# Metagenomic scaffolds enable combinatorial lignin transformation

Cameron R. Strachan[a,b,1], Rahul Singh[a,1], David VanInsberghe[a,1,2], Kateryna Ievdokymenko[a,c], Karen Budwill[d], William W. Mohn[a], Lindsay D. Eltis[a,c,e], and Steven J. Hallam[a,b,c,f,3]

[a]Department of Microbiology and Immunology, University of British Columbia, Vancouver, BC, Canada V6T 1Z3; [b]MetaMixis, Inc., Vancouver, BC, Canada V6J 3R3; [c]Genome Science and Technology Graduate Program, University of British Columbia, Vancouver, BC, Canada V6T 1Z3; [d]Environment and Carbon Management, Alberta Innovates–Technology Futures, Edmonton, AB, Canada T6N 1E4; [e]Department of Biochemistry and Molecular Biology, University of British Columbia, Vancouver, BC, Canada V6T 1Z3; and [f]Graduate Program in Bioinformatics, University of British Columbia, Vancouver, BC, Canada V6T 1Z3

Edited by Richard A. Dixon, University of North Texas, Denton, TX, and approved June 3, 2014 (received for review January 27, 2014)

**Engineering the microbial transformation of lignocellulosic biomass is essential to developing modern biorefining processes that alleviate reliance on petroleum-derived energy and chemicals. Many current bioprocess streams depend on the genetic tractability of *Escherichia coli* with a primary emphasis on engineering cellulose/hemicellulose catabolism, small molecule production, and resistance to product inhibition. Conversely, bioprocess streams for lignin transformation remain embryonic, with relatively few environmental strains or enzymes implicated. Here we develop a biosensor responsive to monoaromatic lignin transformation products compatible with functional screening in *E. coli*. We use this biosensor to retrieve metagenomic scaffolds sourced from coal bed bacterial communities conferring an array of lignin transformation phenotypes that synergize in combination. Transposon mutagenesis and comparative sequence analysis of active clones identified genes encoding six functional classes mediating lignin transformation phenotypes that appear to be rearrayed in nature via horizontal gene transfer. Lignin transformation activity was then demonstrated for one of the predicted gene products encoding a multicopper oxidase to validate the screen. These results illuminate cellular and community-wide networks acting on aromatic polymers and expand the toolkit for engineering recombinant lignin transformation based on ecological design principles.**

environmental genomics | synthetic biology

Lignin is the second-most abundant biopolymer on Earth and a promising feedstock for deriving energy, fine chemicals, and structural materials from renewable plant resources (1, 2). The synthesis of lignin occurs within plant cell walls by free radical reactions that cross-link phenylpropanoids into a heterogeneous matrix that is resistant to microbial and chemical attack (3). Lignin recalcitrance is further reflected in the deposition of coal throughout the Carboniferous period before the emergence of fungal enzymes associated with lignolysis in Permian forest soil ecosystems (4). To date, fungi are the most widely studied lignin-degrading microbes and are the major source of lignin-transforming enzymes, including laccases, manganese-dependent peroxidases, and lignin peroxidases (5–9). Functional characterization of these metalloenzymes is consistent with a model of lignin degradation based on oxidative combustion mediated by a broad range of small molecule oxidants, such as veratryl alcohol and Mn(III) (9).

Despite this mechanistic understanding, however, there are numerous challenges associated with engineering lignin bioprocess streams, including the genetic intractability of many fungal systems and barriers to scale-up expression of active fungal-derived enzymes in heterologous systems, such as *Escherichia coli*. Although a handful of bacterial isolates, including *Enterobacter lignolyticus* SCF1 and *Rhodococcus jostii* RHA1, encode lignin-transforming phenotypes, the functional diversity of bacterial lignin-transforming enzymes in the environment remains low (5–9). Implementing high-throughput methods to expedite the discovery of bacterial lignin transformation pathways provides one promising route toward overcoming these challenges (10–13); however, efforts to develop such functional screens have been unreliable, owing to the inherent complexity of the lignin polymer (14).

Here we develop and validate a biosensor responsive to lignin transformation products. We use this biosensor in functional metagenomic screens of fosmid libraries sourced from coal beds to recover scaffolds conferring an array of lignin transformation phenotypes. Genetic and biochemical approaches were used to characterize these scaffolds resulting in an ecological model for bacterial lignin transformation with potential application to biorefining process streams.

## Results and Discussion

We reasoned that sensing lignin transformation products rather than labeling the lignin polymer itself might improve signal detection across a wide range of enzymatic activities. To this end,

### Significance

Plant biomass conversion into biofuels and chemicals can reduce human reliance on petroleum and promote sustainable biorefining processes. The structural polymer lignin can comprise up to 40% of plant biomass, but resists decomposition into valuable monoaromatic compounds. In this study, we devised a previously unidentified biosensor responsive to lignin transformation products. We used this biosensor in a functional screen to recover metagenomic scaffolds sourced from coal bed bacterial communities. Genetic and biochemical analyses revealed six functional classes mediating lignin transformation that are mobilized in nature via horizontal gene transfer. Our results suggest that bacterial lignin transformation is an adaptive trait that can be exploited to engineer combinatorial arrays with defined product profiles, a prerequisite for scale-up production using different plant sources.

Author contributions: C.R.S., R.S., D.V., L.D.E., and S.J.H. designed research; C.R.S., R.S., D.V., and K.I. performed research; K.B. contributed new reagents/analytic tools; C.R.S., R.S., D.V., K.I., and S.J.H. analyzed data; C.R.S., R.S., D.V., W.W.M., L.D.E., and S.J.H. wrote the paper; W.W.M. and L.D.E. helped supervise the project; and S.J.H. supervised the project.

Conflict of interest statement: C.R.S. and S.J.H. are cofounders of MetaMixis, Inc., a synthetic biology company that uses coculture-based biosensor screening to interrogate metagenomic libraries for the production of industrial enzymes and active pharmaceutical intermediates. In addition, US provisional patent applications describing components of this work have been filed (with serial nos. 61/936,448 and 62/013,369).

This article is a PNAS Direct Submission.

Freely available online through the PNAS open access option.

Data deposition: Sequences of active fosmid clones were deposited in GenBank database (accession nos. KJ802934–KJ802957).

[1]C.R.S., R.S., and D.V. contributed equally to this work.

[2]Present address: Graduate Program in Microbiology, Massachusetts Institute of Technology, Cambridge, MA 02139.

[3]To whom correspondence should be addressed. E-mail: shallam@mail.ubc.ca.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1401631111/-/DCSupplemental.

we interrogated an *E. coli* clone library of fluorescent transcriptional reporters with a mixture of lignin transformation products including vanillin, vanillate, and *p*-coumarate (Fig. 1*A*) (15). The most responsive clone harbored a promoter regulating the *emrRAB* operon, encoding a negative feedback regulator (EmrR) and multidrug resistance pump (EmrAB) acting on various structurally unrelated antibiotics (16, 17).

Because the compounds used to identify this promoter had not yet been shown to induce *emrRAB* expression, we evaluated response specificity using a library of monoaromatic compounds (Fig. 1*B*). Sensitivity of detection increased with promoter copy number, reaching a lower detection threshold of 50 μM using the three most active lignin transformation products (Fig. 1*C* and *SI Appendix*, Fig. S1). We then demonstrated the capacity of this promoter to detect in vitro lignin transformation by monitoring the formation of monoaromatic products from a solvent-fractionated hardwood kraft lignin (SF-HKL) using an engineered manganese-oxidizing dye decolorizing peroxidase, DypB N246A (Fig. 1*D*) (5). Coculture-based detection of lignin transformation products was also demonstrated using bacterial isolates affiliated with multiple phyletic groups, including *E. lignolyticus* SCF1 and *R. jostii* RHA1 (*SI Appendix*, Fig. S2).

Given these initial screening results, we evaluated the specific role of the EmrR transcriptional regulator in responding to lignin transformation products demonstrating that *emrR* is necessary and sufficient for the compound-dependent activation of the *emrRAB* operon (*SI Appendix*, Fig. S3*A*). Abolishing *emrR* activity, but not *emrB* activity, caused slightly impaired growth kinetics in the presence of several lignin transformation products (*SI Appendix*, Fig. S3*B*). Moreover, a dramatic lag phase was consistently observed in *emrR* loss-of-function mutants exposed to SF-HKL pretreated with DypB N246A (*SI Appendix*, Fig. S4*A*). Complementation by overexpression not only rescued the impaired growth phenotype in the presence of monoaromatic compounds, but also increased growth rate and final biomass accumulation (*SI Appendix*, Fig. S4*B*). Taken together, these results are consistent with a role for EmrR in regulating a metabolic network responsive to monoaromatic exposure in the environment, and reinforce the potential use of EmrR and its promoter as a versatile biosensor (P*emrR*-GFP) in functional screens for lignin transformation.

High molecular weight coal is derived from lignin but made increasingly recalcitrant through the processes of coalification (18). We reasoned that coal beds would be enriched for bacterial genes encoding lignin transformation proteins, because, unlike in forest soils, here primary transformation is not likely to be fungal-mediated (4, 19). Therefore, we chose to interrogate bacterial-dominated metagenomic libraries sourced from Rockyford Standard (CO182) and Basal (CO183) coal zones in Alberta, Canada with the P*emrR*-GFP biosensor (*SI Appendix*, Fig. S5) (19). Metagenomic libraries from CO182 and CO183 were constructed using the Fosmid CopyControl system (pCC1FOS; Epicentre), based on previous reports suggesting that increased copy number enhances heterologous gene expression in the EPI300 *E. coli* host (20). In parallel, the P*emrR*-GFP biosensor was transferred to the pCC1FOS vector used in library production to facilitate coculture-based screening using shared antibiotic selection. A total of 46,000 fosmids arrayed in 384-well plates were grown in the presence of SF-HKL overnight before biosensor addition. Cocultures were subsequently grown for 3 h before GFP fluorescence was measured. Fluorescent signals were normalized to background and corrected for edge effects. Finally, 24 fosmids activating the P*emrR*-GFP biosensor (16 from CO182 and 8 from CO183) were selected for downstream sequencing and functional characterization (*SI Appendix*, Fig. S5 *B* and *C*).

To verify the production of lignin transformation products activating the P*emrR*-GFP biosensor, nine of the most active fosmid clones were incubated in the presence of SF-HKL and a second industrially purified high-performance lignin (HP-L)

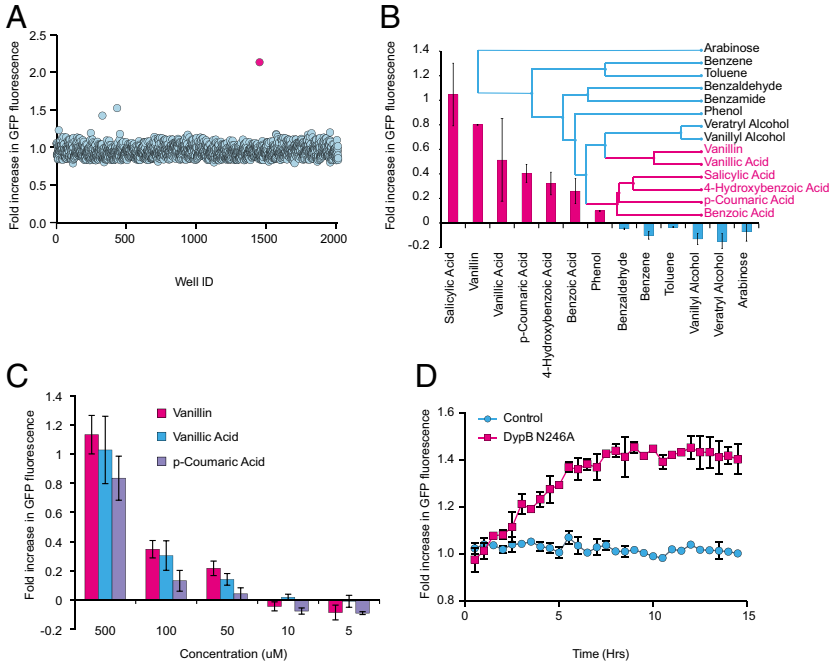


**Fig. 1.** P*emrR*-GFP biosensor discovery and characterization. (*A*) Screening *E. coli* intragenic regions with monaromatic lignin transformation products including vanillin, vanillic acid, *p*-coumaric acid, vanillyl alcohol, and veratryl alcohol. (*B*) Relative reporter signal after incubation with 0.5 mM of select benzene derivatives for 2 h. Tree represents hierarchical clustering of the compound similarity using the single linkage algorithm. (*C*) Reporter sensitivity after 2 h. (*D*) Monitoring of in vitro lignin oxidation by DypB N246A in the presence of glucose oxidase, manganese, and hydrogen peroxide. Controls did not contain manganese. Error bars represent 95% confidence intervals ($n = 3$).

substrate (21). Lignin transformation products, including vanillin, syringaldehyde, and syringic acid, were then measured by gas chromatography-mass spectrometry (GC-MS). An array of monoaromatic compound profiles were observed for single fosmid incubations, which varied between SF-HKL and HP-L, with certain lignin transformation products reaching up to 15 mg $L^{-1}$, consistent with different substrate properties or varying specificities of fosmid-encoded enzymes (Fig. 2 and Dataset S1).

Curiously, fosmid cocultures exhibited synergy in combination, producing monoaromatic compound profiles that differed from individual fosmid incubation profiles in unexpected ways (Fig. 2). Moreover, whereas single fosmid incubations with SF-HKL led to precipitate formation, only coculture fosmid incubations were capable of forming precipitates with HP-L (*SI Appendix*, Fig. S6). These observations suggest that fosmids recovered in the P*emrR*-GFP biosensor screen confer lignin transformation phenotypes with different end product profiles, similar to observations made for fungal lignin transformation (22, 23).

Random transposon mutagenesis identified genes encoded on 11 fosmids necessary for activating the P*emrR*-GFP biosensor. Nine out of 11 mutagenized fosmids contained insertions that reduced biosensor activation in two or more genes, suggesting that observed lignin transforming phenotypes require multiple pathway components (Fig. 3 and *SI Appendix*, Fig. S7). A comparison of transposon insertion sites across the 11 fosmids identified six functional classes mediating lignin transformation phenotypes. These included genes predicted to encode oxidoreductase activity (associated with lignin transformation), cosubstrate generation (hydrogen peroxide formation), protein secretion (secretion apparatus or signal peptide), small molecule transport (multidrug efflux superfamily), motility [methyl-accepting chemotaxis proteins (MCPs)], and signal transduction (PAS domain-containing sensors) (Fig. 3 and *SI Appendix*, Fig. S7). Complete sequencing and comparative analysis of all 24 fosmids activating the P*emrR*-GFP biosensor identified recurring subsets of genes on typically nonsyntenic clones encoding one or more of the six functional classes identified by transposon mutagenesis (Fig. 3 and *SI Appendix*, Fig. S7).

To further validate the screen, we biochemically characterized a gene product whose disruption in fosmid 182_08_C21 resulted in reduced biosensor activation and the loss of lignin transformation activity (*SI Appendix*, Fig. S8). The predicted protein had 100% and 80% amino acid sequence identity with CopA in *Pseudomonas stutzeri* ATCC 14405 and *Pseudomonas putida*, respectively. Both strains are known to degrade a range of aromatic compounds (8, 24). CopA is a TAT-secreted multicopper oxidase (MCO) that provides resistance to copper in *Pseudomonas syringae*, a plant pathogen, and has been described as a pseudolaccase owing to the requirement of exogenous Cu(II) for oxidase activity (25, 26). Interestingly, *copA* is located at one

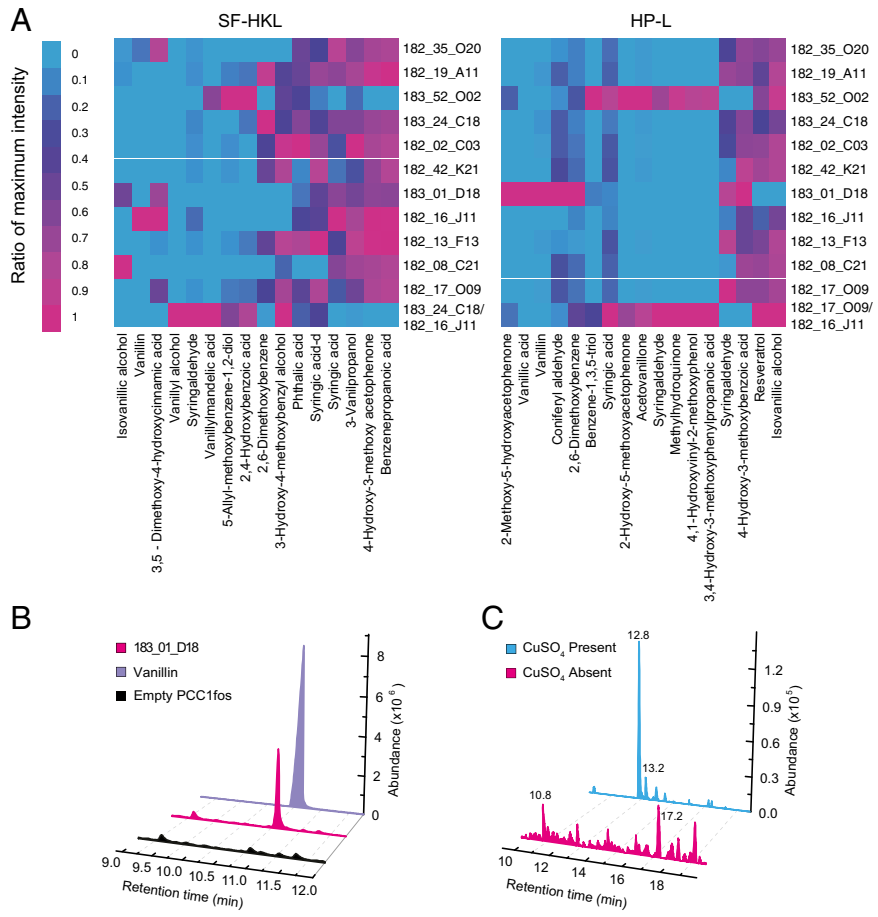


**Fig. 2.** Profiling monoaromatic compounds by GC-MS. (*A*) Intensity of lignin-related monoaromatic compounds in culture supernatant as a ratio of the maximum. All values are normalized to a control strain harboring an empty fosmid and an internal standard. Clones were incubated with both SF-HKL and HP-L in minimal media. (*B*) Chromatogram showing production of vanillin on incubation of the clone 183_01_D18 with HP-L. The empty fosmid control and an authentic standard of vanillin are shown for the comparison. (*C*) Difference chromatogram showing CopA-catalyzed transformation of HP-L in the presence and absence of $CuSO_4$.

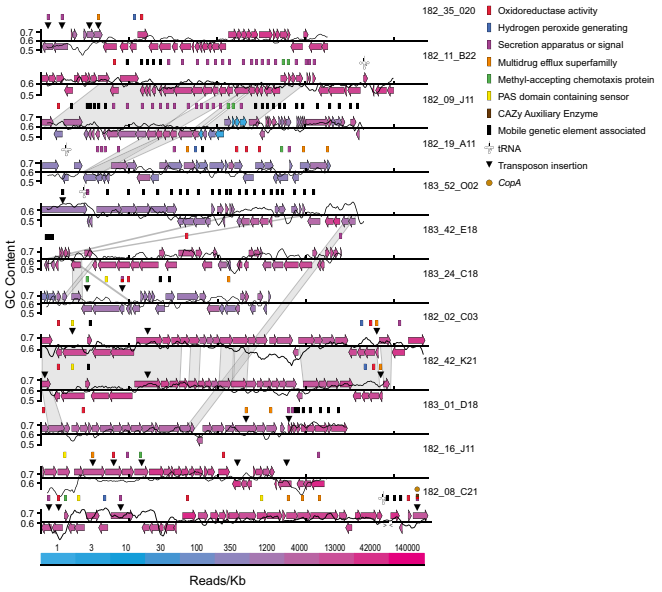


**Fig. 3.** Genetic context maps for select active fosmids. Functional classes related to lignin degradation, Carbohydrate-active enzymes (CAZy) auxiliary enzymes, mobile genetic element associations, transposon mutagenesis insertions, and tRNAs are annotated. Transposon mutagenesis insertions that caused a statistically significant decrease in GFP signal were identified using a z-score ratio. The G+C ratio for every 200 nucleotides and gene abundance determined by mapping more than 500 million illumina reads sourced from the coal bed milieu is displayed as well. Connections represent protein homologs with minimum 50% identity and an E-value of 10E-20.

end of the active fosmid clone, and the encoded protein has a 39-residue C-terminal truncation compared with reference sequences. This truncation includes a conserved motif (HCHXXXHXXXXM/L/F) required for binding of type 1 (T1) and type 3 (T3) copper centers. Initial efforts to characterize this protein based on oxidation of 2,6-DMP were unsuccessful; thus, a full-length *copA* with 100% nucleotide sequence identity to the active fosmid gene was amplified from CO182 genomic DNA and used to produce a poly-histidine-tagged protein in *E. coli* (*SI Appendix*, Fig. S8*B*). The affinity-purified CopA (*SI Appendix*, Fig. S8*C*) catalyzed the oxidation of 2,2′-azino-bis(3-ethylbenzothiazoline-6-sulphonic acid (ABTS; 1,000 ± 30 nmol $min^{-1}$ $mg^{-1}$, pH 4.5) and 2,6-DMP (700 ± 40 nmol $min^{-1}$ $mg^{-1}$, pH 8) in the presence of added Cu (II). Interestingly, CopA also catalyzed the oxidation of 2,6-DMP in the absence of Cu(II) (130 ± 10 nmol $min^{-1}$ $mg^{-1}$ at pH 8 after a lag phase of ∼10 min (*SI Appendix*, Fig. S8 *D* and *E*).

To evaluate the lignin-transforming activity of CopA, samples of HP-L were incubated for 3 h with the enzyme in the absence and presence of exogenous Cu(II). GC-MS analysis revealed the appearance of new peaks in both the presence and absence of Cu(II) compared with controls containing no enzyme (Fig. 2*C*). Products identified included 6-acetyl-2,5-dihydroxy-1,4-naphthoquinone ($t_r$ = 10.8 min), 2,6-dimethoxybenzene-1,4-diol ($t_r$ = 12.8 min), 4-hydroxy-3-methoxybenzoic acid ($t_r$ = 13.4 min), and 4-hydroxy-3,5-dimethoxybenzoic acid ($t_r$ = 17.2 min). Efforts to elucidate the molecular basis for CopA activity and to characterize the truncated enzyme in whole cells continue. Lignin transformation by a CopA-type MCO is reminiscent of the lignin-transforming activity of DypB, a dye-decolorizing peroxidase that has been implicated in iron metabolism (27).

Although oxidoreductase activity, hydrogen peroxide generation, and protein secretion have been previously implicated in lignin transformation, the specific roles of the remaining three functional classes remain uncertain (2, 8). It is notable that several of the fosmids identified with the P*emrR*-GFP biosensor actually encode small molecule transport systems homologous to *emrR* and *emrB*, further reinforcing a role for these genes in regulating microbial responses to monoaromatic exposure in the environment (Dataset S2). Cell motility could facilitate optimal cell positioning along microscale transformational gradients. This relationship between lignin transformation and cell motility is highlighted by a recent study that found an enrichment of MCP-encoding genes and transcripts in the microbiota of wood-feeding termites relative to dung-feeding termites (28). Finally, signal transduction proteins could play a role in mediating lignin substrate specificity among and between microbial groups and contribute to gradient formation. Indeed, recent cultivation-dependent studies using nitrated lignin substrates from wheat, *Miscanthus*, and pine identified multiple transformation phenotypes among and between bacterial and fungal isolates (14). The need for genes encoding both MCP and signal transduction on the fosmids for activating the biosensor implicates both of these functional classes in mediating lignin transformation phenotypes (Fig. 3 and *SI Appendix*, Fig. S7).

In addition to the six functional classes described above, 16 of the 24 completely sequenced fosmids harbored genes associated with mobile genetic elements (MGEs), including plasmids and phage. These genes were typically positioned proximal to one or more of the six functional classes, suggesting a role for horizontal gene transfer (HGT) in propagating lignin transformation phenotypes in the environment (Fig. 3 and *SI Appendix*, Figs. S7 and S9). To further explore the relationship between lignin transformation phenotypes and putative gene transfer networks, we examined coverage depth, G+C content variation, and tRNA positioning (Fig. 3 and *SI Appendix*, Fig. S7). Although no consistent pattern of G+C content variation or tRNA positioning was observed for functional classes meditating lignin transformation phenotypes, fragment recruitment of 500 million unassembled Illumina reads sourced from CO182 and CO183 identified low to intermediate coverage in genomic intervals containing MGE-associated genes and one or more of the six functional classes (Fig. 3 and *SI Appendix*, Fig. S7). For example, *copA* on fosmid 182_08_C21 was linked to a genomic interval exhibiting decreased coverage in proximity to three MGE-associated genes and one tRNA. To expand on this observation, we selected active clones from the CO182 fosmid library affiliated with *Pseudomonas* and mapped predicted protein sequences onto selected reference genomes (*SI Appendix*, Fig. S10). The resulting coverage plot revealed syntenic intervals shared between active fosmid clones and the reference genomes. In most cases, these syntenic intervals were interrupted by unique gene content encoding MGE-associated genes and one or more of the six functional classes consistent with island or islet formation (*SI Appendix*, Fig. S10).

Because genomic regions are more likely than whole genomes to sweep through populations, genome coverage can provide insight into both the ecological and the functional importance of environmental DNA (29). Islands and islets have been shown to transfer ecologically important traits throughout a habitat-specific horizontal gene pool, with notable examples in symbiotic and marine ecosystems (30–32). We have retrieved a number of coal bed-derived fosmids that confer lignin transformation phenotypes; share common enzymatic, regulatory, and transport features; and display genomic features associated with HGT. The fosmids exhibit differential coverage patterns among and between themselves in the coal bed milieu, suggesting that lignin transformation is a habitat-specific trait shaped by local interactions and frequency-dependent selection (30, 32). Based on the foregoing observations, we propose a model for lignin transformation in the environment in which reusable functional classes are shuffled through HGT to generate adaptive combinatorial arrays that promote niche adaptation based on a range of lignin substrates and transformation products

(*SI Appendix*, Fig. S11). Although rational engineering has driven the development of modern biorefining processes, our results suggest that there may be utility in exploiting ecological design principles to build a new generation of biorefining microorganisms through the use of naturally assembled genetic parts.

## Methods

DNA was manipulated according to standard procedures. Fosmid library preparation, transposon mutagenesis, and purification were performed with kits sourced from Epicentre and sequenced as described previously (29). All chemicals used were of analytical grade and purchased from Sigma-Aldrich.

**Strains and Growth Conditions.** Minimal media consisted of M9 minimal media supplemented with (0.4% glucose, 100 μg $mL^{-1}$ arabinose, 40 μg $mL^{-1}$ leucine, 1 mM $MgSO_4$, and 2 μM thiamine. Lysogeny broth (LB) and minimal media were supplemented with 50 μg $mL^{-1}$ kanamycin, 12.5 μg $mL^{-1}$ chloramphenicol, and 100 μg $mL^{-1}$ ampicillin to maintain pUA66, PCC1fos, and pBAD24, respectively. The *emrR* and *emrB* KO and cognate WT strains were obtained from the Keio collection through the Coli Genetic Stock Center. All cultures were grown at 37 °C in a 220-rpm rotary shaker unless stated otherwise.

**Plasmid Construction.** The *emrRAB* promoter region and GFP were amplified from the pUA66 backbone with primers Frep (GCGGAATTCCGCAGCATTATCATCC) and Rrep (GCGAAGCTTCCTGCAGGTCTGGACATTTAT). The PCR product was digested with EcoRI and HindIII and then ligated with EcoRI/HindIII-digested PCC1fos to generate PCC1reporter. The PCC1fos vector is under a copy-control system, which is inducible in the EPI300 background host (Epicentre). PCC1reporter under high copy number in EPI300 is referred to as the P*emrR*-GFP biosensor. For inducible overexpression of emrR, *emrR* was amplified from *E. coli* K12 genomic DNA using the primers FemrR (GCGGAATTCatgGATAGTTCGTTTACGCCCA) and RemrR (GCGAAGCTTTTAGCTCATCGCTTCGAGAACC). The resulting PCR product was also digested with EcoRI and HindIII and then ligated with EcoRI/HindIII-digested pBAD24, yielding pBAD24emrR. The *copA* gene (including the TAT sequence) was amplified from metagenomic DNA (CO182) using PCR and the primers *CopA-For* (GCTGAACCATATGCAT-AGAAAACCCA) and *CopA-Rev* (GCTGAACCATATGCATAGAAAACCCA). The amplicon was cloned into pET15b$^+$ (Novagen) using NdeI and XhoI restriction sites, and its nucleotide sequence was confirmed to generate pETCopA$_{C21}$.

**Screening the *E. coli* Reporter Library.** A library of 1,820 *E. coli* K12 MC1655 intragenic regions fused to gfpmut2 on low-copy plasmids was replicated into 96-well round-bottom culture plates containing M9 minimal medium supplemented with glucose. After growth overnight, a compound pool comprising 1 mM vanillin, vanillic acid, p-coumaric acid, vanillyl alcohol, and veratryl alcohol was added. The plates were then incubated, and GFP fluorescence was measured by reading excitation at 481 nm and emission at 508 nm on a Varioscan Flash Spectral Scanning Multimode Reader (Thermo Scientifc) before selecting the most active clone. All GFP measurements were made as described herein.

**PemrR-GFP Biosensor Characterization.** The P*emrR*-GFP biosensor was grown overnight and diluted 1/10 in 180 μL of LB. The compound of interest, dissolved in 20 μL of 30% DMSO was then added before a 2-h incubation and subsequent reading of GFP fluorescence. Arabinose was removed from the media when comparing the effect of plasmid copy number. For in vitro enzyme assays, P*emrR*-GFP was diluted 1/10 and incubated in M9 minimal medium with 0.5% glucose, 0.5 μg SF-HKL, 40 mM manganese, 100 nM glucose oxidase, and 50 nM DypB N246A. GFP measurements were recorded every 30 min.

**Environmental Isolate Screening.** In vivo isolate cultures were carried out in 1/10 diluted low-salt (50 mg/L $CaCl_2$*$2H_2O$) LB, 1 g $L^{-1}$ of HP-L, and 1% DMSO that was filtered through a 0.2-μm filter (ExpressPlus; Milipore). Cultures were inoculated 1/100 from saturated cultures grown in 1/10 diluted LB and grown in stationary flasks for 2 wk before cells were centrifuged at 1,600 × *g*, and culture supernatant was removed and filtered through a 0.2- μm filter (ExpressPlus; Milipore) before being assayed with the P*emrR*-biosensor. The supernatant from duplicate cultures was aliquoted in 180-μL volumes in triplicate. To this, 20 μL of the P*emrR*-GFP biosensor, diluted 1/4 from an overnight culture in LB, was added. The mixture was allowed to incubate while stationary for 2 h before GFP measurements were made.

**GC-MS of Clones Incubated with Lignin.** Lignin-containing growth medium was prepared by adding HP-L or SF-HKL (1 g $L^{-1}$) dissolved in DMSO (3% final concentration) to minimal medium under stirring at 37 °C. Lignin-amended medium was stirred for another 1 h, followed by filtration (0.2 μm; ExpressPlus; Millipore) to remove any precipitate. The EPI300 strains harboring fosmids were then inoculated 1/10,000 in 5 mL of lignin medium from an overnight culture in LB. The cultures were grown as described above for 16 h, after which cells were spun down (16,000 × *g* for 10 min), and the culture supernatant was removed. The culture supernatant was acidified using formic acid (10% vol/vol). Acidification precipitated the residual lignin in all of the clones, which was removed by centrifugation (16,000 × *g* for 10 min) and filtration (0.2 μm DMSO safe filter; Pall). The clear supernatant was extracted three times using ethyl acetate (1:1). The extracts were dried over anhydrous magnesium sulfate, and the solvent was evaporated under the stream of nitrogen. The air-dried samples were resuspended in 300 μL of pyridine. To each sample, 100 μg of 4-chlorobenzoic acid was added as an internal standard. Subsequently, the samples were derivatized using BSTFA+ TMCS- (99:1).

GCMS was performed using an HP 66890 series GC system fitted with an HP 5973 mass selective detector and a 30 × 250 μm HP-5MS Agilent column. The operating conditions were $T_{GC}$ (injector), 280 °C; $T_{MS}$ (ion source), 230 °C; oven time program ($T_0$ min), 120 °C; $T_2$ min, 120 °C; $T_{45}$ min, 300 °C (heating rate 4 °C $min^1$); and $T_{54}$ min, 300 °C. The injector volume was 1 μL. With the exception of those confirmed and quantified in Dataset S1, compounds were identified through a catalog search.

**Enzyme-Catalyzed Lignin Transformation.** The 1-mL reactions containing 0.5 mg/mL HP-L, 10% DMSO, 90 nM CopA, and 20 mM potassium phosphate (pH 8) were incubated for 3 h at 30 °C. Reactions were performed in duplicate, with and without 5 mM $CuSO_4$. Controls contained no enzyme. Extraction and GC-MS analyses were performed as described above.

**emrRAB Characterization.** All time course $OD_{600}$ and fluorescence measurements were made on an Infinite 200 PRO plate reader (TECAN) with WT (BW25112), *emrR*(–), and *emrB*(–) *E. coli* strains. Monoaromatic compounds were added in DMSO to a final concentration of 3%. The pUA66 vector harboring the reporter construct was used in the BW25112 background for monitoring fluorescence. For studying the expression of *emrR*, pBAD24-expressing *emrR* was induced with 0.06 mM arabinose.

**Fosmid Library Production.** A fosmid library was prepared from coal bed cuttings provided by Alberta Innovates Technology Futures, and DNA was extracted from the homogenized samples using previously described methods (33). The environmental DNA was cloned into the PCC1fos copy control vector and transformed into the EPI300 host (Epicentre) as described previously (34). Both CO182 and CO183 samples yielded ∼60,000 fosmid clones with an average insert size of 42 kb. Approximately 20,000 clones from each sample were Sanger end-sequenced (Applied Biosystems 3730 system) at the Michael Smith Genome Science Center, Vancouver, BC, Canada.

**High-Throughput Functional Screening.** For fosmid library screening, 60,000 clone libraries were replicated using a Qpix2 robotic colony picker (Genetix) in 384-well black plates (34). Clones were grown in 45 μL of LB for 12 h, and 20 μL of LB containing HP-L (added as described in the GC-MS profile section) was added, followed by another 5-h incubation. The P*emrR*-GFP biosensor was then added by diluting an overnight culture 1/4 and adding 20 μL, followed by a 3-h incubation before florescent measurements were taken.

**Full Fosmid Sequencing.** Once 24 active clones were selected, fosmid DNA was extracted using the FosmidMax DNA preparation kit (Epicentre) according to the manufacturer's protocols. Contaminating *E.coli* DNA was removed using Plasmid-Safe DNase (Epicentre) (34). All DNA concentrations were determined using Quant-iT PicoGreen (Invitrogen), and 500 ng of each fosmid was sent to the Michael Smith Genome Science Center for sequencing on an Illumina GAIIx sequencer.

**Transposon Mutagenesis.** For 11 of the active clones, a Tn5 transposon mutagenesis library was created using the EZ-Tn5 kan insertion kit (Epicentre) (34). Approximately 384 mutants were arrayed for rescreening as described in the high-throughput screening section. Mutants were Sanger-sequenced (Applied Biosystems 3730 system) at the Michael Smith Genome Science Center, and activity was mapped to fosmid position using BLAST. Statistically significant decreases in P*emrR*-GFP activity were then selected using a *z*-score ratio.

**Production and Purification of Recombinant CopA.** *E. coli* BL21(DE3) was freshly transformed with pET15Cop$A_{C21}$. The cells were grown in LB containing 100 μg/mL ampicillin and 2 mM $CuSO_4$ at 37 °C with shaking to an $OD_{600}$ of 0.5, at which time isopropyl β-D-thiogalactopyranoside was added to 1 mM. The cells were incubated for another 16 h at 20 °C, then harvested by centrifugation at 3,220 × *g*. The cells were washed, suspended in 15 mL of 20 mM potassium phosphate, 20 mM imidazole (pH 8) containing 0.1 μg/mL DNase, and disrupted using an Emulsi Flex-C5 homogenizer (Avestin) operated at 4 °C. Cell debris was removed by centrifugation (31,000 × *g* for 30 min).

The soluble portion was filtered through a 0.45-μm filter and incubated with 2 mL of Ni-NTA resin for 2 h at 4 °C. The resin was packed into a column (Qiagen), drained, and washed with 5 column volumes each of 20 mM potassium phosphate, 20 mM imidazole (pH 8), and 20 mM potassium phosphate, 60 mM imidazole (pH 8). CopA was eluted using 5 column volumes of 20 mM potassium phosphate, 300 mM imidazole (pH 8). The eluted protein was concentrated using a 30K NMWL Amicon Ultra-15 Centrifugal Filter Device (Millipore) and exchanged into 20 mM sodium phosphate (pH 7.5).

**Enzyme Activity Assays.** Assays were performed in triplicate in 200 μL at 25.0 ± 0.5 °C using a Cary 5000 spectrophotometer fitted with a thermostatted cuvette holder. The oxidation of ABTS was followed by monitoring the change in absorbance at 414 nm (ε 36.6 $mM^{-1}$ $cm^{-1}$) in a reaction mixture containing 2 mM ABTS, 5 mM $CuSO_4$, 0.5 μg CopA, and 50 mM sodium acetate at pH 4.5. The oxidation of 2,6-dimethoxyphenol (DMP) was followed by monitoring the change in absorbance at 470 nm (ε 49.6 $mM^{-1}$ $cm^{-1}$) in a reaction mixture containing 2 mM 2,6-DMP, 5 mM $CuSO_4$, 30 nM CopA, and 20 mM potassium phosphate (pH 8). In the absence of $CuSO_4$, reaction mixtures contained 3.3 μg CopA.

**Bioinformatics.** All ORFs were determined using Prodigal (http://prodigal.ornl.gov/) and annotated using BLAST or National Center for Biotechnology Information (NCBI) nr databases, with functional class assignments curated with the NCBI conserved domain database. Mobile element annotations were also curated using the ACLAME database (http://aclame.ulb.ac.be/). A custom perl script was designed that uses Circos (http://circos.ca) to visualize homology between fosmids. BLAST was used to map the location of the metagenomic reads (E-value cutoff of 1E-10) to the fosmid ORFs, and custom python scripts were used to visualize the abundance of each ORF in the metagenome. Phylogenetic assignment binning was done using Sort-ITEMS (http://metagenomics.atc.tcs.com/binning/SOrt-ITEMS).

**ACKNOWLEDGMENTS.** We thank Keith Mewis, Payal Sipahimalani, Melanie Scofield, and Sam Kheirandish for technical support related to library production and screening. This work was performed under the auspices of the Natural Sciences and Engineering Research Council (NSERC) of Canada, Genome British Columbia, Genome Alberta, Genome Canada, the Canada Foundation for Innovation, the Tula Foundation-funded Centre for Microbial Diversity and Evolution, and the Canadian Institute for Advanced Research. The Western Canadian Research Grid provided access to high-performance computing resources. C.R.S. and D.V. were supported by the NSERC, and K.I. was supported by the NSERC CREATE Genome Sciences and Technology training program at the University of British Columbia.

1. Zakzeski J, Bruijnincx PCA, Jongerius AL, Weckhuysen BM (2010) The catalytic valorization of lignin for the production of renewable chemicals. *Chem Rev* 110(6): 3552–3599.
2. Ruiz-Dueñas FJ, Martínez AT (2009) Microbial degradation of lignin: How a bulky recalcitrant polymer is efficiently recycled in nature and how we can take advantage of this. *Microb Biotechnol* 2(2):164–177.
3. Boerjan W, Ralph J, Baucher M (2003) Lignin biosynthesis. *Annu Rev Plant Biol* 54:519–546.
4. Floudas D, et al. (2012) The Paleozoic origin of enzymatic lignin decomposition reconstructed from 31 fungal genomes. *Science* 336(6089):1715–1719.
5. Singh R, et al. (2013) Improved manganese-oxidizing activity of DypB, a peroxidase from a lignolytic bacterium. *ACS Chem Biol* 8(4):700–706.
6. Khudyakov JI, et al. (2012) Global transcriptome response to ionic liquid by a tropical rain forest soil bacterium, *Enterobacter lignolyticus*. *Proc Natl Acad Sci USA* 109(32): E2173–E2182.
7. Brown ME, Barros T, Chang MCY (2012) Identification and characterization of a multifunctional dye peroxidase from a lignin-reactive bacterium. *ACS Chem Biol* 7(12):2074–2081.
8. Bugg TDH, Ahmad M, Hardiman EM, Rahmanpour R (2011) Pathways for degradation of lignin in bacteria and fungi. *Nat Prod Rep* 28(12):1883–1896.
9. Brown ME, Chang MC (2014) Exploring bacterial lignin degradation. *Curr Opin Chem Biol* 19:1–7.
10. Wargacki AJ, et al. (2012) An engineered microbial platform for direct biofuel production from brown macroalgae. *Science* 335(6066):308–313.
11. Bokinsky G, et al. (2011) Synthesis of three advanced biofuels from ionic liquid-pretreated switchgrass using engineered *Escherichia coli*. *Proc Natl Acad Sci USA* 108(50):19949–19954.
12. Wang X, et al. (2013) Engineering furfural tolerance in *Escherichia coli* improves the fermentation of lignocellulosic sugars into renewable chemicals. *Proc Natl Acad Sci USA* 110(10):4021–4026.
13. Choi YJ, Lee SY (2013) Microbial production of short-chain alkanes. *Nature* 502(7472): 571–574.
14. Ahmad M, et al. (2010) Development of novel assays for lignin degradation: Comparative analysis of bacterial and fungal lignin degraders. *Mol Biosyst* 6(5):815–821.
15. Zaslaver A, et al. (2006) A comprehensive library of fluorescent transcriptional reporters for *Escherichia coli*. *Nat Methods* 3(8):623–628.
16. Brooun A, Tomashek JJ, Lewis K (1999) Purification and ligand binding of EmrR, a regulator of a multidrug transporter. *J Bacteriol* 181(16):5131–5133.
17. Xiong A, et al. (2000) The EmrR protein represses the *Escherichia coli* emrRAB multidrug resistance operon by directly binding to its promoter region. *Antimicrob Agents Chemother* 44(10):2905–2907.
18. Strapoc D, et al. (2008) Methane-producing microbial community in a coal bed of the Illinois basin. *Appl Environ Microbiol* 74(8):2424–2432.
19. An D, et al. (2013) Metagenomics of hydrocarbon resource environments indicates aerobic taxa and genes to be unexpectedly common. *Environ Sci Technol* 47(18): 10708–10717.
20. Martinez A, Bradley AS, Waldbauer JR, Summons RE, DeLong EF (2007) Proteorhodopsin photosystem gene expression enables photophosphorylation in a heterologous host. *Proc Natl Acad Sci USA* 104(13):5590–5595.
21. Arato C, Pye EK, Gjennestad G (2005) The lignol approach to biorefining of woody biomass to produce ethanol and chemicals. *Appl Biochem Biotechnol* 121-124:871–882.
22. Arfi Y, et al. (2013) Characterization of salt-adapted secreted lignocellulolytic enzymes from the mangrove fungus *Pestalotiopsis* sp. *Nat Commun* 4:1810.
23. Yakovlev IA, et al. (2013) Genes associated with lignin degradation in the polyphagous white-rot pathogen *Heterobasidion irregulare* show substrate-specific regulation. *Fungal Genet Biol* 56:17–24.
24. Lalucat J, Bennasar A, Bosch R, García-Valdés E, Palleroni NJ (2006) Biology of *Pseudomonas stutzeri*. *Microbiol Mol Biol Rev* 70(2):510–547.
25. Cha JS, Cooksey DA (1991) Copper resistance in *Pseudomonas syringae* mediated by periplasmic and outer membrane proteins. *Proc Natl Acad Sci USA* 88(20):8915–8919.
26. Hoegger PJ, Kilaru S, James TY, Thacker JR, Kües U (2006) Phylogenetic comparison and classification of laccase and related multicopper oxidase protein sequences. *FEBS J* 273(10):2308–2326.
27. Ahmad M, et al. (2011) Identification of DypB from *Rhodococcus jostii* RHA1 as a lignin peroxidase. *Biochemistry* 50(23):5096–5107.
28. He S, et al. (2013) Comparative metagenomic and metatranscriptomic analysis of hindgut paunch microbiota in wood- and dung-feeding higher termites. *PLoS ONE* 8(4):e61126.
29. Shapiro BJ, et al. (2012) Population genomics of early events in the ecological differentiation of bacteria. *Science* 336(6077):48–51.
30. Polz MF, Alm EJ, Hanage WP (2013) Horizontal gene transfer and the evolution of bacterial and archaeal population structure. *Trends Genet* 29(3):170–175.
31. Oliver KM, Degnan PH, Hunter MS, Moran NA (2009) Bacteriophages encode factors required for protection in a symbiotic mutualism. *Science* 325(5943):992–994.
32. Cordero OX, Polz MF (2014) Explaining microbial genomic diversity in light of evolutionary ecology. *Nat Rev Micr* 12(4):263–273.
33. Lee S, Hallam SJ (2009) Extraction of high molecular weight genomic DNA from soils and sediments. *J Vis Exp* (33):e1569.
34. Mewis K, et al. (2013) Biomining active cellulases from a mining bioremediation system. *J Biotechnol* 167(4):462–471.