

Bio-based Oxalic Acid Production in *Issatchenkia orientalis* Enables Sustainable Rare Earth Recovery

Jingxia Lu^{1,2}, Wenjun Guo^{3,4}, Ziyue Dong⁵, Sarang S. Bhagwat^{3,4}, Shih-I Tan¹, Zhixin Zhu^{1,2}, Andrew Johnson⁶, Jeremy S. Guest^{3,4,7}, Rick Honaker⁶, Dan McFarland Park⁵, Yongqin Jiao⁵, and Huimin Zhao^{1,2,3,8*}

¹The Grainger College of Engineering, Department of Chemical and Biomolecular Engineering, University of Illinois Urbana-Champaign, Urbana, IL 61801, USA

²Carl R. Woese Institute for Genomic Biology, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA

³DOE Center for Advanced Bioenergy and Bioproducts Innovation, University of Illinois Urbana-Champaign, Urbana, IL, 61801, USA

⁴The Grainger College of Engineering, Department of Civil and Environmental Engineering, University of Illinois Urbana-Champaign, Urbana, IL 61801, USA

⁵Physical and Life Sciences Directorate, Lawrence Livermore National Laboratory, Livermore, CA 94550, USA

⁶Department of Mining Engineering, University of Kentucky, Lexington, KY 40506, USA

⁷Institute for Sustainability, Energy, and Environment (iSEE), University of Illinois Urbana-Champaign, Urbana, IL 61801, USA

⁸NSF Molecule Maker Lab Institute, University of Illinois Urbana-Champaign, Urbana, IL 61801, USA

*Corresponding author Email: zhao5@illinois.edu

Supplementary Method 1 - Plasmid construction

Plasmids for gene overexpression were constructed using HiFi assembly in *E. coli*. PCR-amplified DNA fragments (50–100 ng) and PCR-amplified backbones were assembled via HiFi assembly (New England Biolabs, Ipswich, MA) and transformed into *E. coli*. Plasmids from several *E. coli* colonies were extracted and verified by restriction digestion. To construct plasmids for gene deletion, gBlocks containing homology donors and spacers were synthesized by Integrated DNA Technologies (Coralville, IA). The homology donor contained two 50 bp homology arms flanking the Cas9 cutting site, and a restriction enzyme, such as *Xho*I, was in the middle of the two homology arms. The gBlocks were then cloned into the CRISPR/Cas9 plasmid by Golden Gate assembly. Detailed construction of each plasmid is described as following:

p_Pcoah. pVT101a containing TDH3p promoter and terminator 3376 was linearized by primer pairs of Pcoah_BG_F and Pcoah_BG_R. Gene *Pcoah* synthesized by Twist Bioscience (San Francisco, CA) was amplified using primer of Pcoah_F and Pcoah_R. HiFi assembly was used to assemble fragments.

p_Fpoah. pVT101a containing TDH3p promoter and terminator 3376 was linearized by primer pairs of Pcoah_BG_F and Pcoah_BG_R. Gene *Fpoah* synthesized by Twist Bioscience (San Francisco, CA) was amplified using primer of Pcoah_F and Pcoah_R. HiFi assembly was used to assemble fragments.

p_FpOAR. pVT101a containing TDH3p promoter and terminator 3376 was linearized by primer pairs of FpOAR_BG_F and FpOAR_BG_R. Gene *FpOAR* synthesized by Twist Bioscience (San Francisco, CA) was amplified using primer of FpOAR_F and FpOAR_R. HiFi assembly was used to assemble fragments.

p_HboT1. pVT101a containing TDH3p promoter and terminator 3376 was linearized by primer pairs of HboT1_BG_F and HboT1_BG_R. Gene *HboT1* synthesized by Twist Bioscience (San Francisco, CA) was amplified using primer of HboT1_F and HboT1_R. HiFi assembly was used to assemble fragments.

p_HboT2. pVT101a containing TDH3p promoter and terminator 3376 was linearized by primer pairs of HboT2_BG_F and HboT2_BG_R. Gene *HboT2* synthesized by Twist Bioscience (San Francisco, CA) was amplified using primer of HboT2_F and HboT2_R. HiFi assembly was used to assemble fragments.

p_PYK1. *PYK1* was amplified from *S. cerevisiae* genomic DNA using primer pairs of PYK1_F and PYK1_R. pVT101a backbone containing TDH3p promoter and terminator 3376 was linearized by primer pairs of PYK1_BG_F and PYK1_BG_R. HiFi assembly was used to assemble fragments.

p_g4285. *g4285* was amplified from *I. orientalis* genomic DNA using primer pairs of g4285_F and g4285_R. pVT101a backbone containing TDH3p promoter and terminator 3376 was linearized by primer pairs of g4285_BG_F and g4285_BG_R. HiFi assembly was used to assemble fragments.

Supplementary Method 2 - Strain construction

For gene deletion, CRISPR/Cas9 plasmid was transformed to *I. orientalis* via a lithium acetate-mediated method¹, and transformants were plated on an SC-URA plate and incubated at 30 °C. Several colonies were randomly picked and cultivated in liquid SC-URA medium. Then, colony PCR was performed or then restriction digestion was performed to verify the deletion. For gene overexpression, the cassette for gene overexpression was PCR-amplified to contain 40 bp to 42-bp homology arms flanking the integration site. The PCR product and CRISPR/Cas9 plasmid targeting the integration site were co-transformed to *I. orientalis* via the lithium acetate-mediated method. Transformants were plated on an SC-URA plate and incubated at 30 °C. Several colonies were randomly picked and cultivated in liquid SC-URA medium. Then, colony PCR was performed for verifying gene integration. Plasmid removal was performed using FOA selection. The *ura3* expression cassette was further transformed to the engineered strain. Detailed construction of each strain is described as following:

OA_Pcoah_C1/ura3Δ. oah_LHR_F and oah_LHR_R were used to amplify the *Pcoah* expression cassette from p_Pcoah. Amplified product was then transformed into LPB/ura3Δ.

OA_Pcoah_C2/ura3Δ. oah_EG4_LHR_F and oah_EG4_RHR_F were used to amplify the *Pcoah* expression cassette from p_Pcoah. Amplified product and pVT36b-EG4 plasmid were then transformed into LPB/ura3Δ.

OA_Pcoah/ura3Δ. oah_EG6_LHR_F and oah_EG6_RHR_F were used to amplify the *Pcoah* expression cassette from p_Pcoah. Amplified product and pVT36b-EG6 plasmid were then transformed into LPB/ura3Δ.

OA_Fpoah_C1/ura3Δ. oah_LHR_F and oah_LHR_R were used to amplify the *Fpoah* expression cassette from p_Fpoah. Amplified product was then transformed into LPB/ura3Δ.

OA_Fpoah_C2/ura3Δ. oah_EG4_LHR_F and oah_EG4_RHR_F were used to amplify the *Fpoah* expression cassette from p_Fpoah. Amplified product and pVT36b-EG4 plasmid were then transformed into LPB/ura3Δ.

OA_Fpoah_C3/ura3Δ. oah_EG6_LHR_F and oah_EG6_RHR_F were used to amplify the *Fpoah* expression cassette from p_Fpoah. Amplified product and pVT36b-EG6 plasmid were then transformed into LPB/ura3Δ.

OA_Pcoah_PYC_C1/ura3Δ. PYC_LP3_LHR_F and PYC_LP3_LHR_R were used to amplify the *PYC* expression cassette from pRS416-SA-site1. Amplified product and pVT36b-LP3 plasmid were then transformed into OA_Pcoah/ura3Δ.

OA_Pcoah_PYC/ura3Δ. PYC_EG4_LHR_F and PYC_EG4_LHR_R were used to amplify the *PYC* expression cassette from pRS416-SA-site1. Amplified product and pVT36b-EG4 plasmid were then transformed into OA_Pcoah/ura3Δ.

OA_Pcoah_PYC_C3/ura3Δ. PYC_GD5_LHR_F and PYC_GD5_LHR_R were used to amplify the *PYC* expression cassette from pRS416-SA-site1. Amplified product and pVT36b-GD5 plasmid were then transformed into OA_Pcoah/ura3Δ.

OA_Pcoah/FpOAR/ura3Δ. TDH3p_int5_F and ter3376_int5_R were used to amplify the *Pcoah* expression cassette from p_FpOAR. Amplified product and pVT36b-int5 plasmid were then transformed into OA_Pcoah/ura3Δ.

OA_Pcoah/Hbot1/ura3Δ. TDH3p_int5_F and ter3376_int5_R were used to amplify the *Pcoah* expression cassette from p_Hbot1. Amplified product and pVT36b-int5 plasmid were then transformed into OA_Pcoah/ura3Δ.

OA_Pcoah/Hbot2/ura3Δ. TDH3p_int5_F and ter3376_int5_R were used to amplify the *Pcoah* expression cassette from p_Hbot2. Amplified product and pVT36b-int5 plasmid were then transformed into OA_Pcoah/ura3Δ.

JL01/ura3Δ. TDH3p_int5_F and ter3376_int5_R were used to amplify the *Hbot2* expression cassette from p_Hbot2. Amplified product and pVT36b-int5 plasmid were then transformed into OA_Pcoah_PYC/ura3Δ.

JL02/ura3Δ. pVT36b-gpd which contained the homology arm targeting to *gpd* gene was transformed into JL01/ura3Δ.

JL03/ura3Δ. int3_LHR_F and int3_RHR_R were used to amplify the *PYK1* expression cassette from p_PYK1. Amplified product and pVT36b-int3 plasmid were then transformed into JL02/ura3Δ.

JL02/g4285/ura3Δ. int3_LHR_F and int3_RHR_R were used to amplify the *g4285* expression cassette from p_g4285. Amplified product and pVT36b-int3 plasmid were then transformed into JL02/*ura3*Δ.

JL04/ura3Δ. pVT36b-g4567 which contained the homology arm targeting to the *g4567* gene was transformed into JL03/*ura3*Δ.

Supplementary Table 1. Strains used in this study.

Strains	Features	Sources
<i>E. coli</i> DH5α	cloning host	NEB
SD108/ <i>ura3</i> Δ	<i>I. orientalis</i> SD108 with uracil autotrophy	Ref. 2
LPB/ <i>ura3</i> Δ	SD108/ <i>ura3</i> Δ with landing pad integration	Ref. 3
OA_Pcoah_C1/ <i>ura3</i> Δ	<i>Pcoah</i> integrated to LPB/ <i>ura3</i> Δ at <i>URA3</i>	This study
OA_Pcoah_C2/ <i>ura3</i> Δ	<i>Pcoah</i> integrated to LPB/ <i>ura3</i> Δ at <i>EG4</i>	This study
OA_Pcoah/ <i>ura3</i> Δ	<i>Pcoah</i> integrated to LPB/ <i>ura3</i> Δ at <i>EG6</i>	This study
OA_Fpoah_C1/ <i>ura3</i> Δ	<i>Fpoah</i> integrated to LPB/ <i>ura3</i> Δ at <i>URA3</i>	This study
OA_Fpoah_C2/ <i>ura3</i> Δ	<i>Fpoah</i> integrated to LPB/ <i>ura3</i> Δ at <i>EG4</i>	This study
OA_Fpoah_C3/ <i>ura3</i> Δ	<i>Fpoah</i> integrated to LPB/ <i>ura3</i> Δ at <i>EG6</i>	This study
OA_Pcoah_PYC_C1/ <i>ura3</i> Δ	<i>PYC</i> integrated to OA_Pcoah/ <i>ura3</i> Δ at <i>LP3</i>	This study
OA_Pcoah_PYC/ <i>ura3</i> Δ	<i>PYC</i> integrated to OA_Pcoah/ <i>ura3</i> Δ at <i>EG4</i>	This study
OA_Pcoah_PYC_C3/ <i>ura3</i> Δ	<i>PYC</i> integrated to OA_Pcoah/ <i>ura3</i> Δ at <i>GD5</i>	This study
OA_Pcoah/FpOAR/ <i>ura3</i> Δ	<i>FpOAR</i> integrated to OA_Pcoah/ <i>ura3</i> Δ at <i>int5</i>	This study
OA_Pcoah/HboT1/ <i>ura3</i> Δ	<i>HboT1</i> integrated to OA_Pcoah/ <i>ura3</i> Δ at <i>int5</i>	This study
OA_Pcoah/HboT2/ <i>ura3</i> Δ	<i>HboT2</i> integrated to OA_Pcoah/ <i>ura3</i> Δ at <i>int5</i>	This study
JL01/ <i>ura3</i> Δ	<i>HboT2</i> integrated to OA_Pcoah_PYC/ <i>ura3</i> Δ at <i>int5</i>	This study
JL02/ <i>ura3</i> Δ	JL01/ <i>ura3</i> Δ with <i>gpd</i> knockout	This study
JL03/ <i>ura3</i> Δ	<i>PYK1</i> integrated to the <i>EG4</i> site in JL02/ <i>ura3</i> Δ	This study
JL02/g4285/ <i>ura3</i> Δ	<i>g4285</i> integrated to the <i>EG4</i> site in JL02/ <i>ura3</i> Δ	This study
JL04/ <i>ura3</i> Δ	JL03/ <i>ura3</i> Δ with <i>g4567</i> knockout	This study
OA_Pcoah_C1	OA_Pcoah_C1/ <i>ura3</i> Δ with <i>ura3</i> recovered	This study
OA_Pcoah_C2	OA_Pcoah_C2/ <i>ura3</i> Δ with <i>ura3</i> recovered	This study
OA_Pcoah	OA_Pcoah/ <i>ura3</i> Δ with <i>ura3</i> recovered	This study
OA_Fpoah_C1	OA_Fpoah_C1/ <i>ura3</i> Δ with <i>ura3</i> recovered	This study
OA_Fpoah_C2	OA_Fpoah_C2/ <i>ura3</i> Δ with <i>ura3</i> recovered	This study
OA_Fpoah_C3	OA_Fpoah_C3/ <i>ura3</i> Δ with <i>ura3</i> recovered	This study
OA_Pcoah_PYC_C1	OA_Pcoah_PYC_C1/ <i>ura3</i> Δ with <i>ura3</i> recovered	This study
OA_Pcoah_PYC	OA_Pcoah_PYC/ <i>ura3</i> Δ with <i>ura3</i> recovered	This study
OA_Pcoah_PYC_C3	OA_Pcoah_PYC_C3/ <i>ura3</i> Δ with <i>ura3</i> recovered	This study
OA_Pcoah/FpOAR	OA_Pcoah/FpOAR/ <i>ura3</i> Δ with <i>ura3</i> recovered	This study
OA_Pcoah/HboT1	OA_Pcoah/HboT1/ <i>ura3</i> Δ with <i>ura3</i> recovered	This study
OA_Pcoah/HboT2	OA_Pcoah/HboT2/ <i>ura3</i> Δ with <i>ura3</i> recovered	This study
JL01	JL01/ <i>ura3</i> Δ with <i>ura3</i> recovered	This study
JL02	JL02/ <i>ura3</i> Δ with <i>ura3</i> recovered	This study
JL03	JL03/ <i>ura3</i> Δ with <i>ura3</i> recovered	This study
JL02/g4285	JL02/g4285/ <i>ura3</i> Δ with <i>ura3</i> recovered	This study
JL04	JL04/ <i>ura3</i> Δ with <i>ura3</i> recovered	This study

Supplementary Table 2. Plasmids used in this study.

Plasmids	Features	Sources
pVT36b	<i>I. orientalis</i> plasmid for gene deletion and insertion	Ref. 4
pVT36b-LP3	pVT36b with insertion of spacer for integration at site <i>LP3</i>	Ref. 3
pVT36b-EG4	pVT36b with insertion of spacer for integration at site <i>EG4</i>	Ref. 3
pVT36b-EG6	pVT36b with insertion of spacer for integration at site <i>EG6</i>	Ref. 3
pVT36b-GD5	pVT36b with insertion of spacer for integration at site <i>GD5</i>	Ref. 3
pVT36b-int3	pVT36b with insertion of spacer for integration at site <i>int3</i>	Ref. 5
pVT36b-int5	pVT36b with insertion of spacer for integration at site <i>int5</i>	Ref. 5
pVT36b- <i>gpd</i>	pVT36b with insertion of spacer for deletion of <i>gpd</i>	Ref. 5
pVT36b- <i>g4567</i>	pVT36b with insertion of spacer for deletion of <i>g4567</i> with 50 bp upstream and downstream homology arm targeting to <i>g4567</i> gene	
pRS416-SA-site1	pRS416 harboring reductive TCA pathway flanked with homology arms for integration to site 1	Ref. 5
pVT101a	<i>I. orientalis</i> plasmid for gene deletion and insertion	Ref. 5
p_Pcoah	pVT101a with insertion of <i>TDH3p-Pcoah-3376t</i>	This study
p_Fpoah	pVT101a with insertion of <i>TDH3p-Fpoah-3376t</i>	This study
p_FpOAR	pVT101a with insertion of <i>TDH3p-FpOAR-3376t</i>	This study
p_HboT1	pVT101a with insertion of <i>TDH3p-Hbot1-3376t</i>	This study
p_HboT2	pVT101a with insertion of <i>TDH3p-Hbot2-3376t</i>	This study
p_PYK1	pVT101a with insertion of <i>TDH3p-PYK1-3376t</i>	This study
p_g4285	pVT101a with insertion of <i>TDH3p-g4285-3376t</i>	This study

Supplementary Table 3. Comparison of XRD peak positions (d-spacing) and relative intensities for $\text{Nd}_2(\text{C}_2\text{O}_4)_3 \cdot 10\text{H}_2\text{O}$ synthesized using Sigma oxalic acid, bio-oxalic acid, and the standard reference PDF-00-020-0764. The d-spacing and intensity values for the samples synthesized with Sigma oxalic acid and bio-oxalic acid are compared to evaluate the structural similarity of the products.

Nd ₂ (C ₂ O ₄) ₃ ·10H ₂ O standard, PDF-00-020-0764			Nd ₂ (C ₂ O ₄) ₃ ·10H ₂ O, synthesized using Sigma Oxalic acid in this study			Nd ₂ (C ₂ O ₄) ₃ ·10H ₂ O, synthesized using Bio Oxalic acid in this study		
hkl	d-Spacing, Å	Intensity	hkl	d-Spacing, Å	Intensity	hkl	d-Spacing, Å	Intensity
(100)	10.2	80	(100)	10.2	70	(100)	10.2	46
(110)	6.97	50	(110)	7.0	18	(110)	7.0	17
(011)	6.66	100	(011)	6.70	19	(011)	6.70	14
($\bar{1}11$)	6.51	100	($\bar{1}11$)	6.57	21	($\bar{1}11$)	6.56	16
($\bar{1}02$)	5.11	80	($\bar{1}02$)	5.10	22	($\bar{1}02$)	5.10	15
(111)	4.95	100	(111)	4.91	50	(111)	4.97	13
($\bar{2}11$)	4.77	100	($\bar{2}11$)	4.82	100	($\bar{2}11$)	4.82	100
($\bar{3}02$)	3.52	80	($\bar{3}02$)	3.51	7	($\bar{3}02$)	3.51	5
(311)	2.725	100	(311)	2.727	7	(311)	2.725	6
(023)	2.614	100	(023)	2.612	7	(023)	2.611	5

Supplementary Table 4. Representative studies on oxalic acid production by various microorganisms.

Microbe	Feedstock/substrate	pH and Time (Days)	Fermentation mode	Titer	References
<i>Aspergillus niger</i> ATCC 9029	Sucrose, Lactose permeate, Molasses	6; 7 days	Shake flask	15.70 g/L; 16.49 g/L; 1.80 g/L	6
	Sucrose	6; 10 days	Fed-batch bioreactor	38.44 g/L	
<i>Aspergillus niger</i> 1120	Sucrose	pH shift; 8 days	Stirred tank	33.8 g/L	7
<i>Aspergillus niger</i> W78C	Sucrose	6; 10 days	Stirred tank reactor	64.3 g/L	8
<i>Aspergillus niger</i> 1120	Milk whey	6; 13 days	Stirred tank reactor	41.4 g/L	9
<i>Aspergillus niger</i> XP	Crude rapeseed oil	4.0; 7 days	Shake flask	64.2 g/L	10
	Fatty acid	4.0; 7 days	Shake flask	68.1 g/L	
<i>Aspergillus niger</i> NCIM 548	Glucose, sucrose, and lactose	6-7; 7-9 days per cycle	Immobilized cells (PUF), submerged	20.6 g/L	11
<i>Aspergillus niger</i>	Whey permeate	6.0; 7 days	Shake flask	1-7 mM	12
<i>Aspergillus niger</i> XP	Biodiesel-derived waste composed of glycerol	pH 4.5; 7 days	Stirred tank reactor	48.9 g/L	13
<i>Aspergillus niger</i> LFMB 1	Crude glycerol	4.8-6.0; 10.9 days	Shake flask	21.5 g/L	14
<i>Aspergillus niger</i>	Cashew apple juice (CAJ)	5.4; 7.31 days	Shake flask	122.68 g/L	15
<i>Trichoderma asperellum</i> MG323528	Corn stover	3.7; 10 days	Solid-state fermentation	14.5 g/L	16
<i>Penicillium oxalicum</i>	Potato dextrose	4.5; 5 days	Shake flask	0.17 g/L	17
<i>Penicillium purpurogenum</i> MN3	Rice husk	Not controlled; 7 days	Solid-state fermentation	43.20 mg/g	18
<i>Issatchenkia orientalis</i> SD108	Glucose	4.0; 10 days	Fed-batch bioreactor	39.53 g/L	This study

Supplementary Table 5. Economic assumptions for techno-economic analysis (TEA).

				Assumption	Reference
TEA parameters					
Plant lifetime (duration)				30 years (2016-2046)	This study
Depreciation method				MACRS7	19
Steam power depreciation method				MACRS20	19
Operating days				200 days (baseline value)	20
Internal rate of return (IRR)				10% (baseline value)	19
Combined federal and state income tax rate				21%	19
Construction capital investment fractions				8% for 1 st year, 60% for 2 nd year, 32% for 3 rd year	19
Startup time				3 months	19
Finance interest				8% of capital cost financing	19
Finance years				10 years	19
Finance fraction				40% of capital cost	19
Cost components ²					
Fixed capital investment (FCI)	Total depreciable capital (TDC)	Direct permanent investment (DPI)	Warehouse	4% of ISBL DPI ¹	19
			Site development	9% of ISBL DPI ¹	19
			Additional piping	4.5% of ISBL DPI ¹	19
		Depreciable indirect costs	Proratable costs	10% of DPI	19
			Field expenses	10% of DPI	19
			Construction fee	20% of DPI	19
			Contingency	10% of DPI	19
	Non-depreciable indirect costs		Other indirect costs	10% of DPI	19
Working capital				5% of FCI	19
Fixed operating cost (FOC)	Labor cost and labor burden			6.4 million USD (baseline value), 90% labor burden	19
	Property insurance			0.7% of FCI	19
	Maintenance			3% of ISBL DPI ¹	19

¹ ISBL DPI represents the total installed costs (i.e., direct permanent investment) of ISBL (inside battery limits) units. In a biomanufacturing design, ISBL units are the core process operations, such as batch fermentor, cell mass separator, and acidulation reactor in this work. By contrast, OSBL (outside battery limits) units include supporting and auxiliary facilities, such as storage units, wastewater treatment system, and turbogenerator system.

² Installed cost of ISBL and OSBL units and annual operating cost (AOC) are not presented in this table, and they are calculated from the system's simulation results. Specifically, installed costs are based on the unit design results and cost correlations; while the AOC accounts for material inputs and utilities.

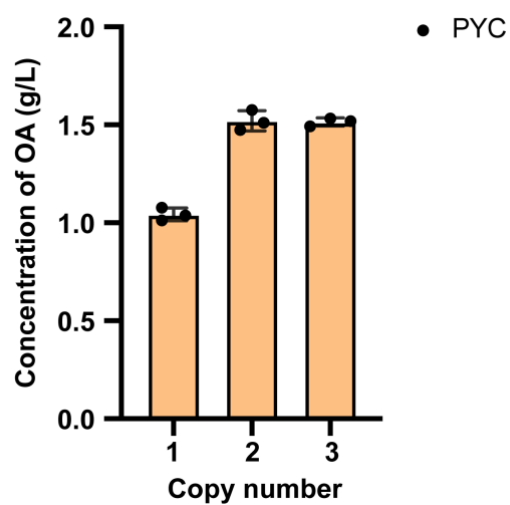
Supplementary Table 6. Baseline material and utility prices.

Materials / utility	Unit	Value	Reference
<i>Input</i>			
Sugarcane	\$·wet kg ⁻¹	0.04	20
Lime	\$·kg ⁻¹	0.08	19
Calcium hydroxide	\$·kg ⁻¹	0.29	19
CO ₂ for fermentation	\$·kg ⁻¹	0.29	21
Corn steep liquor (CSL)	\$·kg ⁻¹	0.08	22
Magnesium chloride (MgCl ₂)	\$·kg ⁻¹	0.38	22
Diammonium sulfate [(NH ₄) ₂ SO ₄]	\$·kg ⁻¹	0.2	22
Monoethanolamine (MEA)	\$·kg ⁻¹	1.52	23
Sulfuric acid (H ₂ SO ₄)	\$·kg ⁻¹	0.1	22
Sodium hypochlorite (NaOCl)	\$·kg ⁻¹	0.14	24
Citric acid	\$·kg ⁻¹	0.34	24
Bisulfite	\$·kg ⁻¹	0.08	24
Flue gas desulfurization (FGD) lime	\$·kg ⁻¹	0.29	19
Cooling tower chemicals	\$·kg ⁻¹	3.0	22
Boiler chemicals	\$·kg ⁻¹	5.0	22
Makeup water	\$·kg ⁻¹	0.0003	25
<i>Output</i>			
Gypsum (disposal)	\$·kg ⁻¹	0	This study
Electricity	\$·kWh ⁻¹	0.07	26

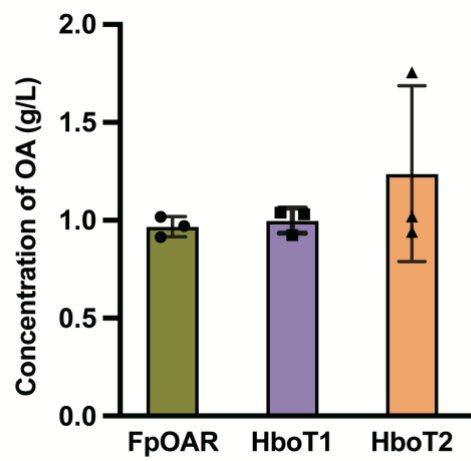
Supplementary Table 7. Distributions of results from the uncertainty analysis for MSP, CI, and FEC at alternative numbers of Monte Carlo simulations (1000, 3000, 6000, and 10,000). This table was used to determine an appropriate number of samples for statistical robustness. Source data for this table are available in the Supplementary Data 6.

Percentile	MSP [\$·kg ⁻¹]				CI [kg CO ₂ e·kg ⁻¹]				FEC [MJ·kg ⁻¹]			
	1000	3000	6000	10000	1000	3000	6000	10000	1000	3000	6000	10000
5 th	1.52	1.51	1.51	1.51	-2.86	-2.80	-2.83	-2.82	-63.22	-62.18	-62.31	-62.50
10 th	1.56	1.57	1.57	1.57	-2.50	-2.49	-2.47	-2.47	-57.95	-58.10	-58.02	-57.95
25 th	1.67	1.68	1.67	1.67	-1.85	-1.87	-1.88	-1.86	-50.12	-50.05	-50.22	-50.01
50 th	1.80	1.80	1.80	1.79	-1.27	-1.23	-1.23	-1.22	-42.17	-41.89	-41.93	-41.92
75 th	1.93	1.92	1.93	1.93	-0.64	-0.68	-0.67	-0.67	-34.45	-34.87	-34.65	-34.71
90 th	2.06	2.05	2.05	2.06	-0.23	-0.20	-0.23	-0.23	-29.12	-29.08	-29.18	-29.29
95 th	2.14	2.13	2.12	2.14	-0.01	0.00	0.02	0.02	-25.93	-26.23	-26.24	-26.29

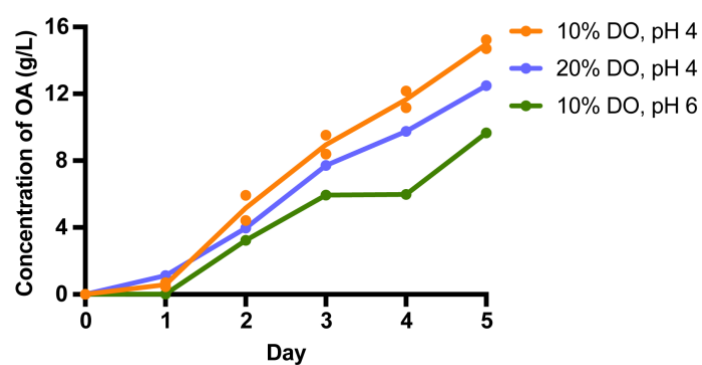
Supplementary Figures



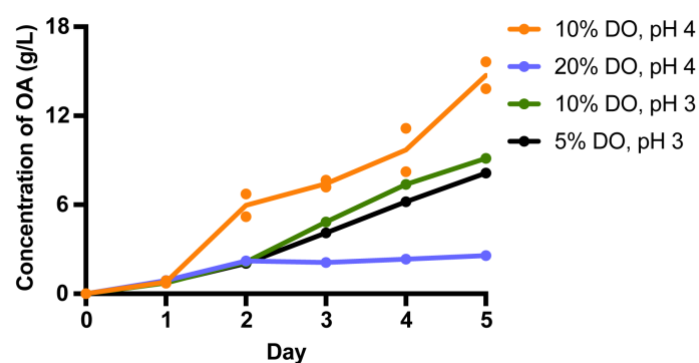
Supplementary Figure 1. Oxalic acid production in strains overexpressing *PYC* with varying copy numbers. Strains were cultured in SC-URA medium with 50 g/L glucose under shake flask fermentation for 5 days, and the concentration of oxalic acid were measured. Error bars represent the mean $\pm$ SD (n = 3, biologically independent samples). Source data for this figure are available in the Source Data file.



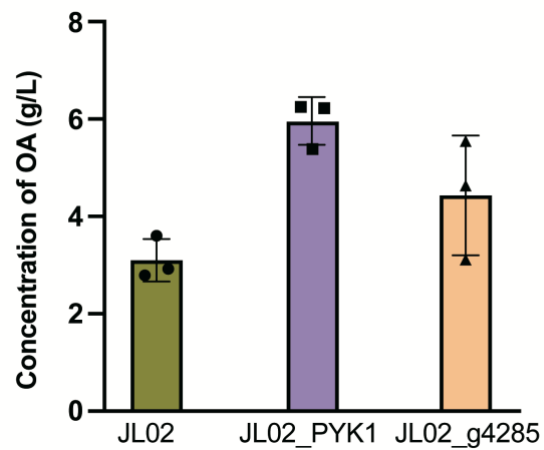
Supplementary Figure 2. Oxalic acid production in strains overexpressing transporter. Strains were cultured in SC-URA medium with 50 g/L glucose under shake flask fermentation for 5 days, and the concentrations of oxalic acid were measured. Error bars represent the mean $\pm$ SD ($n = 3$, biologically independent samples). Source data for this figure are available in the Source Data file.



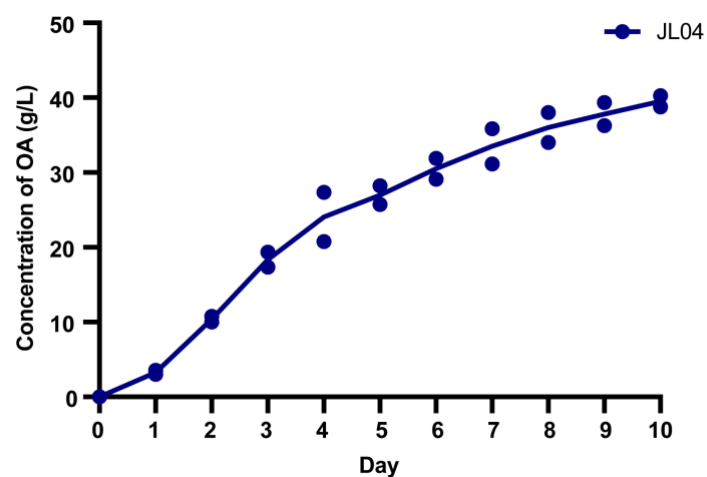
Supplementary Figure 3. Oxalic acid production in strain JL01 under different fermentation conditions. Strains were cultured in SC-URA medium using fed-batch fermentation for 5 days, and oxalic acid concentrations were measured. For pH 4.0 and 10% DO, two biologically independent replicates were performed ($n = 2$), while for the other two conditions, a single experiment was conducted ($n = 1$). Source data for this figure are available in the Source Data file.



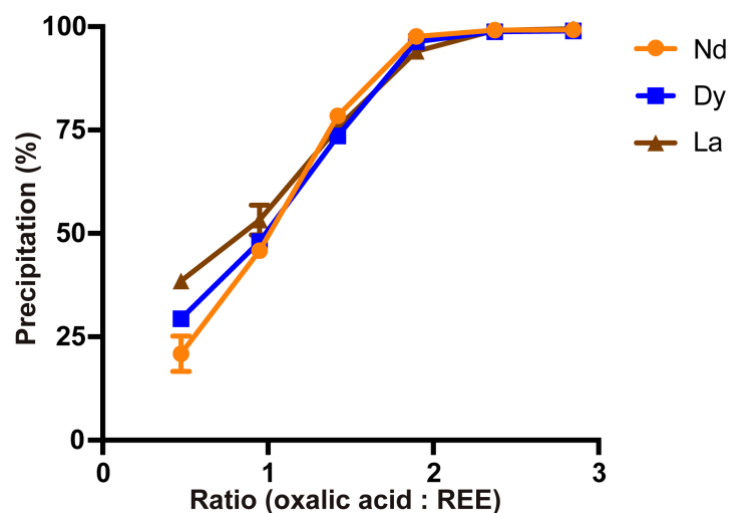
Supplementary Figure 4. Oxalic acid production in strain JL02 under different fermentation conditions. Strains were cultured in SC-URA medium using fed-batch fermentation for 5 days, and oxalic acid concentrations were measured. For pH 4.0 and 10% DO, two biologically independent replicates were performed ($n = 2$), while for the other three conditions, a single experiment was conducted ($n = 1$). Source data for this figure are available in the Source Data file.



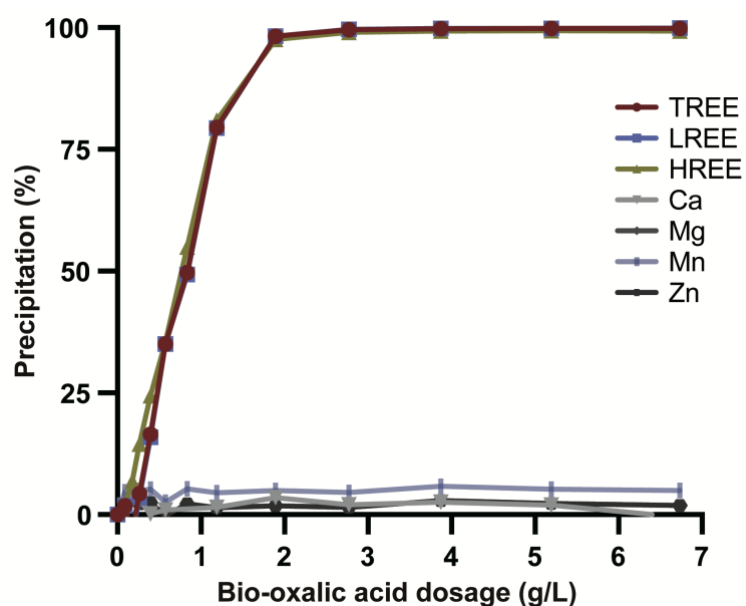
Supplementary Figure 5. Oxalic acid production in the strain JL02 overexpressing pyruvate kinase candidates. Strains were cultured in SC-URA medium with 50 g/L glucose under shake flask fermentation for 5 days, and the concentrations of oxalic acid were measured. The medium was buffered with citric acid- NaH_2PO_4 to maintain pH at 4.0. Error bars represent the mean $\pm$ SD ($n = 3$, biologically independent samples). Source data for this figure are available in the Source Data file.



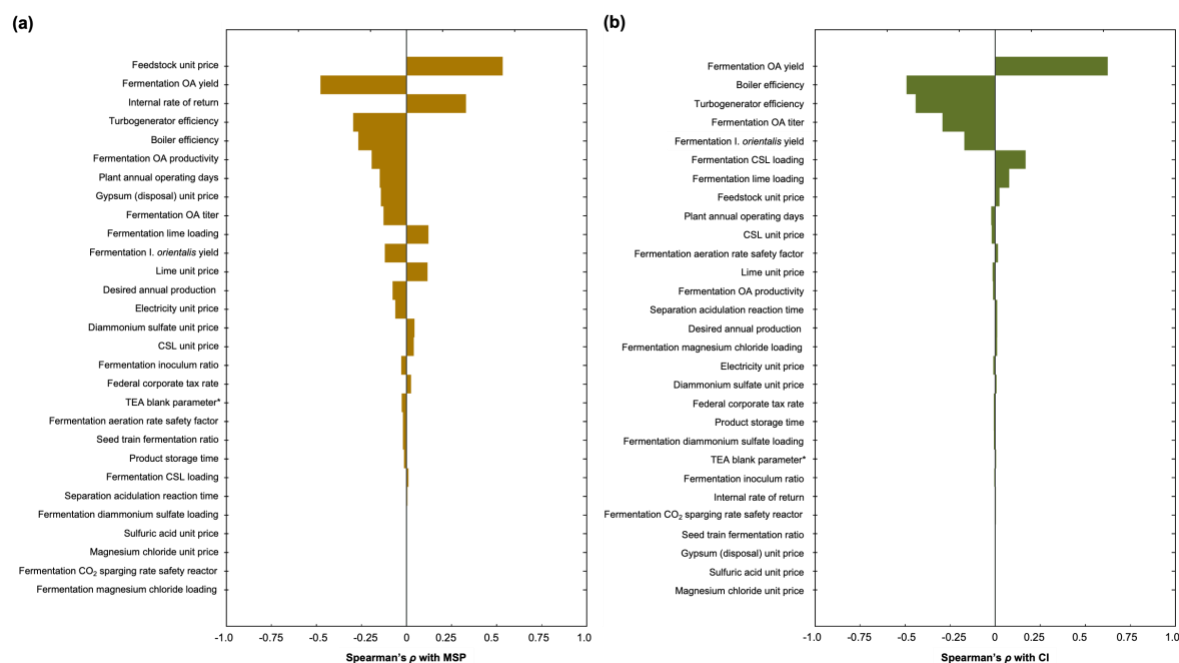
Supplementary Figure 6. Fed-batch fermentation of engineered strain JL04 in a 100 mL bioreactor for 10 days. Each point represents an independent biological replicate ($n = 2$). Source data for this figure are available in the Source Data file.



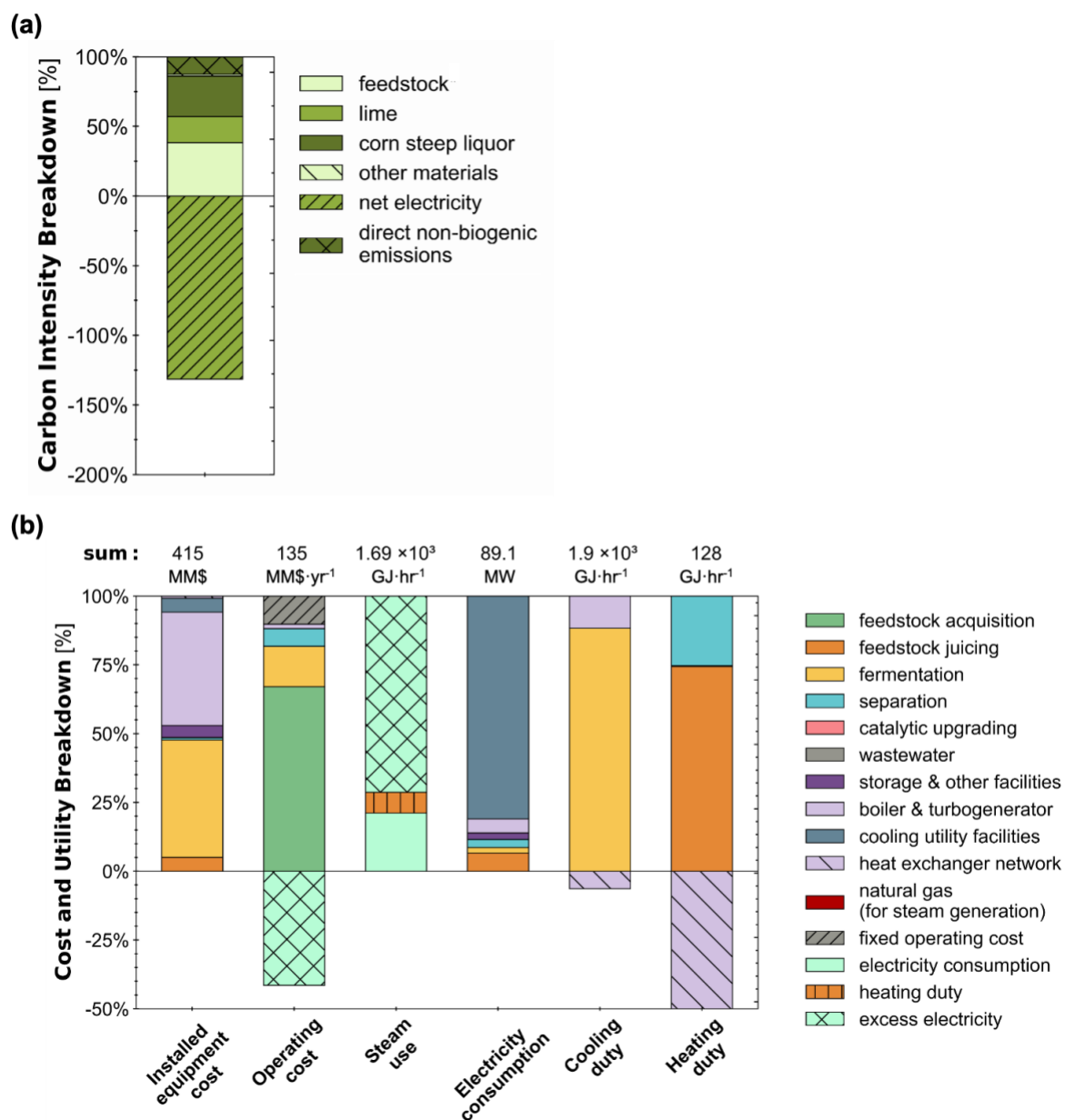
Supplementary Figure 7. Bio-oxalic acid enables efficient precipitation of Nd, La, and Dy. Bio-oxalic acid (42.7 g/L) was mixed with 1 mM REE solutions (NdCl_3 , LaCl_3 , DyCl_3) at varying oxalic acid:REE molar ratios (0 to 3). Residual REE concentrations were measured using a xylenol orange colorimetric assay. Error bars represent the mean $\pm$ SD ($n = 3$). Source data for this figure are available in the Source Data file.



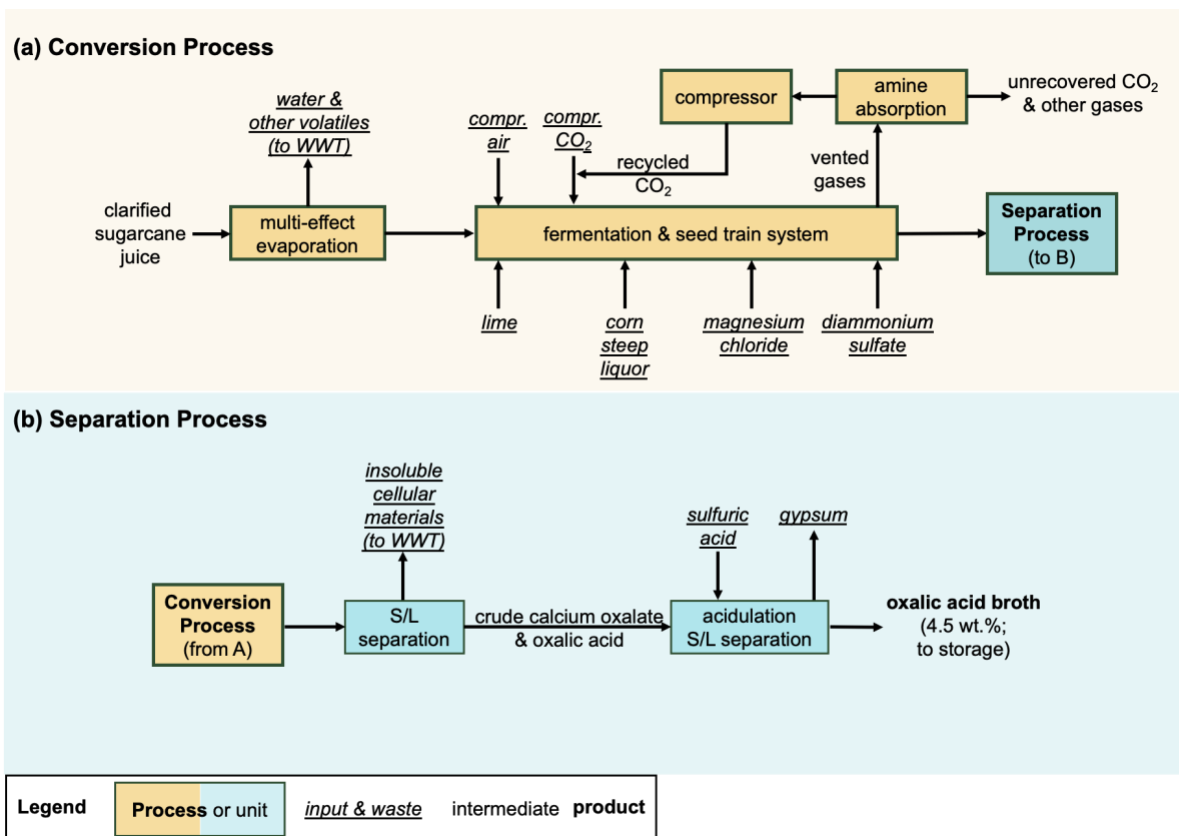
Supplementary Figure 8. Effective precipitation of total REEs (TREE) from allanite leachate using bio-oxalic acid. Allanite leachate was treated with bio-oxalic acid (22.0 g/L feedstock) to evaluate the recovery of TREE. TREE refer to the combination total of light REEs (LREE) and heavy REEs (HREE). The LREEs analyzed include Sc, La, Ce, Pr, Nd, and Sm; the HREEs include Y, Eu, Gd, Tb, Dy, Ho, Er, Tm, Yb, and Lu. Source data for this figure are available in the Source Data file.



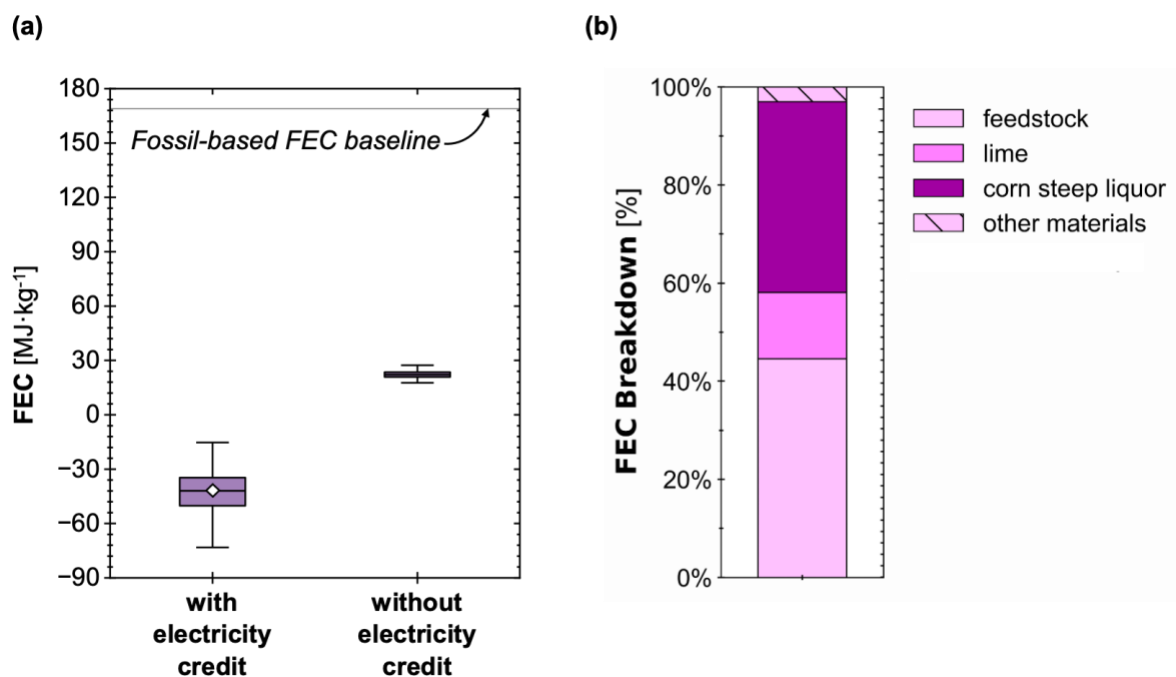
Supplementary Figure 9. Sensitivity analysis results from 6,000 Monte Carlo simulations. Spearman's ρ values between input parameters and (a) MSP and (b) CI. The asterisked entry "TEA blank parameter" is a non-functional placeholder, included solely to represent the Spearman's ρ threshold used to distinguish statistically significant parameters, i.e., those with Spearman's ρ above this threshold. Source data for this figure are available in the Source Data file.



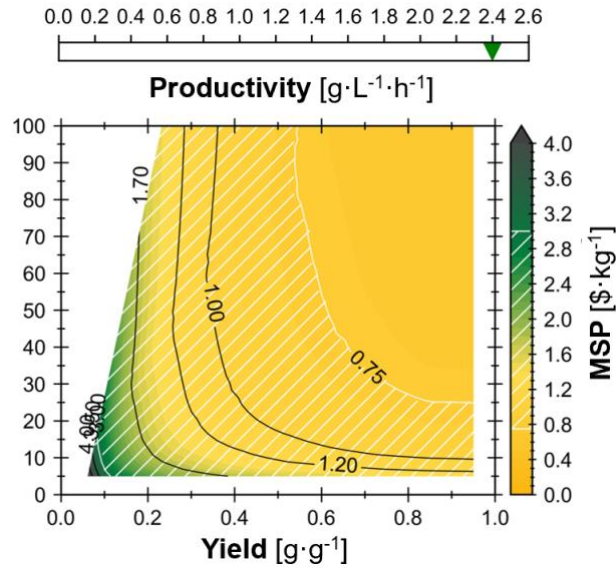
Supplementary Figure 10. Breakdowns (stacked bar charts) of (a) carbon intensity and (b) cost and utility breakdown at baseline. Source data for this figure are available in the Source Data file.



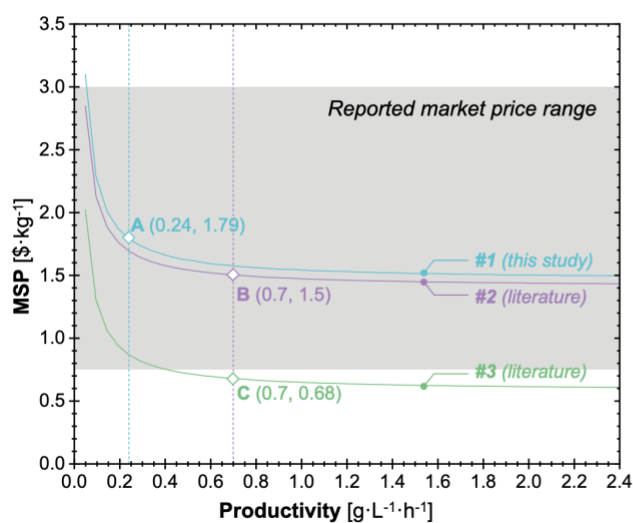
Supplementary Figure 11. Simplified block flow diagram of the designed biorefinery. (a) Conversion process. (b) Separation process. Acronyms denote compressed gas (compr.), wastewater treatment (WWT), and solid-liquid separation (S/L separation).



Supplementary Figure 12. Fuel energy consumption (FEC; cradle-to-gate) performance. (a) Uncertainties (box and whisker plots) of FEC with and without electricity credit. Whiskers, boxes, and middle lines represent 5th/95th, 25th/75th, and 50th percentiles from 6,000 Monte Carlo simulations. Diamond marker represents the baseline result, and the line represents the fossil-based FEC baseline of 169 $\text{MJ}\cdot\text{kg}^{-1}$ ²⁷. (b) FEC breakdown (stacked bar chart). Only positive contributions are presented for transparency (net electricity offsets approximately 300% of total positive contributions). Source data for this figure are available in the Source Data file.

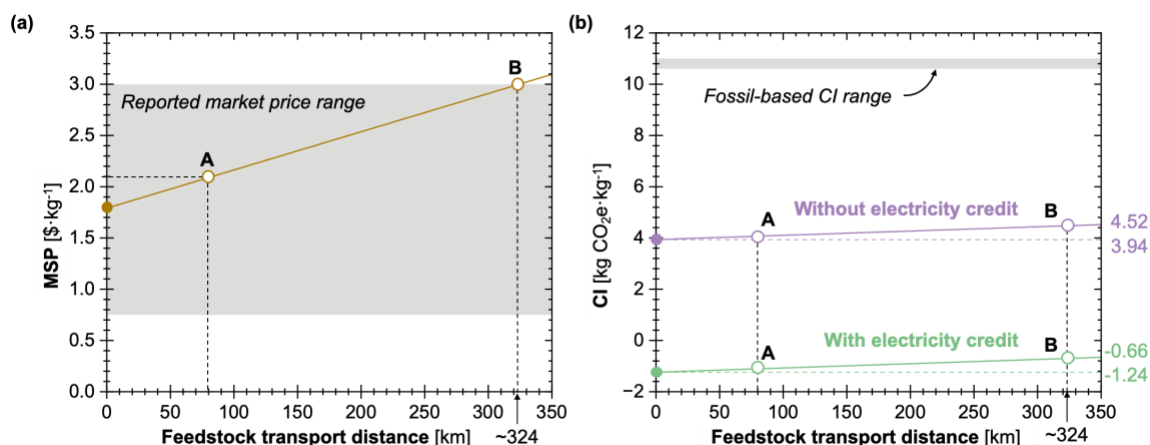


Supplementary Figure 13. MSP across 2,500 fermentation yield-titer combinations at the 10-fold improvement ($2.4 \text{ g} \cdot \text{L}^{-1} \cdot \text{h}^{-1}$) of baseline productivity. This figure complements Figure 4c in the main text but is provided here to avoid redundancy, as the results are similar to those observed at 5-fold productivity improvement. Source data for this figure are available in the Source Data file.



Scenario	#1	#2	#3
Titer [g·L ⁻¹]	28.23	122.72	122.72
Yield [g·g ⁻¹]	0.2	0.818	0.818
Feedstock price adjusted based on sugar content	Yes	Yes	No

Supplementary Figure 14. Minimum selling price (MSP) of bio-oxalic acid as a function of fermentation productivity (ranging from 0.2x to 10x baseline of this study) across three scenarios: (1) this study, (2) *Aspergillus niger* system^{I5} with the highest reported titer of 122.72 g·L⁻¹ and yield of 0.818 g·g⁻¹ (the baseline productivity of 0.7 g·L⁻¹·h⁻¹) with sugar-adjusted feedstock price, and (3) *Aspergillus niger* system *r*^{I5} without feedstock price adjustment. Diamonds represent the baseline performance for each scenario. Source data for this figure are available in the Source Data file.



Supplementary Figure 15. (a) Minimum selling price (MSP) and (b) carbon intensity (CI) of bio-oxalic acid production with feedstock transport impacts included, while all other parameters are held at baseline. Colored points indicate values with no transport (as discussed in the main text). Points A represent a standard sourcing radius of 50 miles (80 km), increasing MSP by 16.7% to approximately $\text{\$2.1}\cdot\text{kg}^{-1}$ and CI by 2.9% to approximately $4.1 \text{ kg CO}_2\text{e}\cdot\text{kg}^{-1}$ (without electricity credit). Points B represent the largest transport distance (approximately 324 km) that remains market competitive. Feedstock transport exhibits a minor impact on CI. Source data for this figure are available in the Source Data file.

Supplementary References

- 1 Gietz, R. D., Schiestl, R. H., Willems, A. R. & Woods, R. A. Studies on the transformation of intact yeast cells by the LiAc/SS-DNA/PEG procedure. *Yeast* **11**, 355-360 (1995).
- 2 Xiao, H., Shao, Z., Jiang, Y., Dole, S. & Zhao, H. Exploiting *Issatchenkia orientalis* SD108 for succinic acid production. *Microb. cell fact.* **13**, 1-11 (2014).
- 3 Fatma, Z., Tan, S.-I., Boob, A. G. & Zhao, H. A landing pad system for multicopy gene integration in *Issatchenkia orientalis*. *Metab. Eng.* **78**, 200-208 (2023).
- 4 Tran, V. G., Cao, M., Fatma, Z., Song, X. & Zhao, H. Development of a CRISPR/Cas9-based tool for gene deletion in *Issatchenkia orientalis*. *Mosphere* **4**, 10.1128/msphere. 00345-00319 (2019).
- 5 Tran, V. G. *et al.* An end-to-end pipeline for succinic acid production at an industrially relevant scale using *Issatchenkia orientalis*. *Nat. Commun.* **14**, 6152 (2023).
- 6 Strasser, H., Burgstaller, W. & Schinner, F. High-yield production of oxalic acid for metal leaching processes by *Aspergillus niger*. *FEMS Microbiol. Lett.* **119**, 365-370 (1994).
- 7 Cameselle, C., Bohlmann, J., Nunez, M. & Lema, J. Oxalic acid production by *Aspergillus niger*: Part I: Influence of sucrose and milk whey as carbon source. *Bioprocess Eng.* **19**, 247-252 (1998).
- 8 Walaszczyk, E., Podgórski, W., Janczar-Smuga, M. & Dymarska, E. Effect of medium pH on chemical selectivity of oxalic acid biosynthesis by *Aspergillus niger* W78C in submerged batch cultures with sucrose as a carbon source. *Chem. Pap.* **72**, 1089-1093 (2018).
- 9 Bohlmann, J., Cameselle, C., Nunez, M. & Lema, J. Oxalic acid production by *Aspergillus niger*: Part II: Optimisation of fermentation with milk whey as carbon source. *Bioprocess Eng.* **19**, 337-342 (1998).
- 10 Rymowicz, W. & Lenart, D. Oxalic acid production from lipids by a mutant of *Aspergillus niger* at different pH. *Biotechnol. Lett.* **25**, 955-958 (2003).
- 11 Mandal, S. K. & Banerjee, P. C. Submerged production of oxalic acid from glucose by immobilized *Aspergillus niger*. *Process Biochem.* **40**, 1605-1610 (2005).
- 12 Brown, K., Harrison, J. & Bowers, K. Production of oxalic acid from *Aspergillus niger* and whey permeate. *Water Air Soil Pollut.* **229**, 1-10 (2018).
- 13 Musiał, I., Cibis, E. & Rymowicz, W. Designing a process of kaolin bleaching in an oxalic acid enriched medium by *Aspergillus niger* cultivated on biodiesel-derived waste composed of glycerol and fatty acids. *Appl. Clay Sci.* **52**, 277-284 (2011).
- 14 André, A. *et al.* Biotechnological conversions of bio-diesel derived waste glycerol into added-value compounds by higher fungi: production of biomass, single cell oil and oxalic acid. *Ind. Crops Prod.* **31**, 407-416 (2010).
- 15 Betiku, E., Emeko, H. A. & Solomon, B. O. Fermentation parameter optimization of microbial oxalic acid production from cashew apple juice. *Heliyon* **2** (2016).
- 16 Al-Askar, A., Saber, W., Ghoneem, K. & Rashad, Y. Oxalic acid as the main molecule produced by *Trichoderma asperellum* MG323528 fermented on corn stover based medium. *Biotechnology* **17**, 95-103 (2018).
- 17 Zhang, B. *et al.* In vitro starch digestion: Mechanisms and kinetic models. *Starch Struct. Funct. Appl. Foods*, 151-167 (2020).
- 18 El-Naggar, N. E.-A. & El-Hersh, M. S. Organic acids associated with saccharification of cellulosic wastes during solid-state fermentation. *J. Microbiol.* **49**, 58-65 (2011).
- 19 Davis, R. E. *et al.* Process design and economics for the conversion of lignocellulosic biomass to hydrocarbon fuels and coproducts: 2018 biochemical design case update; biochemical deconstruction and conversion of biomass to fuels and products via integrated biorefinery pathways. <https://www.nrel.gov/docs/fy19osti/71949.pdf> (2018).
- 20 Huang, H., Long, S. & Singh, V. Techno-economic analysis of biodiesel and ethanol co-production from lipid-producing sugarcane. *Biofuels Bioprod. Bioref.* **10**, 299-315 (2016).
- 21 Chemanalyst. Liquid Carbon Dioxide Price Trend and Forecast. <https://www.chemanalyst.com/Pricing-data/liquid-carbon-dioxide-1090> (2025).
- 22 ICIS. Chemical prices, forecasts and market insights. <https://www.icis.com/explore/commodities/chemicals> (2025).

- 23 Intratec. Monoethanolamine prices worldwide. <https://www.intratec.us/solutions/primary-commodity-prices/commodity/monoethanolamine-prices> (2025).
- 24 Li, Y. *et al.* Design of a high-rate wastewater treatment process for energy and water recovery at biorefineries. *ACS Sustain. Chem. Eng.* **11**, 3861-3872 (2023).
- 25 U.S. Environmental Protection Agency. Data and information used by WaterSense. <https://www.epa.gov/watersense/data-and-information-used-watersense> (2025).
- 26 U.S. Energy Information Administration. Electricity sales, revenue, and average price. https://www.eia.gov/electricity/sales_revenue_price (2025).
- 27 Ecoinvent. ecoinvent cutoff system model, version 3.8. <https://ecoquery.ecoinvent.org/3.8/cutoff> (2025).