Supporting Information:

Transposase-Mediated Chromosomal Integration of Exogenous Genes in *Acidithiobacillus ferrooxidans*

Yuta Inaba, Indrani Banerjee, Timothy Kernan, and Scott Banta*

Department of Chemical Engineering, Columbia University, New York, NY, USA

Table S1. DNA and amino acid sequences for codon-optimized version of the hyperactive mini-Tn5 transposase, Tnp, expressed in *A. ferrooxidans.*

Tnp DNA	ATGATCACATCCGCACTTCACCGAGCAGCAGATTGGGCAAAAAGCGTTTTTAGTTCAGCCGCGCTGGGTGATCCCCGCCGTACCGCCCGTCTGGTGAACGTCGCCGCACAGCTCGCCAAGTACTCGGGCAAATCCATTACAATTTCCAGCGAAGGCTCAAAGGCCATGCAAGAGGGCGCCTATCGCTTCATCCGCAATCCGAATGTCAGCGCCGAGGCCATTCGGAAGGCCGGTGCGATGCAGACCGTCAAATTGGCCCAAGAATTCCCCGAGCTGTTGGCGATCGAAGATACGACGAGCCTTTCATATCGTCATCAAGTGGCGGAAGAACTGGGAAAACTGGGTTCCATCCAGGACAAGTCTCGCGGTTGGTGGGTACACTCGGTGCTGCTCCTGGAAGCTACCACCTTCCGTACCGTGGGCTTGTTGCACCAAGAGTGGTGGATGCGCCCGGATGATCCGGCGGACGCGGATGAGAAAGAATCGGGGAAATGGCTGGCTGCTGCCGCGACCAGCCGGCTGCGGATGGGTAGTATGATGAGCAACGTCATAGCCGTCTGTGACCGCGAAGCCGACATCCATGCCTATCTGCAGGACAAGCTCGCGCATAATGAACGTTTTGTGGTTCGCTCCAAGCACCCGCGCAAAGACGTTGAGTCGGGGCTGTACCTGTACGACCACCTCAAAAACCAGCCTGAGTTGGGCGGATACCAGATTTCCATTGGCCAGAAAGGCGTCGTTGACAAGCGGGGTAAGCGCAAGAATCGCCCCGCCCGGAAAGCCTCTCTCTCCCTGCGGAGTGGGCGTATCACCCTCAAGCAAGGTAATATCACGCTCAACGCCGTACTGGCGGAAGAGATCAACCCCCCCAAAGGTGAGACTCCATTGAAGTGGCTGCTGCTGACCAGCGAGCCCGTGGAGTCTCTGGCCCAAGCGCTGCGCGTGATTGACATTTATACCCACCGGTGGCGAATCGAAGAATTTCATAAAGCCTGGAAAACGGGTGCCGGAGCGGAGCGTCAGCGGATGGAAGAACCTGATAATCTCGAACGCATGGTCAGTATCCTCAGCTTTGTGGCCGTGCGCCTGCTGCAACTGCGTGAGTCGTTCACCCCACCGCAGGCCCTGCGAGCTCAAGGGCTCTTGAAAGAAGCCGAACATGTCGAGTCCCAAAGCGCCGAAACGGTACTGACGCCGGATGAGTGCCAGTTGTTGGGATATCTCGATAAGGGCAAGCGCAAGCGGAAAGAAAAGGCCGGCTCCCTTCAGTGGGCCTACATGGCGATCGCGCGTCTGGGGGGATTCATGGATAGTAAGCGGACTGGGATCGCTTCCTGGGGCGCGCTGTGGGAAGGCTGGGAAGCCCTCCAGAGCAAACTCGATGGGTTCCTGGCAGCAAAAGATCTGATGGCGCAGGGTATCAAGATTTGA
Tnp AA	MITSALHRAADWAKSVFSSAALGDPRRTARLVNVAAQLAKYSGKSITISSEGSKAMQEGAYRFIRNPNVSAEAIRKAGAMQTVKLAQEFPELLAIEDTTSLSYRHQVAEELGKLGSIQDKSRGWWVHSVLLLEATTFRTVGLLHQEWWMRPDDPADADEKESGKWLAAAATSRLRMGSMMSNVIAVCDREADIHAYLQDKLAHNERFVVRSKHPRKDVESGLYLYDHLKNQPELGGYQISIGQKGVVDKRGKRKNRPARKASLSLRSGRITLKQGNITLNAVLAEEINPPKGETPLKWLLLTSEPVESLAQALRVIDIYTHRWRIEEFHKAWKTGAGAERQRMEEPDNLERMVSILSFVAVRLLQLRESFTPPQALRAQGLLKEAEHVESQSAETVLTPDECQLLGYLDKGKRKRKEKAGSLQWAYMAIARLGGFMDSKRTGIASWGALWEGWEALQSKLDGFLAAKDLMAQGIKI

FIG. S1. The Clustal Omega sequence alignment for the transposase sequences used in pBAM1 and pBAM2.

```
pBAM1           ATGATTACCAGTGCACTGCATCGTGCGGCGGATTGGGCGAAAAGCGTGTTTTCTAGTGCT 60
pBAM2           ATGATCACATCCGCACTTCACCGAGCAGCAGATTGGGCAAAAAGCGTTTTTAGTTCAGCC 60
                ***** **    ***** ** ** ** ** ******** ******** ***  *   **

pBAM1           GCGCTGGGTGATCCGCGTCGTACCGCGCGTCTGGTGAATGTTGCGGCGCAACTGGCCAAA 120
pBAM2           GCGCTGGGTGATCCCCGCCGTACCGCCCGTCTGGTGAACGTCGCCGCACAGCTCGCCAAG 120
                ************** ** ******** *********** ** ** ** ** ** *****

pBAM1           TATAGCGGCAAAAGCATTACCATTAGCAGCGAAGGCAGCAAAGCCATGCAGGAAGGCGCG 180
pBAM2           TACTCGGGCAAATCCATTACAATTTCCAGCGAAGGCTCAAAGGCCATGCAAGAGGGCGCC 180
                **    ******  ****** ***  **********   ** ******** ** *****

pBAM1           TATCGTTTTATTCGTAATCCGAACGTGAGCGCGGAAGCGATTCGTAAAGCGGGTGCCATG 240
pBAM2           TATCGCTTCATCCGCAATCCGAATGTCAGCGCCGAGGCCATTCGGAAGGCCGGTGCGATG 240
                ***** ** ** ** ******** ** ***** ** ** ***** ** ** ***** ***

pBAM1           CAGACCGTGAAACTGGCCCAGGAATTTCCGGAACTGCTGGCAATTGAAGATACCACCTCT 300
pBAM2           CAGACCGTCAAATTGGCCCAAGAATTCCCCGAGCTGTTGGCGATCGAAGATACGACGAGC 300
                ******** *** ******* ***** ** ** *** **** ** ******** **

pBAM1           CTGAGCTATCGTCATCAGGTGGCGGAAGAACTGGGCAAACTGGGTAGCATTCAGGATAAA 360
pBAM2           CTTTCATATCGTCATCAAGTGGCGGAAGAACTGGGAAAACTGGGTTCCATCCAGGACAAG 360
                **    *********** ***************** *********  *** ***** **

pBAM1           AGCCGTGGTTGGTGGGTGCATAGCGTGCTGCTGCTGGAAGCGACCACCTTTCGTACCGTG 420
pBAM2           TCTCGCGGTTGGTGGGTACACTCGGTGCTGCTCCTGGAAGCTACCACCTTCCGTACCGTG 420
                   ** *********** **    ******** ******** ******** *********

pBAM1           GGCCTGCTGCATCAAGAATGGTGGATGCGTCCGGATGATCCGGCGGATGCGGATGAAAAA 480
pBAM2           GGCTTGTTGCACCAAGAGTGGTGGATGCGCCCGGATGATCCGGCGGACGCGGATGAGAAA 480
                *** ** **** ***** *********** **************** ******** ***

pBAM1           GAAAGCGGCAAATGGCTGGCCGCTGCTGCAACTTCGCGTCTGAGAATGGGCAGCATGATG 540
pBAM2           GAATCGGGGAAATGGCTGGCTGCTGCCGCGACCAGCCGGCTGCGGATGGGTAGTATGATG 540
                ***   ** *********** ***** ** **    ** *** * ***** ** ******

pBAM1           AGCAACGTGATTGCGGTGTGCGATCGTGAAGCGGATATTCATGCGTATCTGCAAGATAAA 600
pBAM2           AGCAACGTCATAGCCGTCTGTGACCGCGAAGCCGACATCCATGCCTATCTGCAGGACAAG 600
                ******** ** ** ** ** ** ** ***** ** ** ***** ******** ** **

pBAM1           CTGGCCCATAACGAACGTTTTGTGGTGCGTAGCAAACATCCGCGTAAAGATGTGGAAAGC 660
pBAM2           CTCGCGCATAATGAACGTTTTGTGGTTCGCTCCAAGCACCCGCGCAAAGACGTTGAGTCG 660
                ** ** ***** ************** **   *** ** ***** ***** ** **

pBAM1           GGCCTGTATCTGTATGATCACCTGAAAAACCAGCCGGAACTGGGCGGCTATCAGATTAGC 720
pBAM2           GGGCTGTACCTGTACGACCACCTCAAAAACCAGCCTGAGTTGGGCGGATACCAGATTTCC 720
                ** ***** ***** ** ***** *********** **  ******* ** ******  *

pBAM1           ATTCCGCAGAAAGGCGTGGTGGATAAACGTGGCAAACGTAAAAACCGTCCGGCGCGTAAA 780
pBAM2           ATTGGCCAGAAAGGCGTCGTTGACAAGCGGGGTAAGCGCAAGAATCGCCCCGCCCGGAAA 780
                ***   *********** ** ** ** ** ** ** ** ** ** ** ** ** ** ***
```

```
pBAM1       GCGAGCCTGAGCCTGCGTAGCGGCCGTATTACCCTGAAACAGGGCAACATTACCCTGAAC 840
pBAM2       GCCTCTCTCTCCCTGCGGAGTGGGCGTATCACCCTCAAGCAAGGTAATATCACGCTCAAC 840
            **    **   ****** ** ** ***** ***** ** ** ** ** ** ** ** ***

pBAM1       GCGGTGCTGGCCGAAGAAATTAATCCGCCGAAAGGCGAAACCCCGCTGAAATGGCTGCTG 900
pBAM2       GCCGTACTGGCGGAAGAGATCAACCCCCCCAAAGGTGAGACTCCATTGAAGTGGCTGCTG 900
            ** ** ***** ***** ** ** ** ** ***** ** ** **  **** *********

pBAM1       CTGACCAGCGAGCCGGTGGAAAGTCTGGCCCAAGCGCTGCGTGTGATTGATATTTATACC 960
pBAM2       CTGACCAGCGAGCCCGTGGAGTCTCTGGCCCAAGCGCTGCGCGTGATTGACATTTATACC 960
            ************** *****   ****************** ******** *********

pBAM1       CATCGTTGGCGCATTGAAGAATTTCACAAAGCGTGGAAAACGGGTGCGGGTGCGGAACGT 1020
pBAM2       CACCGGTGGCGAATCGAAGAATTTCATAAAGCCTGGAAAACGGGTGCCGGAGCGGAGCGT 1020
            ** ** ***** ** *********** ***** ************** ** ***** ***

pBAM1       CAGCGTATGGAAGAACCGGATAACCTGGAACGTATGGTGAGCATTCTGAGCTTTGTGGCG 1080
pBAM2       CAGCGGATGGAAGAACCTGATAATCTCGAACGCATGGTCAGTATCCTCAGCTTTGTGGCC 1080
            ***** *********** ***** ** ***** ***** ** ** ** ***********

pBAM1       GTGCGTCTGCTGCAACTGCGTGAATCTTTTACTCCGCCGCAAGCACTGCGTGCGCAGGGC 1140
pBAM2       GTGCGCCTGCTGCAACTGCGTGAGTCGTTCACCCCACCGCAGGCCCTGCGAGCTCAAGGG 1140
            ***** ***************** ** ** ** ** ***** ** ***** ** ** **

pBAM1       CTGCTGAAAGAAGCGGAACACGTTGAAAGCCAGAGCGCGGAAACCGTGCTGACCCCGGAT 1200
pBAM2       CTCTTGAAAGAAGCCGAACATGTCGAGTCCCAAAGCGCCGAAACGGTACTGACGCCGGAT 1200
            **  ********** ***** ** **   *** ***** ***** ** ***** ******

pBAM1       GAATGCCAACTGCTGGGCTATCTGGATAAAGGCAAACGCAAACGCAAAGAAAAAGCGGGC 1260
pBAM2       GAGTGCCAGTTGTTGGGATATCTCGATAAGGGCAAGCGCAAGCGGAAAGAAAAGGCCGGC 1260
            ** *****  ** **** ***** ***** ***** ***** ** ******** ** ***

pBAM1       AGCCTGCAATGGGCGTATATGGCGATTGCGCGTCTGGGCGGCTTTATGGATAGCAAACGT 1320
pBAM2       TCCCTTCAGTGGGCCTACATGGCGATCGCGCGTCTGGGGGGATTCATGGATAGTAAGCGG 1320
              *** ** ***** ** ******** *********** ** ** ******** ** **

pBAM1       ACCGGCATTGCGAGCTGGGGTGCGCTGTGGGAAGGTTGGGAAGCGCTGCAAAGCAAACTG 1380
pBAM2       ACTGGGATCGCTTCCTGGGGCGCGCTGTGGGAAGGCTGGGAAGCCCTCCAGAGCAAACTC 1380
            ** ** ** **   ****** ************** ******** ** ** ********

pBAM1       GATGGCTTTCTGGCCGCGAAAGACCTGATGGCGCAGGGCATTAAAATCTAA          1431
pBAM2       GATGGGTTCCTGGCAGCAAAAGATCTGATGGCGCAGGGTATCAAGATTTGA          1431
            ***** ** ***** ** ***** ************** ** ** ** * *
```

FIG. S2. Maps of pBAM2-GFP and pBAM2-KDC plasmids used to generate chromosomally integrated *A. ferrooxidans* strains with the super-folder (sfGFP) and 2-keto decarboxylase (KDC) genes, respectively. Top: pBAM2-GFP contains the codon-optimized version of the transposase and the sfGFP gene. Bottom: pBAM2-KDC contains the codon-optimized version of the transposase and the KDC gene.

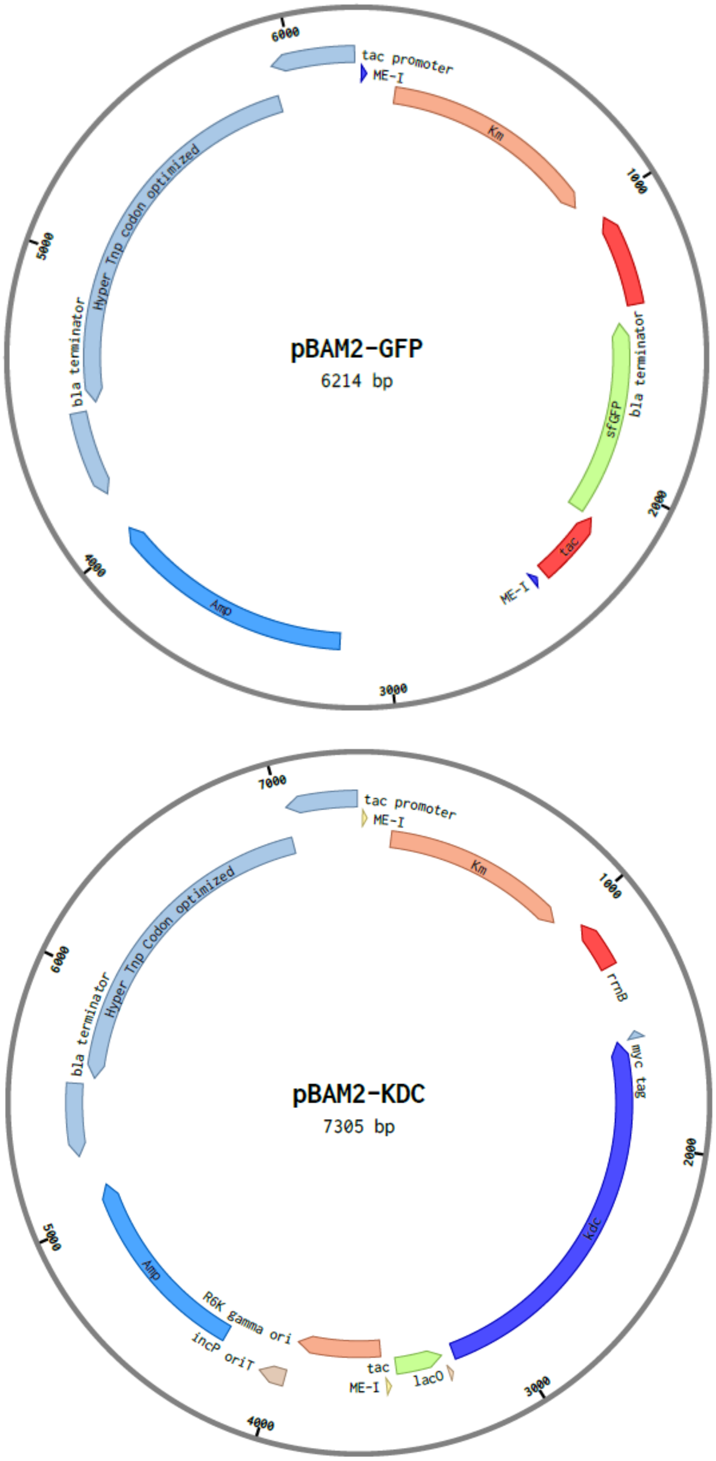

FIG. S3. Schematic representation of the inverse PCR method used in this study. The genomic DNA from the mutant strains were digested overnight with EcoRI. The cut DNA was ligated intramolecularly resulting in two circular pieces of DNA containing a portion of tn5 and the flanking genomic DNA. The ligated DNA was used as a template for four PCR reactions using two sets of nested reactions using the primers shown to amplify the flanking genomic DNA. The shorter fragment of DNA amplified was sequenced with the sequencing primers, tnFSeq and tnRSeq.

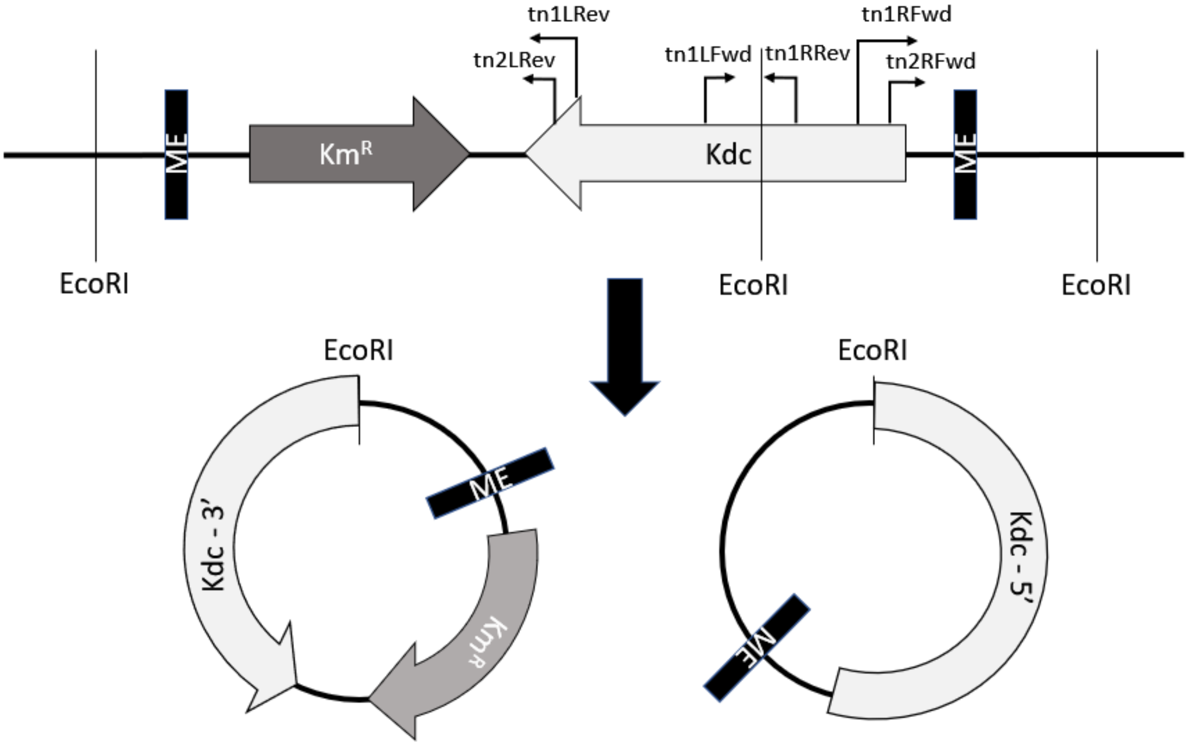

Table S2. Media formulations used in this study. AFM3 and F2S media were adjusted to pH 1.8. S204 and SM4 medium were adjusted to pH 5.0. $Fe_2(SO_4)_3\cdot 5H_2O$ was used for adding Fe(III); however, the weight by volume is calculated for the amount of Fe(III) ions in solution.

AFM3 (F2S)	**Mating**	**S204**	**SM4**
0.8 g/L $(NH_4)_2SO_4$	0.625% (w/v) agar	0.625% (w/v) agar	0.8 g/L $(NH_4)_2SO_4$
0.1 g/L HK_2PO_4	11.25 g/L $(NH_4)_2SO_4$	11.25 g/L $(NH_4)_2SO_4$	0.1 g/L HK_2PO_4
2.0 g/L $MgSO_4\cdot 7H_2O$	0.38 g/L KCl	0.38 g/L KCl	2.0 g/L $MgSO_4\cdot 7H_2O$
5ml/L Trace Mineral Supplement (ATCC® MD-TMS™)	1.86 g/L $MgSO_4\cdot 7H_2O$	1.86 g/L $MgSO_4\cdot 7H_2O$	5ml/L Trace Mineral Supplement (ATCC® MD-TMS™)
10 mM Citric Acid	0.0005% (w/v) Fe(III)	0.04% (w/v) Fe(III)	1 mM Citric Acid
100 mM $FeSO_4\cdot 7H_2O$	0.05% (w/v) NaS_2O_3	0.2% (w/v) NaS_2O_3	0.0005% (w/v) Fe(III)
0.1% (w/v) colloidal sulfur for F2S	100 µM diaminopimelic acid	100 µM diaminopimelic acid	100 µM diaminopimelic acid
	40 µg/mL leucine	40 µg/mL leucine	40 µg/mL leucine
	0.05% (w/v) yeast extract	225 µg/mL kanamycin	0.1% (w/v) colloidal sulfur
			250 µg/mL kanamycin