Title

Observing Biosynthetic Activity Utilizing Next Generation Sequencing and the DNA Linked Enzyme Coupled Assay

Journal

ACS Chemical Biology, 12(1)

ISSN

1554-8929

Authors

de Raad, Markus
Modavi, Cyrus
Sukovich, David J
et al.

Publication Date

2017-01-20

DOI

10.1021/acschembio.6b00652

Peer reviewed

ACS chemical biology

Articles

pubs.acs.org/acschemicalbiology

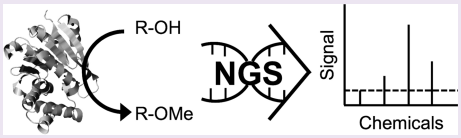
Observing Biosynthetic Activity Utilizing Next Generation Sequencing and the DNA Linked Enzyme Coupled Assay

Markus de Raad,[†] Cyrus Modavi,[†] David J. Sukovich, and J. Christopher Anderson*

Department of Biological Engineering, Synthetic Biology Institute, University of California, Berkeley, Berkeley, California 94704, United States

Supporting Information

ABSTRACT: Currently, the identification of new genes drastically outpaces current experimental methods for determining their enzymatic function. This disparity necessitates the development of high-throughput techniques that operate with the same scalability as modern gene synthesis and sequencing technologies. In this paper, we demonstrate the versatility of the recently reported DNA-Linked Enzyme-Coupled Assay (DLEnCA) and its ability to support high-throughput data acquisition through next-generation sequencing (NGS). Utilizing methyltransferases, we highlight DLEnCA's ability to rapidly profile an enzyme's substrate specificity, determine relative enzyme kinetics, detect biosynthetic formation of a target molecule, and its potential to benefit from the scales and standardization afforded by NGS. This improved methodology minimizes the effort in acquiring biosynthetic knowledge by tying biochemical techniques to the rapidly evolving abilities in sequencing and synthesizing DNA.



The pace of enzyme function elucidation lags far behind the rate of gene and protein sequence discovery.[1] Currently, only a small fraction of all known and predicted enzymes have had their computationally predicted function confirmed or substrate range identified in any detail.[2] For example, about 20% of entries in the Enzyme Commission classification are "orphan enzymes" (known enzyme activities lacking an associated protein sequence);[3] furthermore, the COMBREX project determined that only 0.4% of the 3.3 million proteins identified from the genome sequencing of microbes have had their computationally predicted function confirmed.[4] Such a gap between empirical data and the number of deposited gene sequences can only increase with mining of metagenomic sequences.[5,6]

The need for empirically determined function is especially important for the generation of novel biosynthetic pathways in metabolic engineering. One factor hindering the high-throughput characterization of enzymes, or related pursuits in enzyme engineering, is the available suite of biochemical monitoring methods. Liquid or gas chromatography coupled to mass spectrometry (LC/GC-MS) continues to be the prevailing (and, in many cases, the only) method available for detecting a chemical transformation. However, LC/GC-MS methods are often laborious and low-throughput and therefore severely limit the throughput that can be experimentally achieved. For this reason, various protocols that indirectly monitor the consumption of a reaction cofactor—through leveraging phenomena such as spectrophotometric shifts,[7−9] colorimetric/fluorescent indicators,[10−13] or bioluminescence[14]—have been developed.

Although powerful, the current suite of coupled assays has drawbacks. The first is a frequent need to use reagents that are not commercially available *en bulk*. Specialized and expensive reagents present cost barriers for scaling an assay past the hundreds of samples range. Compensating by reducing reaction volumes is only possible if the signal's limit of detection and precision are not compromised. The second drawback is that typical enzyme coupled assays utilize a cyclical/catalytic detection chemistry that increases signal variability and propensity for false positives. Last, almost all current non-LCMS biochemistry methods are based on utilization of purified proteins, which can be a major hurdle in throughput.

To address these limitations, we previously developed the DNA-Linked Enzyme-Coupled Assay (DLEnCA) methodology,[15] which is compatible with purified *in vitro* transcription-translation (TxTl) mixtures and records enzymatic activity as a DNA modification in a linear/stoichiometric fashion. This paradigm enables a flexible setup where signal readout—in a form lacking a cyclic/catalytic component—can be tied to any method of detecting DNA cleavage: gel electrophoresis, quantitative PCR, Förster resonance energy transfer (FRET) using a DNA-hairpin probe, and, in the work here, next generation sequencing (NGS). As such, while the individual reactions must still be compartmentalized by some method, the leveraging of NGS technology enables the output of hundreds of reactions to be simultaneously processed using a single piece of equipment.

The general concept of DLEnCA (as illustrated in Figure 1) is based on detecting the depletion of a cofactor pool utilized by both the substrate-modifying enzyme of interest and a DNA-modifying protein. First, an enzyme is either generated *in situ* with a TxTl reaction or added as purified recombinant protein and then incubated with substrate and cofactor. If the enzyme is

Received: August 1, 2016
Accepted: November 18, 2016
Published: November 18, 2016

DOI: 10.1021/acschembio.6b00652
ACS Chem. Biol. 2017, 12, 191−199

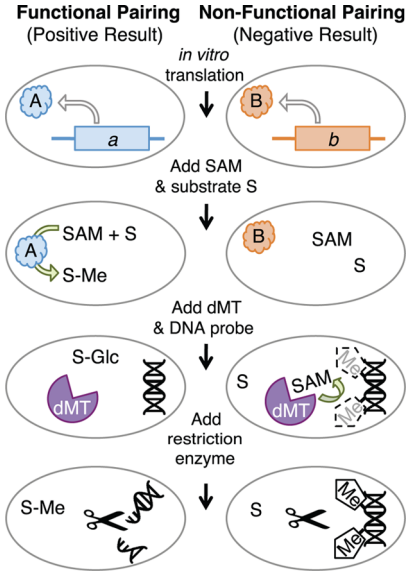


Figure 1. DLEnCA overview. A positive or functional pairing of methyltransferase A and substrate S results in fluorescence *via* probe cleavage. Alternatively, a negative or nonfunctional pairing of methyltransferase B and substrate S does not consume SAM cofactor; the probe is modified, protected from digestion, and remains unable to provide a signal (such as having fluorescence remain quenched *via* FRET or being unable to accept a sequencing adaptor). The first (*in vitro* translation) step is only for cases where purified protein is not being directly used. Technical abbreviations used: S = substrate; SAM = S-adenosylmethionine; dMT = DNA methyltransferase; Probe = DNA FRET- or NGS-probe oligonucleotide; Me = methyl group.

active against the substrate, it will consume the cofactor; this will preclude modification of a DNA probe when the DNA-modifying enzyme is subsequently added. Conversely, if the enzyme failed to show activity against the substrate, the DNA-modifying enzyme will utilize the cofactor to "protect" a restriction site on the DNA probe. In this way, enzyme activity is converted into a DNA modification that can then be read out using restriction endonucleases. In FRET-DLEnCA, a positive signal is given by DNA cleavage that separates a fluorophore from a quencher (yielding a classical fluorometric signal); in NGS-DLEnCA, the DNA cleavage instead allows a barcoded DNA probe to be "tagged" with NGS adaptors that enable generation of a sequencing read.

By exploiting the DLEnCA workflow flexibility, it is further possible to establish an indirect assay for sensing the production of a target molecule (or class of molecules) from either a biosynthetic enzyme or a cluster of enzymes. We refer to this implementation as pathway-DLEnCA (pDLEnCA). Although DLEnCA can only monitor cofactors that can be coupled to a DNA modification reaction, the specific reactions currently detectable with the assay (glucosylation and methylation) are often terminal tailoring steps within biosynthetic pathways.[16] Additionally, the specificity of these transferases ranges from being highly selective to broadly promiscuous; it is often possible to find or engineer an enzyme with activity against a substrate if an appropriate nucleophile is present in the molecule.[17,18]

In this study, we have adapted the DLEnCA protocol for profiling S-adenosylmethionine (SAM)-dependent methyltransferases to expand the scope of accessible chemistries. Beyond demonstrating the extensibility of DLEnCA as a generic method for detecting enzymatic formation of a target chemical, we have also illustrated how an enzymatic assay can be coupled to a deep-sequencing read-out.

RESULTS AND DISCUSSION

Our work is motivated by the need for a scalable methodology that would enable both the profiling of uncharacterized enzymes and the prototyping of enzyme combinations for building biosynthetic pathways in a high-throughput manner. Noting the prominence of tailoring reactions, along with the promiscuity inherent in some of these enzymes, we chose to test whether such reactions could be utilized as a biosensor system. SAM-dependent methyltransferases were chosen as a model system because of their large diversity in terms of family size and the range of chemical targets.[18]

Adapting DLEnCA for Detecting Methylation. Taking advantage of the S-adenosyl methionine (SAM)-dependent nature of restriction-modification systems, we developed a DLEnCA scheme based on the depletion of the common cofactor SAM (Figure 1). For this variant of DLEnCA, DNA-probe protection is accomplished with the *Eco*RI DNA methyltransferase (M.*Eco*RI); the cognate *Eco*RI restriction enzyme (R.*Eco*RI) serves as the enzyme detecting whether the probe was protected or not. In FRET-DLEnCA, the probe is a hairpin forming oligonucleotide conjugated with a 5′ fluorophore and 3′ quencher; in NGS-DLEnCA, the probe is a double-stranded DNA carrying a unique barcode. Digestion by R.*Eco*RI leads to fluorescence or downstream compatibility with sequencing adaptors, respectively.

With the basic protocol framework established, we next identified functional parameters for the assay and conducted preliminary validation reactions (Supporting Information Figure S1). In brief, we found working concentrations of 10 μM of SAM and a 0.5 μM concentration of DNA FRET-probe to be an optimum trade-off between detection sensitivity and robustness against noise from potential enzymatic (and nonenzymatic) hydrolysis of the cofactor; additionally, we noted that our choice of DMSO solvent slowed down degradation of the FRET-probe by an unknown component of the TxTl kit (Supporting Information Figure S2). For NGS-DLEnCA, which operates on a slightly different principle, we also found using 10 μM of SAM and 0.5 μM of DNA barcode to be sufficient. Furthermore, we identified that a 101-cycle protocol of digestion/ligation (for adding the sequencing adaptor) followed by 22 cycles of PCR provided sufficient and nonsaturating quantities of DNA for MiSeq reactions (data not shown).

Leveraging DLEnCA for Profiling Substrate Specificity. To demonstrate the potential of DLEnCA's capabilities and its utilities to researchers, we undertook a benchtop scale analysis of the well-characterized human soluble catechol methyltransferase, *Hs*S-COMT,[19] using FRET-DLEnCA (Figure 2a). To ensure extensibility, the initial substrate specificity testing was also conducted on eight other previously characterized methyltransferases (Supporting Information Figure 3). Our choices covered three of the four possible chemistries (i.e., O-, N-, and S- methylation), and also spanned both eukaryotic and prokaryotic sources. All known substrate-enzyme pairs and obtained positive signals were further assayed by LC-MS (Supporting Data). Overall, DLEnCA was found to have excellent correlation with previous biochemical characterization.

Adapting DLEnCA for Analyzing Relative Turnover Rate. Complementing the ability to qualitatively profile

DOI: 10.1021/acschembio.6b00652
ACS Chem. Biol. 2017, 12, 191−199

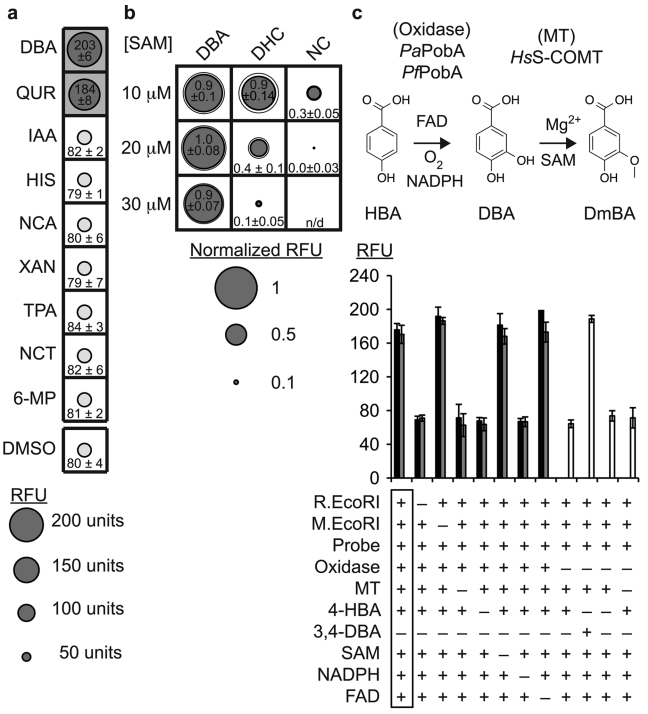


Figure 2. Characterization of *Hs*S-COMT using different modalities of DLEnCA. (a) Substrate specificity heatmap generated with FRET-DLEnCA. Boxes shaded gray represent previously known substrate−enzyme pairs; gray circles indicate reactions that showed significance of $p = 0.01$ against the DMSO blank, as determined by a two-sided Welch's t test on the raw RFU values. All data points represent the average of three technical replicates, with standard deviations too small to illustrate in the figure. Values within each box represent the average RFU value ± standard deviation from three technical replicates. (b) kFRET-DLEnCA output for HsS-COMT against three substrates. Final Min−Max normalized RFU values are based on reacting three different HsS-COMT substrates, each having a turnover rate differing by an order of magnitude, against a titration of SAM and comparing against the appropriate "no chemical" and "no M.EcoRI" reactions. Solid circles present the average of three replicates, and the outer rings represent the standard deviation after accounting for error propagation from the normalization. Values within each box represent the Min−Max normalized RFU values ± propagated standard deviation. (c) Schematic of the pathway reaction followed (Oxidase) and the detecting methyltransferase step (MT), with drop out reaction results below. The tDNA, chemical, SAM, and probe concentrations were 50 ng/reaction/gene, 1 mM, 10 μM, and 0.5 μM, respectively. Data are provided as the average (raw) RFU values, with error bars representing the standard deviation of three technical replicates. Black bars correspond to *Pa*PobA oxidase. Gray bars correspond to *Pf*PobA oxidase, and white bars refer to control runs with only *Hs*S-COMT. Chemical abbreviations used: DBA = 3,4-dihydroxybenzoic acid; QUR = quercetin; IAA = indole-3-acetic acid; HIS = histamine; NCA = nicotinamide; XAN = xanthosine; TPA = tryptamine; NCT = (−)-nicotine; 6-MP = 6-mercaptopurine; DHC = 6,7-dihydroxycoumarin; 4-NC = 4-nitrocatechol; HBA = 4-hydroxybenzoic acid; DmBA = 4-hydroxy-3-methoxybenzoic acid. Technical abbreviations used: R.*Eco*RI = *Eco*RI restriction endonuclease; M.*Eco*RI = *Eco*RI DNA methyltransferase; Probe = DNA FRET-Probe; SAM = S-adenosylmethionine; NADPH = reduced nicotinamide adenine dinucleotide phosphate; FAD = flavin adenine dinucleotide; (+) = component present in reaction; (−) = component absent from reaction.

substrate specificity, we also established that DLEnCA can be used to analyze an enzyme's relative activity against various substrates. We refer to this implementation as kinetic-DLEnCA (kDLEnCA). This secondary profiling modality is accessed by slightly modifying the DLEnCA protocol to expand the assay's dynamic range. The cornerstone theory is that at saturating substrate concentrations, turnover rate is the limiting factor in product formation (and, hence, substrate/cofactor removal). Thus, at increasing concentrations of SAM, slower rates of catalysis will be unable to sufficiently deplete the cofactor pool (within the allotted time) past the threshold required for probe protection.

Visualizing such turnover differences is accomplished using a titration of SAM in combination with a shortened incubation time and reduced amounts of enzyme. For *Hs*S-COMT, preliminary dilution experiments identified that a 30 min reaction period using only 5% (v/v) of the TxTl-generated protein was sufficient to generate a difference in the degree of probe protection observed between two substrates previously characterized with differing catalysis rates: 3,4-dihydroxybenzoic acid and 4-nitrocatechol.[20] By using the relative difference in signal as a "calibration" for rate differences, we then assayed 6,7-dihydroxycoumarin, as a test substrate. This chemical is known to be methylated by *Hs*S-COMT at an intermediate rate relative to the two chosen "calibration" chemicals.[20] By benchmarking the assay reference substrates, we were able to distinguish the relative turnover rate of another substrate (Figure 2b). In this way, it becomes possible to rapidly place all substrates of a promiscuous enzyme into a relative hierarchy once at least two kinetically different chemicals have been identified (most likely through normal enzymology techniques).

Detecting Compound Biosynthesis with DLEnCA. With a methodology in place to identify and characterize terminal methyltransferase reactions, we last verified the ability to utilize such enzymes to observe a putative biosynthetic pathway. The pathway-DLEnCA (pDLEnCA) setup is based on expressing a methyltransferase of known specificity with a variable set of putative biosynthetic enzymes. The major modification to DLEnCA is how the methyltransferase's substrate must now be generated from a provided precursor chemical and auxiliary cofactors (such as NADPH and FAD). From here, the assay would once again directly feed into the SAM-dependent nature of the probe-based output.

The model system chosen to demonstrate this concept was the production of 3,4-dihydroxybenzoic acid from 4-hydroxybenzoic acid through the activity of an oxidase gene involved in aromatic metabolite catabolism. Specifically, we chose the reported *Pseudomonas aeruginosa* gene (*Pa*PobA)[21] and its close homologue from *P. protegens Pf-5* (*Pf*PobA). Strains of *P. protegens* (previously, *P. fluorescens*)[22] have previously had their *pobA* gene confirmed, but not for the one we choose to work with. The selection of this reaction was also motivated by the fact that it is an oxidative bond formation reaction—one of the major reaction types in metabolism—and also involves an enzyme utilizing multiple cofactors. As illustrated by our results (Figure 2c and LC-MS Supporting Data), FRET-pDLEnCA is able to monitor biosynthesis of the target molecule. Furthermore, all controls yielded the correct outcomes, except for the FAD dropout. For this one unexpected result, we suspect that the PURExpress kit may contain traces of FAD that are sufficient to promote catalysis. Personal communication with New England Biolabs found that they measure for neither the presence nor absence of this chemical. However, considering that this behavior was seen in both oxidoreductase homologues tested, along with the fact that the NADPH

DOI: 10.1021/acschembio.6b00652
ACS Chem. Biol. 2017, 12, 191−199

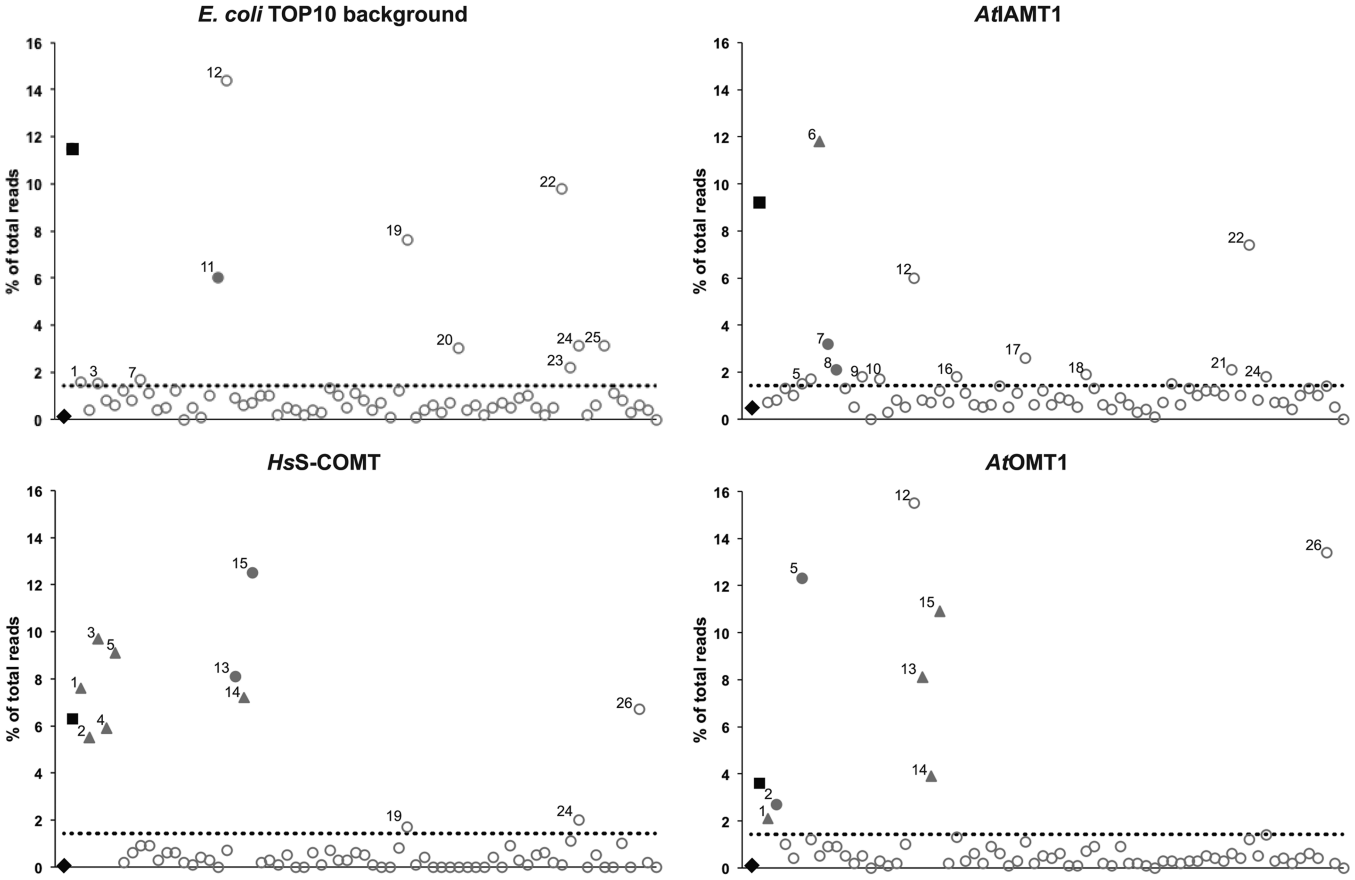


Figure 3. Substrate specificity mapping using NGS-DLEnCA. Three recombinant O-methyltransferases (*At*IAMT1, *Hs*S-COMT and *At*OMT1) and *E. coli* TOP10 background were screened against 68 chemicals. Tests were conducted with three biological replicates. The results were normalized, as described in the Methods section, so that the scatter plots measure the percentage of sequence reads per chemical for each individual methyltransferase or *E. coli* TOP10 background. Dashed line indicates the average percentage of sequence reads per chemical. Triangles represent known substrates; closed circles/triangles represent substrate methylation confirmed by LC-MS. Closed square = positive control; closed diamond = negative control.

cofactor and gene dropouts consistently lack activity, we are confident in attributing 3,4-dihydroxybenzoic acid production to the enzymes under study.

Improving Throughput of DLEnCA: Coupling to Next Generation Sequencing. In order to improve the throughput of DLEnCA, we developed a DLEnCA assay variant based on next generation sequencing (NGS) as a read-out. As with all formulations of DLEnCA, a positive signal is read off as cleavage of the unprotected probe. Unique to NGS-DLEnCA are some modifications and additional processing steps that are required for making the output compatible with deep-sequencing. At the most fundamental level, we chose to utilize recombinantly expressed and purified methyltransferases instead of TxTl-generated protein. This was motivated by the fact that the number of reactions undertaken and the current costs of purified TxTl would together have been prohibitively expensive. Thus, methyltransferases were individually combined during the reaction with a mixture containing chemical and a chemical-specific DNA probe. These probes consisted of a linear dsDNA molecule containing the left-edge sequencing adaptor (for us, the illumina P5 sequence), a barcode unique to each chemical (to facilitate analysis after sequencing, as based on the method of Wetmore *et al.*[23]), and an *Eco*RI restriction site (used as the DNA-based output of DLEnCA).

Beyond these basic protocol changes, the additional processing steps of NGS-DLEnCA begin with quenching and pooling all reactions involving a particular methyltransferase after the M.*Eco*RI treatment. Next, the DNA probes in each pool are processed for deep sequencing by enzymatic digestion with *Eco*RI, so as to cleave unprotected probes representing a positive signal. This is followed by ligating on the right-edge sequencing adapter (the illumina P7 sequence), thereby making only barcodes associated with chemical substrates of the methyltransferase compatible with the NGS reaction. These products were enriched for using a PCR amplification step, which was also utilized to add an indexing sequence unique to that methyltransferase (Supporting Information Figure S4). Finally, the processed pools for each methyltransferase are purified, combined, and subjected to a single run of deep sequencing.

To test the ability of NGS-DLEnCA to monitor enzymatic reactions, we chose to focus on the substrate specificity of three O-methyltransferases—human soluble catechol methyltransferase (*Hs*S-COMT),[19] *Arabidopsis thaliana* flavonol 3′-O-methyltransferase (*At*OMT1),[24] and *Arabidopsis thaliana* IAA carboxyl methyltransferase 1 (*At*IAMT1)[25]—against a 68-member chemical library that contained known and probable substrates (Supporting Information Table S1). After subsequent DLEnCA-NGS, we used an Illumina MiSeq sequencing instrument to identify methyltransferases capable of substrate methylation (Figure 3).

DOI: 10.1021/acschembio.6b00652
ACS Chem. Biol. 2017, 12, 191−199

We were able to identify the previously reported substrates for all three O-methyltransferases using DLEnCA-NGS, and all were confirmed by LC-MS (Table 1 and Supporting Information). Additionally, we were able to identify several previously unidentified substrates. *At*IAMT1 showed activity with indole-3-propionic acid [7] and indole-3-butyric acid [8], both similar to the known substrate indole-3-acetic acid [6]. *Hs*S-COMT methyltransferase showed activity with the flavonoids myrcetin [13] and luteolin [15], both similar to the known flavonol substrate quercetin [14]. *At*OMT1 methyltransferase showed activity with protocatechuic acid [2] and 4-nitrocatechol [5], which mimic the catechol region of quercetin [14] acted upon by *At*OMT1. Last, both *Hs*S-COMT and *At*OMT1 showed activity with crocin [26], a carotenoid, with no similarities to previous reported substrates. However, we were unable to confirm the methylation of crocin by either *Hs*S-COMT or *At*OMT1 using our LC-MS setup.

Table 1. List of Known and Identified Substrates for Each Individual Methyltransferase and *E. coli* TOP10 Background Confirmed by Both DLEnCA-NGS and LCMS (K = Known Substrate; O = Novel Substrate)

compound	*E. coli* TOP10 background	*At*IAMT	*Hs*S-COMT	*At*OMT1
[1] caffeic acid			K	K
[2] 3,4-dihydroxy-benzoic acid			K	O
[3] L-DOPA			K	
[4] dopamine			K	
[5] 4-nitrocatechol			K	O
[6] indole-3-acetic acid		K		
[7] indole-3-propionic acid		O		
[8] indole-3-butyric acid		O		
[11] curcumin	O			
[13] myricetin			O	K
[14] quercetin			K	K
[15] luteolin			O	K

A consequence of our purification setup was the copurification of *E. coli* TOP10 proteins with the recombinant methyltransferases (Supporting Information Figure S5). As a control, the methylation activity of the coeluted *E. coli* proteins was determined. To obtain the coeluted proteins, nontransformed *E. coli* TOP10 cell lysate was subjected to the affinity purification protocol. These coeluted *E. coli* TOP10 proteins generated a NGS signal for multiple substrates, including curcumin [11], kaempferol [12], 2-thiobarbituric acid [19], 2-amino-5-(4-methoxyphenyl)-1,3,4-thiadiazole [20], menadione [22], β-carotene [23], biotin [24], and 4-amino-4H-1,2,4-triazole [25]. The utilized LC-MS setup was only able to detect methylation activity of the coeluted *E. coli* proteins toward curcumin (Supporting Information); other chemicals yielded either no or inconclusive results.

As a result of this observation, any hits found for both *E. coli* TOP10 background and the purified methyltransferases were considered *E. coli* background-positives (kaempferol [12], 2-thiobarbituric acid [19], menadione [22], and biotin [24]). For *At*IAMT1, the limited sequence counts for α-lipoic acid [9], phenol [10], N-acetylserotonin [16], xanthosine [17], 2-mercapto-5-nitrobenzimidazole [18], and 2-mercapto-5-nitrobenzimidazole [21] led us to consider these compounds as background activity. Considering all these facts, NGS-DLEnCA was able to confirm all previously reported substrates and identify five novel substrates for the three methyltransferases assayed.

Experimental Considerations. The work presented here builds on our previous report of the DLEnCA methodology (as originally implemented for glucosyltransferases)[15] and adapts the technique toward another major clade of transferases. Complimenting this, we have demonstrated the extensibility and flexibility of the DLEnCA assay in the context of enzyme kinetics (kDLEnCA), pathway detection (pDLEnCA), and as a high throughput method coupled to NGS systems. These qualities together enable DLEnCA as a *first-pass platform* for enzymology, protein engineering, and pathway engineering. Furthermore, by coupling DLEnCA to a deep sequencing readout, the throughput of the assay can be improved.

For a majority of the initial FRET-DLEnCA work associated with kDLEnCA and pDLEnCA, we opted to use a TxTl system based on how such kits enable rapid combinatorial testing and prototyping. Our particular selection of PURExpress—which is composed of defined recombinant factors—was to avoid the exonuclease activity and any secondary confounding reaction found in more crude cell-free extracts.[15] However, the use of a defined system presents some drawbacks. Beyond the current costs of such systems, there is the absence of known or unknown chaperonins/post-translational maturation proteins required by some enzymes for their folding/activity. As such, there is potential for false-negatives that result from low enzymatic activity of some gene products. It is likely that many potential cases of false-negatives could be ameliorated with the addition of commercially or laboratory-produced recombinant factors, such as GroEL/ES.[26] The addition of either AdoHcy hydrolase or AdoHcy nucleosidase might also improve the general performance of the assay because these reagents have been used in previous assays to reduce product inhibition caused by spent cofactor.[10,11] Although the existing defined translation system is effective in this assay, DLEnCA could be improved with new TxTl formulations from other organisms or other preparation methods that may reduce cost, decrease noise, and increase robustness.

As previously observed by Sukovich *et al.*,[15] one of the major issues encountered with TxTl systems is the propensity for batch-to-batch variation. Although the degree of variation does not affect standard screening, this limitation is noteworthy for FRET-kDLEnCA due to the need for more standardization and dynamic range than normal DLEnCA. Consequently, we did not attempt to control protein/activity levels through titrating the amount of DNA expression cassette nor do we recommend this approach since protein yield versus DNA added is often nonlinear. Instead, it was found that diluting the completed TxTl incubation with buffer (e.g., PBS) was a decent approach for adjusting enzyme activity to match experimental parameters.

The kDLEnCA protocol has a few more specific considerations to be aware of. The ability to screen turnover is dependent on initially having at least two reference substrates with known rates for "calibrating" the dynamic range and benchmarking the analysis. In considering the batch-to-batch variance (especially when using TxTl), it is advisible to run reference substrates for any experiment meant to involve comparisons. Another point of partial concern is that previous work[27] indicates dsDNA can alter kinetic rates by interacting with chemical substrates, emphasizing the point that DLEnCA

DOI: 10.1021/acschembio.6b00652
ACS Chem. Biol. 2017, 12, 191−199

is meant only to provide a high-throughput first-pass at analysis and does not replace traditional enzyme kinetic methodologies. Furthermore, we are uncertain about the maximum dynamic-range that can be obtained in FRET-kDLEnCA; however, we do suspect that adjusting the baseline DNA probe and SAM concentration could potentially help increase the sensitivity to turnover differences, even for normal FRET-DLEnCA.

With regards to the types of chemistries that can be studied in pDLEnCA, we note that only iron−sulfur or membrane anchored enzymes (such as eukaryotic P450s) could currently be considered incompatible with the assay due to cofactor loading or solubilization issues. However, there are publications[28−31] suggesting that *in vitro* reconstitution of such systems is not impossible. Regardless, our success with flavoprotein oxidoreductases in cell-free, similar to others,[32] indicates that a large subspace of anabolic reactions can be easily reconstituted and assayed with basic TxTl systems.

Overall, DLEnCA can be viewed as a complete methodology for implementing the "design-build-test" cycles needed for the discovery of enzymes involved within a biosynthetic pathway or that function as the final tailoring step. First, where necessary, DLEnCA can be used in a forward screening capacity to identify/engineer and, where desired, characterize methyltransferases that act on a substrate or structural moiety of interest. The obtained biochemical information then enables the screening process to be reversed, whereby the obtained methyltransferase is used as a sensor element for detecting biosynthesis of the chemical substrate. In both cases, this would involve the parallel screening of thousands-member libraries for productive reactions. Such a system is supported by how the assay can be linked to high-throughput next generation sequencing (i.e., NGS-DLEnCA)—a major quality setting it apart from other assays in enzyme discovery.

Using NGS-DLEnCA, we have screened three recombinant O-methyltransferases against a small substrate library, obtaining results that match with all previous known substrates in addition to identifying seven novel substrates. By linking an enzymatic reaction to a deep sequencing read-out, NGS-DLEnCA allows the screening of massive enzyme and substrate libraries in a rapid and high-throughput manner. In order to reduce costs, we used *E. coli* TOP10 to recombinantly express the methyltransferases instead of using purified TxTl. The methyltransferases were affinity purified using a commercially available protein miniprep kit (Thermo Fisher Scientific), avoiding time-consuming preparative chromatography and dialysis. As a consequence, *E. coli* proteins and other contaminants were present, which resulted in background methylation activity. By using TxTl as a source of recombinant enzymes, purification and the presences of impurities can be avoided.

To enable the deep sequencing, NGS-DLEnCA requires linking a barcode/index to both the monitored methyltransferase and substrate. Both can be generated using conventional DNA assembly strategies, although large libraries will require large-scale DNA fabrication; with the declining cost of DNA synthesis, this will become increasingly accessible to the average research group. Furthermore, by using DNA methyltransferases that modify cytosines, the output could be made compatible with disulfide sequencing outputs if desired. Such a scenario enables more semiquantitative high-throughput data. Overall, the ability of NGS-DLEnCA to consolidate and pool many reactions significantly streamlines the discovery process while also increasing the maximum theoretical throughput that can be achieved. Most importantly, a NGS output can theoretically detect (or be made compatible) with a single-molecule event and, thus, enable drastic volume reductions without compromising the output signal.

In summary, we have expanded the DLEnCA methodology toward monitoring biosynthetic pathways and utilizing deep sequencing read-out. This makes DLEnCA an ideal platform for high-throughput gene function elucidation/characterization and for monitoring biosynthetic pathways. By linking enzymatic function to chemical modification of DNA and deep sequencing analysis, this assay can provide the solution toward the screening and characterization of large classes of known, unknown, and *de novo* designed enzymes or biosynthetic pathways. To further expand toward the level of massively scaled screening needed for next-generation biochemical production, it will be necessary to implement the work developed here with automated or microfluidic platforms.

METHODS

Materials. *Eco*RI methyltransferase, *Eco*RI-HF endonuclease, MfeI-HF endonuclease, AlwNI-HF endonuclease, S-adenosyl-methionine (SAM) solution, T4 DNA ligase, Gibson Assembly Master Mix, Phusion High-Fidelity DNA polymerase, dNTP mix, and PURExpress were purchased from New England Biolabs (Ipswich, MA, USA). B-PER Protein Extraction Reagent and HisPur Ni-NTA Resin were purchased from Thermo Fisher Scientific (Waltham, MA, USA). One Shot TOP10 chemically competent *E. coli* was purchased from Life Technologies (Carlsbad, CA, USA). L-Arabinose was purchased from Sigma-Aldrich (St. Louis, MO, USA). Water Optima-LCMS, Formic Acid Optima-LCMS, and Acetonitrile Optima-LCMS were purchased from Fisher Scientific (Waltham, MA, USA). All used chemicals are listed in Supporting Information Table S2. All oligonucleotides were purchased from Integrated DNA Technologies (Coraville, IA, USA). Sequences of all oligonucleotides and DNA templates used in this study are listed in the Supporting Information. Fluorescence measurements for all experiments were obtained with a Tecan Safire2, using CoStar (black, transparent bottom) 384-well plates covered with a Bio-Rad PCR-plate transparent seal.

FRET-DLEnCA Probe. The DNA probe was synthesized by Integrated DNA Technologies as a 60 nucleotide hairpin forming oligo containing a 5′ Cy3 fluorophore and 3′ Iowa Black Fluorescent Quencher, with an *Eco*RI site located 4 bp from the hairpin's terminus (Supporting Information Table S3). The probe was resuspended and stored in water as a standard 100 μM stock solution. Prior to use, aliquots of the probe were diluted 10-fold and adjusted to match IDT's Duplex Buffer (100 μM potassium acetate and 30 μM HEPES buffer) before being reannealed.

Cloning of Plasmids for TxTl Expression of Methyltransferases. All methyltransferase genes used in TxTl were cloned into the pET-15b expression vector between the NcoI and *Bam*HI sites, using either the NcoI/*Bam*HI enzymes themselves or BsaI (a type IIs restriction enzyme) to generate compatible cohesive ends. The sequences of the human and *Arabidopsis* methyltransferases were amplified from mesenchymal stem cell and leaf cDNA, respectively. *Coffea arabica*'s xanthosine methyltransferase was isolated from the iGEM repository part BBa_K801070. All bacterial genes were isolated from genomic DNA. All PCR primers are provided in Supporting Information Table S3. In all cases, silent point mutations were made where necessary to facilitate cloning. All vectors were sequence verified. Information and major literature citations relevant to the genes utilized in study are listed in Supporting Information Tables 4 and 5.

Production of Template DNA for TxTl Expression of Methyltransferase. pET vectors were used as templates for generating the gene expression cassettes (DNA template) utilized in all TxTl experiments. This was accomplished using primers bracketing the region around the T7 promoter and terminator. Final PCR product purity was assessed using gel electrophoresis. For any gel

DOI: 10.1021/acschembio.6b00652
ACS Chem. Biol. 2017, 12, 191−199

purifications of TxTl template DNA, direct use of an intercalating agent on the target DNA (for gel visualization) was avoided, as it is reported in the PURExpress manual that these agents can impact cell-free performance. All DNA was standardized to a working concentration, as reported in the sections below.

Cloning of Plasmids for Recombinant Expression of Methyltransferases. Genetic constructs for recombinant expression of methyltransferases were fabricated as clonal plasmids from pre-existing DNA. All genes were cloned into the pBAD/myc-His A vector (Life Technologies) using Phusion polymerase. Each construct contained the methyltransferase under the transcriptional control of a pBAD promoter and a TrrnB terminator. The PCRs were carried out using a PTC-200 Peltier Thermo Cycler at the following temperatures: 98 °C and 1 min followed by 35 cycles at 98 °C and 20 s, 60 °C and 30 s, and 72 °C for 45 s, with a final single extended elongation phase at 72 °C for 10 min. The amplified DNA was gel purified and added to Gibson Assembly Master Mix (New England Biolabs) and assembled using the recommended protocol. Plasmids were transformed into chemically competent *E. coli* strain DH10B and sequence verified.

Cloning and Production of Probe DNA for NGS-DLEnCA. Genetic constructs containing NGS-DLEnCA DNA probes were fabricated as clonal plasmids (pGG001-Barcode) using Golden Gate assembly. The DNA probe contained a 43 bp common part, a 20 bp barcode, and an *Eco*RI site (Supporting Information Tables S3 and S5). In total, 70 individual plasmids with unique barcodes were sequence verified and selected for this study (Supporting Information Table S4). Linear probe DNA (604 bp) was amplified 235 bp upstream of the common part and 299 bp downstream of the *Eco*RI site. Linear probe DNA was amplified using Phusion polymerase with primers P09 and P10 using the following temperatures: 98 °C and 1 min followed by 35 cycles at 98 °C and 20 s, 55 °C and 30 s, and 72 °C for 50 s, with a final single extended elongation phase at 72 °C for 10 min. DNA was purified using DNA Clean & Concentrator (Zymo Research, Irvine, CA, USA) and eluted in H_2O. DNA was quantified using NanoDrop.

Cloning and Production of P7.1 Adapter for Illumina Sequencing. A genetic construct containing the P7.1 adapter was fabricated as a clonal plasmid pGG001-P7.1 using Golden Gate assembly (Supporting Information Table S5). The P7.1 contained a MfeI site, the Illumina P7 adapter, and part of the origin of replication element ColE1. The P7.1 adapter (532 bp) was digested from vector pGG001-P7.1 using MfeI-HF and AlwNI-HF. DNA of expected size was gel purified using the Zymoclean Gel DNA recovery kit (Zymo Research) and eluted in H_2O. DNA was quantified using Nanodrop and sequence verified.

FRET-DLEnCA. Operationally, FRET-DLEnCA involves iterative additions of liquids to a microtiter plate followed by a fluorescence measurement. To avoid plate evaporation during each step, the experimental wells were bordered by wells filled with 10 μL of water. Additionally, the entire plate was sealed with a transparent sealing film (Phenix) during all incubation steps. The protocol contains four sequential rounds of reagent addition each followed by incubation:

(1) *In vitro* transcription and translation. A master mix of PURE buffer A and B (at 4 μL/reaction and 3 μL/reaction, respectively) was prepared and dispensed into reaction wells (7 μL/reaction) on ice. Next, 1 μL of DNA template at 50 ng/μL was added. Reactions were then incubated at 37 °C for 3 h to produce the encoded enzymes.

(2) *In vitro* activity testing. After *in vitro* translation, 1 μL of a 80% (v/v) DMSO solution containing 10 mM chemical substrate and 100 μM SAM was added. The mixture was then incubated at the enzyme's optimum temperature for 3 h to produce the product and deplete SAM.

(3) Probe protection. Reacted mixtures were augmented with a 1.5 μL solution made from premixing DNA probe and *Eco*RI DNA methyltransferase (at a ratio of 0.5 μL of 10 μM probe and 1 μL of 40 units/μL enzyme). This was left to incubate once more at 37 °C for 3 h. After this step, the result of the reaction in step 2 is recorded as a DNA modification of the probe.

(4) Restriction analysis. Finally, 1 μL of 20 units/μL *Eco*RI endonuclease was added on ice. The plate was then incubated at 37 °C in a Tecan Safire2. Fluorescence measurements were collected every 10 min for a total of 500 min with 5 nm bandwidth filtering set to 550/564 nm wavelengths for excitation/emission. The 2 h (120 min) time point was chosen as the final reported value for all experiments. In all cases, drop-out controls were included in which one of the reagents was replaced by an equal volume of water.

Kinetic and Pathway Experimental Modifications. For the kinetic FRET-DLEnCA experiments, the protocol was altered in four ways. First, step 1 involved a master-mix of PURExpress with DNA maintaining the same PURE buffer A, PURE buffer B, and DNA proportions of 4:3:1. The mix was incubated for 3 h at 37 °C before being diluted to 5% (v/v) with 1 × PBS (pH 7.5) and 8 μL aliquots were transferred to wells. Second, step 2 involved varied SAM concentrations of 100−300 μM corresponding to final solution values between 10 and 30 μM. Third, the incubation time in step 2 was reduced to 30 min. Fourth, the dispensing steps in 2 and 3 were performed on ice to keep all the reactions synchronized. For the pathway FRET-DLEnCA experiments, two noteworthy changes were made to the original protocol: 50 ng of each of the two genes were used in each reaction, and FAD/NADPH (both at 5 mM) was added along with the SAM.

Recombinant Production of Methyltransferases. Plasmids containing methyltransferases under the transcriptional control of a pBAD promoter and a TrrnB terminator were used. Methyltransferases were expressed in TOP10 *E. coli* cultures and induced using 0.2% (w/v) arabinose. The cells were grown at 37 °C in 50 mL of 2YT medium with 100 μg/mL ampicillin to an OD600 of 0.4−0.8. The cells were cooled on ice for 5 min and induced for 16−20 h at 16 °C. The cells were harvested by centrifugation (5000*g*, 5 min, 4 °C), and stored at −20 °C. The bacterial pellets obtained were thawed at RT; resuspended in 10 mL/g wet pellet B-PER Protein Extraction Reagent containing 0.1 mg mL^{-1} lysozyme, DNase I (1 U/mL), and RNase A (10 ug/mL); and incubated at RT for 10 min on a horizontal shaker. Cellular debris was removed by centrifugation (10 000*g*, 15 min, 4 °C). Ni-NTA agarose resin was added to the supernatant (0.15 mL per 50 mL of culture), and the solution was incubated at RT for 15 min. The protein−resin mixture was collected using centrifugation (1200*g*, 5 min, RT), resuspended in 0.25 mL of wash buffer (50 mM Tris, 300 mM NaCl, 25 mM imidazole, 10% (v/v) glycerol, pH 6.8), and loaded onto a Pierce spin column. To remove unbound proteins, 0.5 mL of wash buffer was added to the spin column, incubated for 5 min, and centrifuged at 1000−2000*g* for 5 min at RT and repeated twice for a total of three washes. To elute the His-tagged proteins, 0.15 mL of elution buffer (50 mM Tris, 300 mM NaCl, 200 mM imidazole, 10% glycerol, pH 6.8) was added to the spin column, incubated for 5 min, and centrifuged at 1000−2000*g* for 5 min at RT and repeated twice for a total of three elutions. Protein purity was verified using SDS-PAGE.

NGS-DLEnCA Assay. Eluted recombinant methyltransferases were directly used without normalization. All reactions were performed in triplicate. One microliter of eluted methyltransferase was added to SAM (40 nM), DNA probe (1 nM), and chemical (100 μM, listed in Supporting Information) in a buffered solution (50 mM potassium acetate, 20 mM Tris-acetate, 10 mM magnesium acetate, 100 μg/mL BSA, pH 7.9). The final reaction volume was 10 μL. For both negative and positive controls, no chemical was added. Reactions were incubated for 3 h at 37 °C for *E. coli* TOP 10 background and *Hs*S-COMT, at 30 °C for *At*IAMT1 and *At*OMT1. Four units of *Eco*RI DNA methyltransferase (M.*Eco*RI) were added, followed by a 3 h incubation at 37 °C. To the positive control, no M.*Eco*RI was added. Reactions were terminated by the addition of ADB buffer (Zymo Research) and pooled for each individual methyltransferase. Probe DNA was recovered using DNA Clean & Concentrator and eluted in H_2O.

Illumina Sequencing. Sequencing adapter P7.1 was ligated onto the recovered and individually pooled nonmethylated probe DNA. Probe DNA and adapter P7.1 were digested with EcoRI and MfeI, respectively, in order to create compatible sticky ends. Digestion and ligation reactions were carried out in a serial manner by cycling between 37 and 16 °C, as commonly used in DNA assembly methods, including "Golden Gate" assembly.[33] To the recovered and purified

DOI: 10.1021/acschembio.6b00652
ACS Chem. Biol. 2017, 12, 191−199

probe DNA form the DLEnCA assay, 10 mM ATP, 15 units of *Eco*RI-HF, 15 units of MfeI-HF, 1200 units of T4 DNA ligase, and 1050 ng of P7.1 adapter were added. Total volume was 35 μL, and reactions were performed in CutSmart buffer (NEB). Reactions were carried out using a PTC-200 Peltier Thermo Cycler at the following temperatures: 100 cycles at 37 °C for 2.5 min and 16 °C for 5 min, with a final single step at 37 °C for 2.5 min.

Sequencing/indexing adapters were added to probe DNA using PCR amplification. Each individual methyltransferase was assigned a unique index sequence previously used by Wetmore *et al.*,[23] which was placed within the forward amplification primer (Supporting Information Table S1). Probe DNA was amplified using Phusion High-Fidelity DNA polymerase with either primer SI1 for *E. coli* TOP10, SI2 for *Hs*S-COMT, SI3 for AtIAMT, or SI4 for *At*OMT1 and SI5. The PCRs were carried out using the following temperature program: 98 °C for 30 s followed by 30 cycles at 98 °C for 10 s, 60 °C for 30 s, and 72 °C for 30 s, with a final single extended elongation phase at 72 °C for 10 min. DNA of expected size (198−201 bp) was gel purified using the Zymoclean Gel DNA recovery kit (Zymo Research) and eluted in H_2O. DNA was quantified using the Qubit (Life Technologies) according to the manufacturer's protocol. Equal amounts of eluted DNA were combined and sequenced on the Illumina MiSeq using a 150 cycle, dual read protocol.

Analysis of Illumina Data. Sequence data were analyzed using Python software. First, all sequencing reads containing the common part (42 nt) were identified and pooled. No mismatches were allowed. Next, sequencing reads containing the common part were clustered per methyltransferase group using the indexing barcodes (SI1 for *E. coli* TOP10, SI2 for *Hs*S-COMT, SI3 for *At*IAMT, or SI4 for *At*OMT1). No mismatches were allowed. Then, all reads for each individual methyltransferase group were clustered per chemical group using the chemical barcodes (Supporting Information Table S4). The percentage of reads assigned to a specific chemical group from the total number of reads for a specific methyltransferase group was calculated.

LC-MS Confirmation. To confirm the validity of DLEnCA results, determination of methylation was accomplished by means of an LC-MS system consisting of an Agilent Technologies 1200 series HPLC with an Agilent Technologies 6520 Accurate Mass qTOF MS. A Zorbax Eclipse Plus C18 guard column (4.6 cm × 12.5 cm, 5 μm packing, Agilent Technologies) connected to an Eclipse Plus C18 (4.6 mm × 100 mm i.d., 3.5-μm packing, Agilent Technologies) column was used for sample separation. Water and acetonitrile mobile phases (run at a flow rate of 500 μL/min) were supplemented with 0.1% (v/v) formic acid or 0.05% (v/v) ammonium hydroxide for the positive and negative mode runs, respectively. The elution gradient (water/acetonitrile ratio) was ramped as follows: 98:2 (0−2 min), 98:2−5:95 linearly (2−17 min), 5:95 (17−27 min), and 5:95−98−2 (27−28 min). Full scanning mode (50−750 m/z) was used for data acquisition in either positive-ion or negative-ion mode, with operation parameters set as follows: ESI probe capillary voltage, + 3.5 kV with a scan rate of 1.01 scans/second. The nebulizer gas flow rate was 7 L/min. During the analysis, two references (121.0509 m/z ($C_5H_4N_4$) and 922.0098 m/z ($C_{18}H_{18}O_6N_3P_3F_{24}$)) were continuously measured to allow constant mass correction.

For FRET- and pathway-DLEnCA, LCMS incubations contained 10 mM substrate, 3 mM SAM, and 50 ng of DNA in 10 μL PURExpress reaction. For NGS-DLEnCA, 1 μL of eluted methyltransferase was added to SAM (3.2 mM) and substrate (10 mM) in a buffered solution (50 mM potassium acetate, 20 mM Tris-acetate, 10 mM magnesium acetate, 100 μg/mL BSA, pH 7.9) for a total volume of 10 μL. After 3−16 h incubation at the optimal temperature for each enzyme, 10 μL of water was added and samples were quenched with 20 μL of ethanol. Precipitated protein was removed by centrifugation (14 000*g*, 5 min). Supernatant was added to LC-MS vials, and 10 μL of supernatant was injected onto the LC-MS per run. Compound presence was identified searching for MS parent ions.

Statistical Analyses. All data were processed manually without specialist software. For comparison between the sets of data points in Figure 2 and Supporting Information Figure S2, a two-sided Welch's *t* test was utilized so as to account for possibly unequal variances. For the error propagation after Min−Max normalization in Figure 2b, standard formulas recommended by the National Institute of Standards and Technology[34] were used. P-values chosen for each analysis are reported in the appropriate figures.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acschembio.6b00652.

Figures S1−S5 and Mass Spectra (PDF)
Tables S1−S5 (XLSX)

Accession Codes

The protein sequences for all enzymes expressed and studied are as follows: NP_009294 (*Hs*S-COMT), NP_626585 (*Sc*OMT1), NP_200227 (*At*OMT1), NP_200336 (*At*IAMT), NP_008826 (*Hs*HNMT), NP_006160 (*Hs*NNMT), AFV60449 (*Ca*XMT1), NP_006765 (*Hs*INMT), NP_000358 (*Hs*TPMT), NP_248938 (*Pa*PobA), YP_262248 (*Pf*PobA). The final DNA sequences of cloned constructs are provided in Supporting Information Table S2.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: jcanderson@berkeley.edu.

ORCID

Markus de Raad: 0000-0001-8263-9198

Author Contributions

†These authors contributed equally to this work.

Funding

Work was supported by funding from Agilent Technologies via the Synthetic Biology Institute.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The authors would like to thank the various laboratories for their generous donations of genomic DNA or cDNA that were used for obtaining the genes used in this paper: S. Li and J. Chu for human mesenchymal stem cell cDNA; L. Feldmen and K. Jiang for *Arabidopsis thaliana* leaf cDNA; M. C. Chang for *Streptomyces coelicolor* A3(2) genomic DNA; and W. Zhang for *Pseudomonas protegens* Pf-5 genomic DNA.

■ REFERENCES

(1) Bastard, K., Smith, A. A. T., Vergne-Vaxelaire, C., Perret, A., Zaparucha, A., De Melo-Minardi, R., Mariage, A., Boutard, M., Debard, A., Lechaplais, C., Pelle, C., Pellouin, V., Perchat, N., Petit, J.-L., Kreimeyer, A., Medigue, C., Weissenbach, J., Artiguenave, F., De Berardinis, V., Vallenet, D., and Salanoubat, M. (2013) Revealing the hidden functional diversity of an enzyme family. *Nat. Chem. Biol. 10*, 42−49.

(2) Gerlt, J. A., Allen, K. N., Almo, S. C., Armstrong, R. N., Babbitt, P. C., Cronan, J. E., Dunaway-Mariano, D., Imker, H. J., Jacobson, M. P., Minor, W., Poulter, C. D., Raushel, F. M., Sali, A., Shoichet, B. K., and Sweedler, J. V. (2011) The Enzyme Function Initiative. *Biochemistry 50*, 9950−9962.

(3) Sorokina, M., Stam, M., Médigue, C., Lespinet, O., and Vallenet, D. (2014) Profiling the orphan enzymes. *Biol. Direct 9*, 10.

(4) Anton, B. P., Kasif, S., Roberts, R. J., and Steffen, M. (2014) Objective: biochemical function. *Front. Genet. 5*, DOI: 10.3389/fgene.2014.00210.

(5) Banik, J. J., Craig, J. W., Calle, P. Y., and Brady, S. F. (2010) Tailoring Enzyme-Rich Environmental DNA Clones: A Source of Enzymes for Generating Libraries of Unnatural Natural Products. *J. Am. Chem. Soc. 132*, 15661−15670.
(6) Godzik, A. (2011) Metagenomics and the protein universe. *Curr. Opin. Struct. Biol. 21*, 398−403.
(7) Dorgan, K. M., Wooderchak, W. L., Wynn, D. P., Karschner, E. L., Alfaro, J. F., Cui, Y., Zhou, Z. S., and Hevel, J. M. (2006) An enzyme-coupled continuous spectrophotometric assay for S-adenosylmethionine-dependent methyltransferases. *Anal. Biochem. 350*, 249−255.
(8) Wu, Z. L., Ethen, C. M., Prather, B., Machacek, M., and Jiang, W. (2011) Universal phosphatase-coupled glycosyltransferase assay. *Glycobiology 21*, 727−733.
(9) Graves, T. L., Zhang, Y., and Scott, J. E. (2008) A universal competitive fluorescence polarization activity assay for S-adenosylmethionine utilizing methyltransferases. *Anal. Biochem. 373*, 296−306.
(10) Hendricks, C. L., Ross, J. R., Pichersky, E., Noel, J. P., and Zhou, Z. S. (2004) An enzyme-coupled colorimetric assay for S-adenosylmethionine-dependent methyltransferases. *Anal. Biochem. 326*, 100−105.
(11) Wang, C., Leffler, S., Thompson, D. H., and Hrycyna, C. A. (2005) A general fluorescence-based coupled assay for S-adenosylmethionine-dependent methyltransferases. *Biochem. Biophys. Res. Commun. 331*, 351−356.
(12) Lu, L.-M., Zhang, X.-B., Kong, R.-M., Yang, B., and Tan, W. (2011) A Ligation-Triggered DNAzyme Cascade for Amplified Fluorescence Detection of Biological Small Molecules with Zero-Background Signal. *J. Am. Chem. Soc. 133*, 11686−11691.
(13) Kumagai, K., Kojima, H., Okabe, T., and Nagano, T. (2014) Development of a highly sensitive, high-throughput assay for glycosyltransferases using enzyme-coupled fluorescence detection. *Anal. Biochem. 447*, 146−155.
(14) Hemeon, I., Gutierrez, J. A., Ho, M.-C., and Schramm, V. L. (2011) Characterizing DNA Methyltransferases With An Ultrasensitive Luciferase-Linked Continuous Assay. *Anal. Chem. 83*, 4996−5004.
(15) Sukovich, D. J., Modavi, C., de Raad, M., Prince, R. N., and Anderson, J. C. (2015) DNA-Linked Enzyme-Coupled Assay for Probing Glucosyltransferase Specificity. *ACS Synth. Biol. 4*, 833.
(16) Pickens, L. B., Tang, Y., and Chooi, Y.-H. (2011) Metabolic Engineering for the Production of Natural Products. *Annu. Rev. Chem. Biomol. Eng. 2*, 211−236.
(17) Bowles, D., Isayenkova, J., Lim, E.-K., and Poppenberger, B. (2005) Glycosyltransferases: managers of small molecules. *Curr. Opin. Plant Biol. 8*, 254−263.
(18) Struck, A.-W., Thompson, M. L., Wong, L. S., and Micklefield, J. (2012) S-Adenosyl-Methionine-Dependent Methyltransferases: Highly Versatile Enzymes in Biocatalysis, Biosynthesis and Other Biotechnological Applications. *ChemBioChem 13*, 2642−2655.
(19) Axelrod, J., and Tomchick, R. (1958) Enzymatic O-methylation of epinephrine and other catechols. *J. Biol. Chem. 233*, 702−705.
(20) Lautala, P., Ulmanen, I., and Taskinen, J. (2001) Molecular mechanisms controlling the rate and specificity of catechol O-methylation by human soluble catechol O-methyltransferase. *Mol. Pharmacol. 59*, 393−402.
(21) Entsch, B., Nan, Y., Weaich, K., and Scott, K. F. (1988) Sequence and organization of pobA, the gene coding for p-hydroxybenzoate hydroxylase, an inducible enzyme from Pseudomonas aeruginosa. *Gene 71*, 279−291.
(22) Suemori, A., and Iwakura, M. (2007) A Systematic and Comprehensive Combinatorial Approach to Simultaneously Improve the Activity, Reaction Specificity, and Thermal Stability of p-Hydroxybenzoate Hydroxylase. *J. Biol. Chem. 282*, 19969−19978.
(23) Wetmore, K. M., Price, M. N., Waters, R. J., Lamson, J. S., He, J., Hoover, C. A., Blow, M. J., Bristow, J., Butland, G., Arkin, A. P., and Deutschbauer, A. (2015) Rapid Quantification of Mutant Fitness in Diverse Bacteria by Sequencing Randomly Bar-Coded Transposons. *mBio 6*, e00306.
(24) Muzac, I., Wang, J., Anzellotti, D., Zhang, H., and Ibrahim, R. K. (2000) Functional Expression of an Arabidopsis cDNA Clone Encoding a Flavonol 3′-O-Methyltransferase and Characterization of the Gene Product. *Arch. Biochem. Biophys. 375*, 385−388.
(25) Qin, G. (2005) An Indole-3-Acetic Acid Carboxyl Methyltransferase Regulates Arabidopsis Leaf Development. *Plant Cell 17*, 2693−2704.
(26) Li, J., Gu, L., Aach, J., and Church, G. M. (2014) Improved Cell-Free RNA and Protein Synthesis System. *PLoS ONE* (Jeltsch, A., Ed.) *9*, e106232.
(27) Lin, J.-L., and Wheeldon, I. (2013) Kinetic Enhancements in DNA−Enzyme Nanostructures Mimic the Sabatier Principle. *ACS Catal. 3*, 560−564.
(28) Unciuleac, M.-C., Chandramouli, K., Naik, S., Mayer, S., Huynh, B. H., Johnson, M. K., and Dean, D. R. (2007) In Vitro Activation of Apo-Aconitase Using a [4Fe-4S] Cluster-Loaded Form of the IscU [Fe−S] Cluster Scaffolding Protein †. *Biochemistry 46*, 6812−6821.
(29) Kwon, Y.-C., Oh, I.-S., Lee, N., Lee, K.-H., Yoon, Y. J., Lee, E. Y., Kim, B.-G., and Kim, D.-M. (2013) Integrating cell-free biosyntheses of heme prosthetic group and apoenzyme for the synthesis of functional P450 monooxygenase. *Biotechnol. Bioeng. 110*, 1193−1200.
(30) Schwarz, D., Dötsch, V., and Bernhard, F. (2008) Production of membrane proteins using cell-free expression systems. *Proteomics 8*, 3933−3946.
(31) Boyer, M. E., Stapleton, J. A., Kuchenreuther, J. M., Wang, C., and Swartz, J. R. (2008) Cell-free synthesis and maturation of [FeFe] hydrogenases. *Biotechnol. Bioeng. 99*, 59−67.
(32) Quertinmont, L. T., Orru, R., and Lutz, S. (2015) RApid Parallel Protein EvaluatoR (RAPPER), from gene to enzyme function in one day. *Chem. Commun. 51*, 122−124.
(33) Engler, C., Kandzia, R., and Marillonnet, S. (2008) A One Pot, One Step, Precision Cloning Method with High Throughput Capability. *PLoS One 3*, e3647.
(34) Ku, H. H. (1966) Notes on the Use of Propagation of Error Formulas. *J. Res. Natl. Bur. Stand., Sect. C 70C*, 263−273.

DOI: 10.1021/acschembio.6b00652
ACS Chem. Biol. 2017, 12, 191−199