods* **14**, 587–589 (2017).
- Letunic, I. & Bork, P. Interactive Tree Of Life (iTOL) v4: recent updates and new developments. *Nucleic Acids Res.* **47**, 256–259 (2019).
- Fero, M. J., Craft, J. K., Trang, V. & Hillson, N. J. Combinatorial-hierarchical DNA library design using the TeselaGen DESIGN Module with j5. *Methods Mol. Biol.* **2205**, 19–47 (2020).
- Liew, F. M. et al. Gas fermentation—a flexible platform for commercial scale production of low-carbon-fuels and chemicals from waste and renewable feedstocks. *Front. Microbiol.* **7**, 694 (2016).
- Valgepea, K. et al. Arginine deiminase pathway provides ATP and boosts growth of the gas-fermenting acetogen *Clostridium autoethanogenum*. *Metab. Eng.* **41**, 202–211 (2017).
- Ebrahim, A., Lerman, J. A., Palsson, B. O. & Hyduke, D. R. COBRAPy: Constraints-Based Reconstruction and Analysis for Python. *BMC Syst. Biol.* **7**, 74 (2013).
- Gonçalves, E., Pereira, R., Rocha, I. & Rocha, M. Optimization approaches for the in silico discovery of optimal targets for gene over/underexpression. *J. Comput. Biol.* **19**, 102–114 (2012).
- Kwon, Y. C. & Jewett, M. C. High-throughput preparation methods of crude extract for robust cell-free protein synthesis. *Sci. Rep.* **5**, 8663 (2015).
- Jewett, M. C. & Swartz, J. R. Mimicking the *Escherichia coli* cytoplasmic environment activates long-lived and efficient cell-free protein synthesis. *Biotechnol. Bioeng.* **86**, 19–26 (2004).
- Jewett, M. C., Calhoun, K. A., Voloshin, A., Wu, J. J. & Swartz, J. R. An integrated cell-free metabolic platform for protein production and synthetic biology. *Mol. Syst. Biol.* **4**, 220 (2008).
- Karim, A. S., Rasor, B. J. & Jewett, M. C. Enhancing control of cell-free metabolism through pH modulation. *Synth. Biol.* **5**, ysz027 (2020).
- Batth, T. S. et al. Protein aggregation capture on microparticles enables multipurpose proteomics sample preparation. *Mol. Cell. Proteomics* **18**, 1027–1035 (2019).
- Walker, C., Ryu, S., Giannone, R. J., Garcia, S. & Trinh, C. T. Understanding and eliminating the detrimental effect of thiamine deficiency on the oleaginous yeast *Yarrowia lipolytica*. *Appl. Environ. Microbiol.* **86**, e02299–19 (2020).
- Taverner, T. et al. DanteR: an extensible R-based tool for quantitative analysis of -omics data. *Bioinformatics* **28**, 2404–2406 (2012).
- Tyanova, S. et al. The Perseus computational platform for comprehensive analysis of (prote)omics data. *Nat. Methods* **13**, 731–740 (2016).

## Acknowledgements

We would like to thank members of LanzaTech's Synthetic Biology, Strain Development, Process Integration, Analytics, Engineering Design & Development, AI & Modeling, Computational Biology and Freedom Pines teams for their support and conversations about this work, in particular A. Juminaga, A. Quattlebaum, A. Shah, J. Winkler, J. Cogan, L. Fantroy, M. Maas, M. Martin, N. Gayner, N. Fackler, R. C. Tappel, S. Nagaraju, S. Chong, V. Reynoso, W. P. Mitchell and W. Allen. We would also like to thank the Joint Genome Institute DNA synthesis team for their support and conversations on this work, in particular J.-F. Chen, M. Harmon-Smith, R. Evans and Y. Yoshikuni. Funding: Acetone strain and process development, genome-scale modeling, LCA work and initial pilot runs were supported by the U.S. Department of Energy Bioenergy Technologies Office under contract nos. DE-EE0007566 and CRADA/NFE-16-06364 between LanzaTech and the Oak Ridge National Laboratory (F.L., R.N., T.A., C.C., R.O.J., L.W., J.S., P.C., S.D.T., Z.R., A.G., L.T., N.L.E., J.C.B., J.D., R.C., T.J.T., R.J.G., R.L.H., S.D.S., S.D.B., C.L. and M.K.). Cell-free prototyping work was funded by the U.S. Department of Energy Office of Science, Biological and Environmental Research Division, Genomic Science Program, under contract nos. DE-SC0018249 and FWP ERKP903 (F.L., B.J.R., R.O.J., N.L.E., T.J.T., R.J.G., R.L.H., A.S.K., S.D.S., S.D.B., M.C.J. and M.K.). This manuscript was co-authored by UT-Battelle under contract no. DE-AC05-00OR22725 with the U.S. Department of Energy (T.J.T., N.L.E., R.J.G. and R.L.H.). DNA synthesis for the gene libraries was supported by the Joint Genome Institute Community Science Program under award no. CSP-503280; <https://doi.org/10.46936/10.25585/60001121> (M.C.J. and M.K.). The work conducted by the U.S. Department of Energy Joint Genome Institute (<https://ror.org/04xm1d337>), a DOE Office of Science User Facility, is supported by the Office of Science of the U.S. Department of Energy operated under contract no. DE-AC02-05CH11231. B.J.R. is supported by a National Defense Science and Engineering Graduate Fellowship (award ND-CEN-017-095). We also thank the following investors in LanzaTech's technology: BASE, CICC Growth Capital Fund I, CITIC Capital, Indian Oil Company, K1W1, Khosla Ventures, the Malaysian Life Sciences Capital Fund, L. P. Mitsui, the New Zealand Superannuation Fund, Novo Holdings A/S, Petronas Technology Ventures, Primetals, Qiming Venture Partners, Softbank China and Suncor.

## Author contributions

T.J.T., A.S.K., S.D.S., S.D.B., C.L., M.C.J. and M.K. designed the study. R.O.J. and F.L. performed the sequence mining. F.L. developed the combinatorial library framework. F.L., R.N., R.O.J., L.W., C.L. and M.K. performed all strain engineering. A.P.M. and M.K. performed in silico analysis. J.S. and J.D. performed modeling. B.J.R. and A.S.K. performed all cell-free experiments. T.A., Z.R., L.T., J.C.B. and S.D.B. performed all fermentation experiments. C.C., A.G. and R.C. performed LCA. P.C., N.L.E., R.J.G. and R.L.H. performed all omics experiments. E.L., R.N., T.A., B.J.R., A.S.K., C.L. and M.K. performed data analysis and visualization. E.L., R.N., T.A., C.L., B.J.R., A.S.K., M.C.J. and M.K. wrote the manuscript. All authors edited and reviewed the manuscript.

### Competing interests

F.L., R.N., T.A., C.C., R.O.J., L.W., J.S., S.T., A.P.M., A.G., L.T., J.C.B., J.D., R.C., S.D.S., S.D.B., C.L. and M.K. are current or former employees of LanzaTech, a for-profit company pursuing commercialization of the acetone and isopropanol gas fermentation process discussed here. M.C.J. consults for and has joint funding with LanzaTech. F.L. and M.K. are co-inventors on granted US patent 9,365,868 (assigned to LanzaTech) related to the production of acetone and isopropanol by fermentation of a gaseous substrate comprising CO that incorporates discoveries incorporated in this manuscript. M.K. is co-inventor on granted US patent 9,550,979 (assigned to LanzaTech) related to alcohol dehydrogenases with engineered metabolic activity that incorporates discoveries incorporated in this manuscript. R.O.J., A.P.M. and M.K. are co-inventors on US patent application 2015/0152445 (assigned to LanzaTech) related to inactivation of secondary alcohol dehydrogenases that incorporates discoveries incorporated in this manuscript. C.L. is an inventor on provisional US patent application 63/035,739 (assigned to LanzaTech) related to integration of acetone and isopropanol biosynthesis gene variants that incorporates discoveries incorporated in this manuscript. R.N. is an inventor on provisional US patent application 63/083,257 (assigned to LanzaTech)

related to expressing multiple copies of acetone and isopropanol biosynthesis gene variants that incorporates discoveries incorporated in this manuscript. M.C.J. and A.S.K. are co-inventors on a US provisional patent application that incorporates discoveries described in this manuscript. Their interests are reviewed and managed by Northwestern University in accordance with their conflict of interest policies. All other authors declare no competing interests.

### Additional information

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41587-021-01195-w>.

**Correspondence and requests for materials** should be addressed to Ching Leang, Michael C. Jewett or Michael Köpke.

**Peer review information** *Nature Biotechnology* thanks Corinne Scown, Shang-Tian Yang and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

## Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see our [Editorial Policies](#) and the [Editorial Policy Checklist](#).

### Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a Confirmed

- The exact sample size ( $n$ ) for each experimental group/condition, given as a discrete number and unit of measurement
- A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- The statistical test(s) used AND whether they are one- or two-sided  
*Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
- A description of all covariates tested
- A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- For null hypothesis testing, the test statistic (e.g.  $F$ ,  $t$ ,  $r$ ) with confidence intervals, effect sizes, degrees of freedom and  $P$  value noted  
*Give  $P$  values as exact values whenever suitable.*
- For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- Estimates of effect sizes (e.g. Cohen's  $d$ , Pearson's  $r$ ), indicating how they were calculated

*Our web collection on [statistics for biologists](#) contains articles on many of the points above.*

### Software and code

Policy information about [availability of computer code](#)

#### Data collection

Gene/Protein sequences were extracted using the JGI Integrated Microbial Genomes & Microbiomes (IMG/M) platform as described in the Online Methods. DNA sequence data was generated using MiSeq and built in MiSeq Control Software (MCS) (v3.1.0) as described in the Online Methods.

Metabolite data from HPLC was measured on Agilent HPLC 1100 and 1200 series systems and gas composition data was measured using gas chromatography via Agilent 190 Micro GC systems. Data from both systems was collected using Agilent OpenLab CDS ChemStation (C.01.07) as described in the Online Methods

Proteomics LC-MS/MS data was collected on a Q Exactive Plus mass spectrometer running Xcalibur (v4.2.47; Thermo Scientific) data collection software. MS/MS data were searched using Proteome Discoverer (v2.3; Thermo Scientific) as described in the Online Methods.

Cell-free protein expression quantification data was generated using a MicroBeta2 instrument (Perkins Elmer) and built-in software.

No custom software was used for data collection.

#### Data analysis

Extracted and generated sequences were analyzed using Geneious (R9.1.8; Biomatters Ltd), ClustaW (v2.1), IQ-TREE (v1.6.12) as described in the Online Methods. DNA designs were generated using TeselaGen DESIGN module (v19) (TeselaGen Biotechnology; <https://teselagen.com/>).

Protein abundance data was log<sub>2</sub> transformed, normalized, and QA/QC'd with InfernoRDN (v1.1.6687.24196). Statistical analyses of proteomics data were performed with Perseus (v1.6.14.0) as described in the Online Methods.

Metabolic modeling was performed using COBRApy (v0.8.2) and OptFlux (v3.2.10) with Gurobi Optimizer (v7.0.2), as described in the Online methods.

Microsoft Excel was used for general data analysis. No custom software was used for data analysis.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research [guidelines for submitting code & software](#) for further information.

## Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

All data presented in this manuscript are available as supplementary data files or specified databases below.

All genome sequences are available through the JGI Integrated Microbial Genomes & Microbiomes system (IMG/M) platform or NCBI Genbank under accession numbers provided in Supplementary Tab. 7.

All proteomics raw data is available at the ProteomeXchange Consortium via the MassIVE repository (<ftp://massive.ucsd.edu/MSV000085940/>) MassIVE accession: MSV000085940; ProteomeXchange accession: PXD020853.

Emission factors for LCA analysis are available and were taken directly from the GREET 1 model (<https://greet.es.anl.gov/>) and ecoinvent database (<https://ecoinvent.org/the-ecoinvent-database/>).

Any additional data may be available from the authors upon reasonable request.

## Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

Life sciences  Behavioural & social sciences  Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/documents/nr-reporting-summary-flat.pdf](https://www.nature.com/documents/nr-reporting-summary-flat.pdf)

## Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	All sample sizes used are listed in the manuscript. No sample size calculation was performed. Sample size (n>=3) represents field standard and regular practice in literature. Quantification of cell-free expression levels to determine enzyme concentrations in cell-free acetone biosynthesis reactions was carried out as n = 2 as highly reproducible and small differences in levels would not impact conclusions.
Data exclusions	No data was excluded (i.e. no data was 'cherry-picked').
Replication	All attempts at replication were successful and reproducibility is built in to the biological replicates measured in this study. All strains were run more than once (i.e. a separate biologically replicated set) and data was reproducible. Key findings were confirmed through a combination of experiments and models.
Randomization	Samples were organized by experimental variables then characterized fully and reported in completion. Therefore, samples were not randomized.
Blinding	Blinding was not relevant. Animal or human participants were not used in the study, a distribution of biomolecules were measured that is unbiased and does not need blinding.

## Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

## Materials & experimental systems

- | n/a                                 | Included in the study                                  |
|-------------------------------------|--|
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Antibodies                    |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Eukaryotic cell lines         |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Palaeontology and archaeology |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Animals and other organisms   |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Human research participants   |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Clinical data                 |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Dual use research of concern  |

## Methods

- | n/a                                 | Included in the study                           |
|-------------------------------------|---|
| <input checked="" type="checkbox"/> | <input type="checkbox"/> ChIP-seq               |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Flow cytometry         |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> MRI-based neuroimaging |