ACS SyntheticBiology

Letter

pubs.acs.org/synthbio

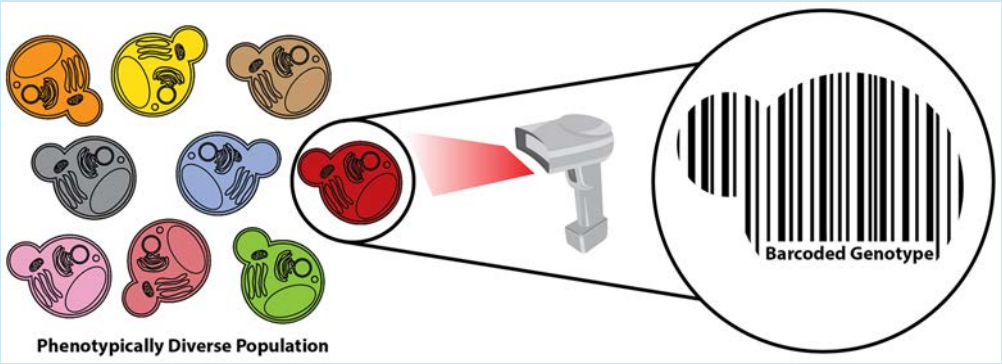
A Barcoding Strategy Enabling Higher-Throughput Library Screening by Microscopy

Robert Chen,[†,‡] Harneet S. Rishi,[†,§] Vladimir Potapov,[¶] Masaki R. Yamada,[§] Vincent J. Yeh,[§] Thomas Chow,[‡] Celia L. Cheung,[‡] Austin T. Jones,[⊥] Terry D. Johnson,[‡] Amy E. Keating,[¶,#] William C. DeLoache,*[,‡] and John E. Dueber*[,‡]

[‡]Department of Bioengineering, [§]Department of Chemical and Biomolecular Engineering, and [⊥]Department of Molecular and Cell Biology, University of California, Berkeley, Berkeley, California 94720, United States

[¶]Department of Biology, and [#]Department of Biological Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, United States

Supporting Information



ABSTRACT: Dramatic progress has been made in the design and build phases of the design−build−test cycle for engineering cells. However, the test phase usually limits throughput, as many outputs of interest are not amenable to rapid analytical measurements. For example, phenotypes such as motility, morphology, and subcellular localization can be readily measured by microscopy, but analysis of these phenotypes is notoriously slow. To increase throughput, we developed microscopy-readable barcodes (MiCodes) composed of fluorescent proteins targeted to discernible organelles. In this system, a unique barcode can be genetically linked to each library member, making possible the parallel analysis of phenotypes of interest *via* microscopy. As a first demonstration, we MiCoded a set of synthetic coiled-coil leucine zipper proteins to allow an 8 × 8 matrix to be tested for specific interactions in micrographs consisting of mixed populations of cells. A novel microscopy-readable two-hybrid fluorescence localization assay for probing candidate interactions in the cytosol was also developed using a bait protein targeted to the peroxisome and a prey protein tagged with a fluorescent protein. This work introduces a generalizable, scalable platform for making microscopy amenable to higher-throughput library screening experiments, thereby coupling the power of imaging with the utility of combinatorial search paradigms.

KEYWORDS: *synthetic biology, library screening, microscopy barcodes, coiled-coil zippers*

The inherent complexity of biological systems and our limited understanding of forward engineering principles usually require the testing of a large number of designs in order to find one that yields the desired phenotype, much less learn design principles to gain predictability.[1] Engineering cellular behaviors often demands the ability to measure a phenotype in individual cells. Accordingly, methods such as cytometry and fluorescence-activated cell sorting (FACS) have become powerful single-cell analytical tools that can be applied in high-throughput. A limitation of these methods is the requirement of their outputs of interest to be fluorescent or be able to link them to a fluorescent output. Many desirable cellular phenotypes such as motility, morphology, and subcellular localization are not currently amenable to these high-throughput strategies but can be clearly visualized by microscopy. However, the throughput of microscopy experiments is not sufficient for rapid design−build−test cycles.

To screen for these phenomena, it is often necessary to link the observable phenotype to the underlying genotype. To this end, many methods for barcoding cells for flow cytometry or microscopy have been devised, often involving the labeling of cells with unique fluorescent markers.[2−5] Two early methods of uniquely tagging cells, Brainbow and red-green-blue (RGB) labeling, use random combinations of fluorescent reporter

Received: March 27, 2015

DOI: 10.1021/acssynbio.5b00060
ACS Synth. Biol. XXXX, XXX, XXX−XXX

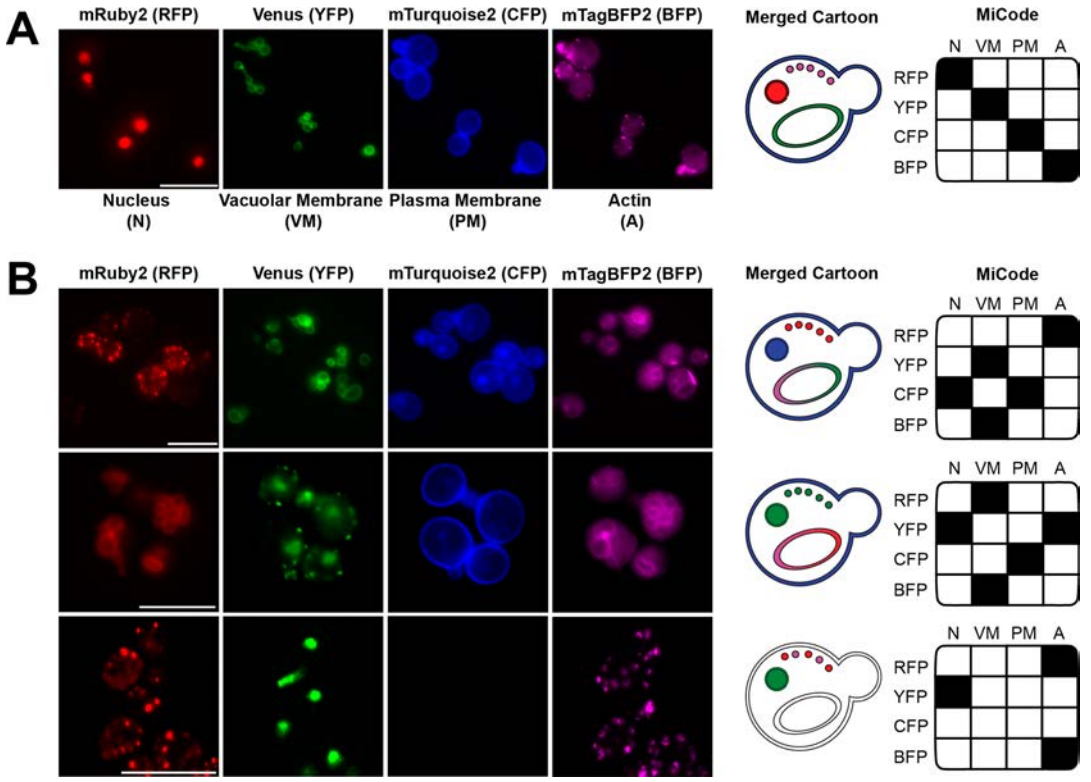


Figure 1. MiCode identification. (A) Cartoon yeast cells and fluorescent micrographs depicting the various subcellular organelle tags and fluorescent proteins used in constructing a MiCode. Four organelle tags are used: nucleus (HTA2), vacuolar membrane (ZRC1), plasma membrane (CIIC tag), and actin (ABP1). Four fluorescent proteins are used: mRuby2 (RFP), Venus (YFP), mTurquoise2 (CFP), and mTagBFP2 (BFP). (B) Several examples of randomly selected MiCodes for demonstration purposes. All scale bars are 10 μm, and all micrographs were contrast-adjusted for clarity.

cassettes to create a unique color for each labeled cell.[2,3] While such approaches provide a striking visualization for distinguishing clonal cells, their stochastic nature makes it difficult to programmatically link the cell's genotype to the observed phenotype. Another technology, fluorescent cell barcoding, uniquely labels each cell type by using a distinct dye with further multiplexing accomplished by using different concentrations of each dye to create more unique fluorescent signatures.[4] While this approach allows for the pooled screening of mixed populations, it requires that each cell type be fixed then individually labeled, thereby limiting the method's utility for live cell screening and pooled experiments of large combinatorial libraries. Recent work with surface zinc fingers has progressed toward genetically encoded cell surface scaffolds that hybridize with fluorophore-labeled DNA probes.[5] While this technique has been used to identify up to six unique genotypes in a pooled screening assay with live, immobilized cells, expanding to larger library sizes ($\geq 10^2$) would require increasing numbers of hybridization/quenching steps that would limit experimental throughput.

To address these limitations in existing methods for high-throughput cell labeling, we have developed a microscopy-readable library of barcodes (MiCodes) that enables live-cell tracking and time-lapse imaging and can be used to increase the throughput of microscopy experiments. Each MiCode consists of a combination of spectrally distinct fluorescent proteins targeted to visually discernible organelles. In this manner, a large set of unique MiCodes can be obtained with the total number of MiCodes defined by the following equation:

$$m = f(s, c, o) = (s^{\wedge}c)^{\wedge}o \quad (1)$$

where m is the number of MiCodes, s is the number of channel states, in this case two for OFF and ON, c is the number of fluorescent channels, and o is the number of organelles. A MiCode library can be physically coupled to a DNA-encoded library such that each library member is genetically linked to a unique MiCode, forming the aforementioned necessary link between phenotype and genotype. Here, we demonstrate a proof-of-concept MiCoding screen with four fluorescent proteins and four organelles using a subset of the possible $(2^4)^4 = 65\,536$ MiCodes. Micrographs can be taken to measure, in parallel with the desired phenotype, the MiCodes of all cells in the field of view.

As an example application, we tested candidate interactions among a set of synthetic coiled-coil proteins that were designed to maximize orthogonality. Coiled coils have been deployed in many synthetic biology applications as modular, high affinity protein−protein interaction parts for engineering protein assemblies in a cell.[6−8] Increased numbers of specific, orthogonal interaction pairs would enable synthetic biologists to engineer multiple independent complexes in the same cell with minimal cross-talk. Achieving interaction orthogonality appears to be challenging.[9,10] For example, although orthogonality was not an explicit goal of a previous effort by Reinke and co-workers when they constructed a set of 23 synthetic coiled-coils with high affinity for their cognate partners, it is still instructive that a maximum set of only three orthogonal pairs was observed.[11] When characterizing candidate sets of orthogonal pairs, the size of the interaction matrix that must be tested increases as n^2, where n is the number of protein pairs to be tested. To address this, we MiCoded a matrix of 64 candidate coiled-coil interactions and developed a two-hybrid

DOI: 10.1021/acssynbio.5b00060
ACS Synth. Biol. XXXX, XXX, XXX−XXX

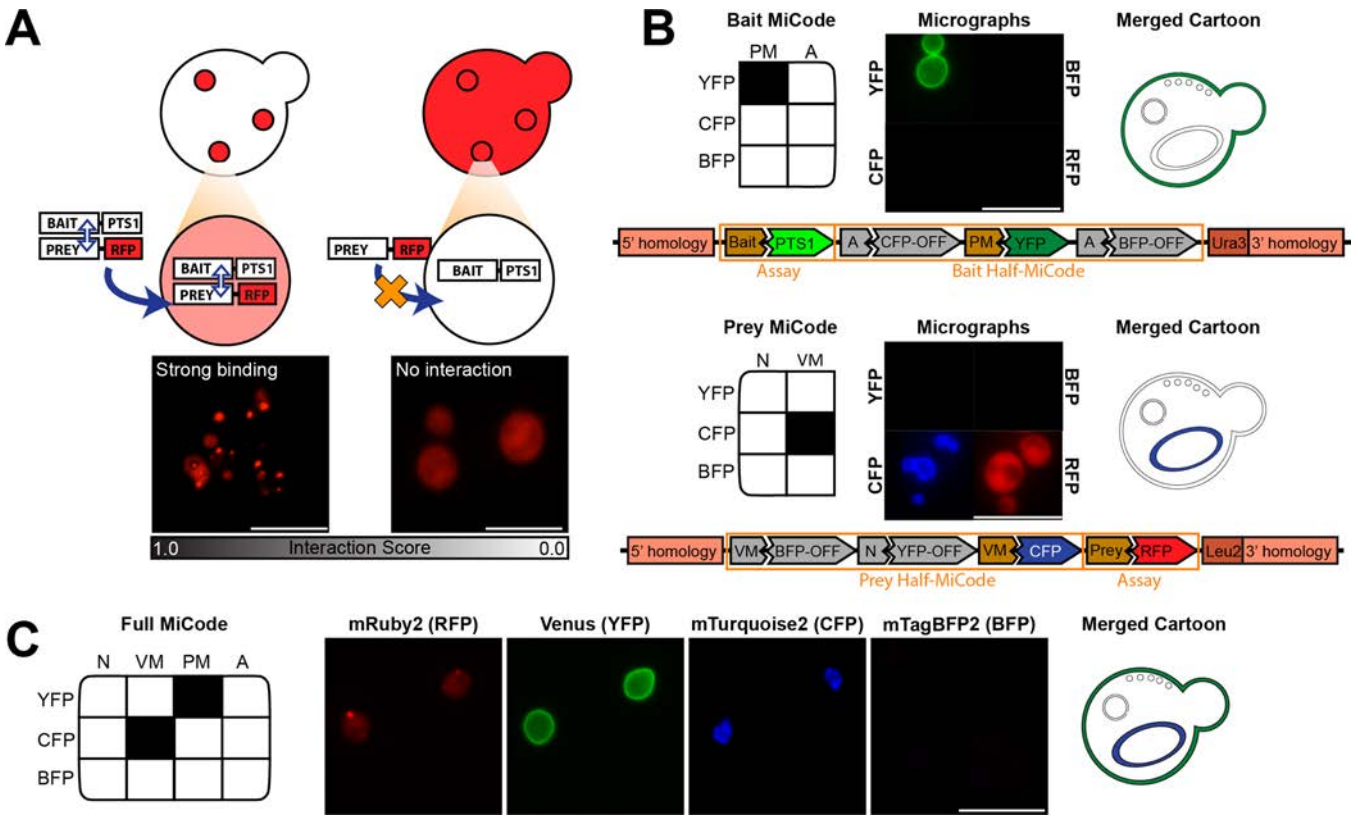


Figure 2. Two-hybrid fluorescence localization assay for probing protein−protein interactions. (A) Cartoon yeast cells and fluorescent micrographs depicting strong and weak binding scenarios. The bait protein is also fused to a photoactivatable GFP protein (not shown for clarity) that is used for verifying peroxisome localization of the bait protein (SI Figure S4). The strong binding example strain is bait/prey SYNZIP2/SYNZIP1 with a reported K_d of <10 nM, while the weak binding example is of SYNZIP18/SYNZIP21 with a reported K_d of >300 nM.[21] (B) Cartoon yeast cells and fluorescent micrographs of two strains each containing a half-MiCode, either only the bait or prey half. The cells were digitally magnified above the microscope's 1000× magnification to better illustrate the organelle targeting. Because the prey is tagged with RFP, when only it is expressed, there is diffuse red fluorescence. Below the micrographs, we present gene locus maps showing the half-MiCode and bait/prey expression cassettes *in situ* in the yeast genome. Gray regions in the maps indicate spacer sequences. The order of the fluorescent proteins and localization tags reflects their position as a result of the Golden Gate cloning strategy (SI Figure S2). (C) Cartoon yeast cell and fluorescent micrographs of a strain containing a full MiCode with both the bait and prey. In this case, there is a strong observable interaction between bait/prey 100A/100B. All scale bars are 10 μm, and all micrographs were contrast-adjusted for clarity.

assay using subcellular localization as an output to simultaneously determine if an interaction occurred in the cytosol *in vivo* for all members of this matrix. Thus, the MiCodes developed herein can enable microscopy to be used as an output for a genetically coupled DNA-encoded library.

RESULTS AND DISCUSSION

I. Construction of MiCodes. Identifiable microscopy-readable barcodes require clearly discernible morphological features. As such, we targeted organelles that produced the following phenotypes: a large ring (plasma membrane), a smaller ring (vacuole membrane), a large single punctate spot (nucleus), and peripherally clustered small punctae (F-actin patches) (Figure 1A). To achieve these phenotypes, we tested a series of organelle targeting tags (Supporting Information (SI), Table S1) described in the Yeast GFP Fusion Localization Database, (yeastgfp.yeastgenome.org/) as fusions to each of the fluorescent proteins used in this study (SI Table S2) under a range of expression levels. The targeting tags showing the clearest organelle targeting with minimal cytosolic background fluorescence were the C-terminal peptide tag CIIC for the plasma membrane, zinc transporter ZRC1 for the vacuole membrane, histone HTA2 for the nucleus, and actin binding protein ABP1 for F-actin patches (SI Table S1). To scale the diversity of possible barcodes, we next increased the number of fluorescent channels under which each targeting tag could be visualized. The compatible four channels mRuby2 red fluorescent protein (RFP), Venus yellow fluorescent protein (YFP), mTurquoise cyan fluorescent protein (CFP), and mTagBFP2 blue fluorescent protein (BFP) were used (SI Table S2). The expression level yielding a distinct fluorescence signal at the desired location with minimal background cytosolic fluorescence was determined. For initial expression level optimization, we used five promoters spanning approximately three orders of magnitude in strength[8] to drive the expression of each of the organelle targeting tags fused to GFP. In most cases, we found moderate expression levels (pRPL18B) to be best for yielding a clear positive signal for the given target organelle while minimizing diffuse, cytosolic fluorescence (SI Figure S1, Table S3). In general, the precise optimal expression level of each tag will also likely be dependent on the imaging protocol used since minimization of cytosolic background is critical for unambiguous MiCoding.

Synthesis of the MiCode was done using a MoClo Golden Gate assembly strategy[12,13] and the constituent parts of the MiCode were integrated into the *URA3* and *LEU2* loci of the

DOI: 10.1021/acssynbio.5b00060
ACS Synth. Biol. XXXX, XXX, XXX−XXX

chromosome (SI Figure S2). In preliminary design iterations, we observed undesired homologous recombination among the MiCode constructs due to repeated use of the same promoters, fluorescent proteins, and terminators. Accordingly, alternate promoters with similar transcriptional strengths to pRP18B, but with sequence-independent nucleotide composition, were used to drive the various targeting tag/fluorescent protein fusions (SI Table S3). Similarly, different terminators were used for each fusion protein (SI Table S4). Finally, the organelle targeting tags (SI Table S1) and the fluorescent protein reporter genes were recoded to make an alternate primary sequence for Venus, mTurquoise2, and mTagBFP2 to further reduce the risk of recombination (SI Table S2).

A first test of the compatibility of the four organelle tags was done by targeting a unique fluorescent protein to each organelle in the same cell—specifically, RFP to the nucleus, YFP to the vacuolar membrane, CFP to the plasma membrane, and BFP to actin (Figure 1A). The resultant cell showed the expected MiCode phenotype when visualized under each nonoverlapping fluorescent channel. A few arbitrary MiCodes were then synthesized as a further test of feasibility (Figure 1B). In this small set, we included strains that had organelles simultaneously targeted with two distinct fluorophores as well as organelles without a targeted fluorescent protein. In each of these MiCode examples, the correct genotype was accurately discernible. The most ambiguous phenotype we observed was for vacuolar membrane targeting, which was sometimes similar in appearance to the nucleus. With experience, however, the vacuolar membrane and nucleus could be clearly distinguishable by the intensity of fluorescence in the central region of the organelle or by the quantity of structures (SI Figure S3).

II. Development of a Two-Hybrid Fluorescence Localization Assay for Measuring Protein−Protein Interactions in the Yeast Cytosol. As a first proof-of-concept of applying MiCodes for improved throughput of analyzing a cellular process, we decided to test protein−protein interactions in the cell. Although MiCoding genetic libraries could be useful for a variety of outputs best measured by microscopy, we used subcellular localization as a readout of protein−protein interactions *in vivo* using a novel two-hybrid assay. Many analytical methods can measure these interactions *in vitro* such as fluorescence anisotropy, pulldown assays, co-immunoprecipitation, and surface plasmon resonance.[14] Although these methods have proven immensely powerful, they provide indirect measurements and do not capture the cytosolic context of the cell (i.e., measure the protein−protein interaction of interest in the presence of potential off-target interactions in the cytosol).

In our two-hybrid fluorescence localization assay, we used the ability of peroxisomes to import folded proteins, including protein assemblies, to directly probe whether two proteins interact in the cytosol. Import efficiency likely will depend on the size of the complex formed, though the peroxisome import complex has been shown to be capable of accommodating cargo with diameters as wide as 9 nm.[15] A bait protein was tagged with a UV-photoactivatable (PA) GFP and a three amino acid (SKL) C-terminal peroxisome targeting tag (PTS1), while a candidate prey protein was tagged with a fluorescent reporter protein (RFP) (Figure 2A). The UV-photoactivatable GFP provided the ability to verify that the bait protein was both expressed and properly targeted to the peroxisome without interfering with MiCode analysis (SI Figure S4). An interaction between the bait and prey led to red punctae resulting from coimport of the RFP-tagged prey protein along with the PTS1-tagged bait protein into the peroxisome (left panel, Figure 2A). The absence of an interaction resulted in diffuse red fluorescence in the cytosol (right panel, Figure 2A). We used a binary scoring metric for each cell analyzed where observation of a distinct punctate phenotype was scored as a 1 and a completely diffuse phenotype was scored as a 0. Since MiCoding individual strains allows for the assaying of many cells in parallel, the final scores represent averages of multiple cells for increased confidence. Because this assay utilizes the red channel to measure interactions *via* peroxisome targeting, we used the remaining YFP, CFP, and BFP channels for the MiCode.

III. SYNZIPs and Design of New Specific Coiled-Coils. We next sought to design coiled-coil zippers, which are short protein domains of approximately 30 amino acids that can form either coiled-coil homo- or hetero-oligomers with α-helical structure and a range of affinities. Our eventual goal was to screen a library of newly designed zippers using MiCodes in conjunction with the two-hybrid fluorescence localization assay. New orthogonal synthetic coiled-coils would further expand the synthetic biology toolbox of protein−protein interaction parts to allow more protein complexes to be formed in the yeast cytosol with minimal crosstalk. Several groups have successfully engineered synthetic coiled-coil interaction partners.[16−20] In particular, Thompson *et al.* reported 23 synthetic coiled-coil leucine zippers that can form many distinct pairwise complexes.[21] These zippers are referred to as SYNZIPS, a designation given to coiled-coil zippers whose biophysical characteristics and interactions with other zippers have been extensively characterized.

Reinke *et al.* (2010)[11] and Thompson *et al.* (2012)[21] discussed the use of both *in vitro* (coiled−coiled protein microarray and fluorescence polarization) and *in vivo* (transcription-based yeast two-hybrid and MAPK signaling) assays to measure interactions among SYNZIP protein pairs.[11,21] The *in vitro* microarray assay detected interactions between protein pairs by printing one partner onto an aldehyde-derivatized slide and testing it for binding to fluorescent dye-labeled SYNZIP partners.[8] The fluorescence polarization assay mixed maltose-binding protein-SYNZIP fluoresceinated proteins with increasing concentrations of unlabeled maltose-binding protein-SYNZIPs. Dissociation constants (K_d) were determined by performing nonlinear least-squares regression on curves of the fraction of fluorescein-labeled protein bound for various titrations of unlabeled protein. The transcription-based yeast two-hybrid (Y2H) assay used fusions of proteins to either the Gal4 DNA-binding domain or the Gal4 activation domain. Successful reconstitution of the Gal4 transcription factor, which was mediated by an interaction between partners of interest (SYNZIP pairs), drove expression of auxotrophic marker genes such as *URA3* and *HIS3*. For this assay, colony size was used as a proxy for binding affinity. Finally, the MAPK signaling assay employed the Ste5 protein scaffold and Msg5 phosphatase, a negative modulator of the MAPK-Ste5 signaling cascade in yeast. Individual SYNZIPs were fused to either Ste5 or Msg5 and interactions were read out by expression of a GFP reporter gene controlled by Fus3 activity (Fus3 activity was inhibited by localization of Msg5 to the Ste5 scaffold). An interaction between a protein pair resulted in decreased expression of the GFP reporter. The fractional expression of GFP in comparison to a control with an Ste5 scaffold with no protein fusion was used as a quantitative metric in this assay.

DOI: 10.1021/acssynbio.5b00060
ACS Synth. Biol. XXXX, XXX, XXX−XXX

To design new orthogonal coiled coils, we used a scoring function previously developed for predicting bZIP coiled-coil interactions.[22] This scoring function predicts binding strength for a pair of sequences and is based on considering pairs and triplets of residues that would interact if the proteins formed a parallel coiled-coil dimer. The model was trained on a large experimental set of quantified bZIP coiled-coil interactions,[23] and it showed good agreement with experimental binding data in rigorous cross-validation testing. The scoring function can be evaluated quickly because it requires only sequences as input, and it was previously used to design tight and specific binders for four human bZIP proteins.[22] To design orthogonal pairs of coiled-coil zippers for this work, sequences were assembled from a library of short 7-residue fragments (heptads) taken from known bZIP proteins, as done previously by Potapov et al.[22] Integer linear programming (ILP) was used to solve for small sets of eight optimal designed sequences. These sets were combined to give larger sets of 20 peptides predicted to form orthogonal interactions, as described in the Methods.

We used the reported heterodimer dissociation constants of SYNZIPs obtained from a fluorescence polarization assay (measured *in vitro* with purified protein) to calibrate the peroxisome localization assay.[21] Representative interactions spanning a range of affinities, some of which were also quantified in the MAPK signaling assay, are shown in SI Figure S5. Two-hybrid fluorescence localization interaction scores—obtained by averaging the scores of at least five cells for each strain of a SYNZIP subset—could distinguish interactions with different reported affinities. We determined that the range of interaction affinities that could be measured by the two-hybrid fluorescence localization assay was ∼10−400 nM. On the basis of the comparison of our interaction scores to the results of the other three assays (SI Figures S5−S7), we defined an interaction score threshold of ≥0.4 (K_d < 200 nM) for a strong interaction, and interaction scores 0−0.4 (K_d > 200 nM) for weak interaction.

To further compare our two-hybrid fluorescence localization assay to previous assays, we tested 68 pairwise interactions between 14 SYNZIPs previously analyzed by Reinke et al. and by Thompson et al. (SI Figure S6).[11,21] We cloned the proteins involved in 27 interaction pairs in both bait−prey and prey−bait orientations, for a total of 54 heterodimer interactions, and additionally tested 14 homodimer interactions. Our two-hybrid fluorescence localization assay agrees moderately well with the interaction profiles reported using fluorescence polarization assays, transcription-based Y2H, and MAPK signaling assays. As shown in SI Figures S6 and S7, most SYNZIPs showed similar interaction strengths for the microarray and two-hybrid fluorescence localization assays; however, there were exceptions. Of the 68 candidate SYNZIP interactions tested using both of these methods, two were diagnosed as strongly interacting by the microarray assay but categorized as weakly interacting by the two-hybrid fluorescence localization assay, while six candidate interactions were diagnosed as extremely weak by the microarray assay but strong by the two-hybrid fluorescence localization assay. Overall this 8 of the 68 disagreement with the microarray results (12%) is similar to the range of variation seen between the coiled-coil microarray and transcription-based Y2H assays (10−13%) (SI Figure S7B).[11] As discussed by Thompson et al., discrepancies between the *in vitro* and *in vivo* interaction assays demonstrate the need to test interactions in the context in which they will be used.[21] To this end, the two-hybrid fluorescent localization assay serves as an additional platform for validation as well as discovery of protein−protein interactions in the cytosol; it should not be used as a sole measurement of a protein−protein interaction.

Various comparisons can be made between our two-hybrid fluorescence localization assay and the *in vivo* and *in vitro* assays described by Thompson et al.[21] Although the *in vitro* fluorescence polarization assay provides quantitative binding affinities, these affinities may not be a precise predictor of protein interaction behavior in a cytosolic context. Y2H resolves some of these potential problems and can be assayed in a quantitative manner, but its output is an indirect measure of the binding and can be confounded by other factors such as nonspecific or indirect interactions.[24,25] The MAPK signaling assay is similar in this regard. The two-hybrid fluorescence localization assay provides information on the SYNZIP binding performance *in vivo* with single-cell resolution. Instead of measuring the output of a linked transcriptional response, this assay directly couples a protein interaction to colocalization, giving an output that is a direct readout of steric interaction. An additional benefit of the two-hybrid fluorescence localization assay is the detection of off-target localization. In our initial screen of newly designed coiled-coils, we observed a few of our newly designed prey proteins fused to RFP mistargeting to the nucleus and peroxisome (R.C. unpublished data). In sum, our two-hybrid fluorescence localization assay expands the toolkit of protein interaction assays by providing a direct *in vivo* measurement of protein binding that is amenable to microscopy.

IV. Application of MiCoding toward the Screening of a Library of Coiled-Coil Zippers. Toward enabling higher-throughput microscopy screening of candidate protein interaction pairs, we added MiCode tags to a single pair of newly designed coiled-coil proteins, 100A and 100B, designed to heterodimerize (see Methods). For testing of the interaction between 100A and 100B using the two-hybrid fluorescence localization assay, we designated 100A as the bait and 100B as the prey. The newly designed bait zipper (100A) and prey zipper (100B) were cloned separately and each was incorporated into a DNA segment that included half of the eventual full MiCode (Figure 2B and SI Figure S2). In this single-pair mock assembly (i.e., the identities of the bait and prey were known *a priori*; however, the bait−prey binding affinity was uncharacterized by other assays), the gene expressing the bait construct was physically linked to the genes expressing the plasma membrane and actin-targeting half of the MiCode. Similarly, the gene expressing the prey construct was cloned next to the genes expressing the other half of the MiCode (nucleus and vacuole membrane targeting). The prey zipper and its half-MiCode were integrated into the *LEU2* locus, followed by integration of the bait zipper and other half-MiCode into the *URA3* locus. Figure 2B shows representative strains containing either the bait-half MiCode (top panel) or prey-half MiCode (bottom panel) constructs integrated in the chromosome. A full MiCode was created through cointegration of the bait and prey half-MiCoded constructs into the same strain (Figure 2C). Upon inspection of the fully-MiCoded test strain *via* microscopy, we observed punctate fluorescence in the RFP channel yielding an interaction score of 1 between bait 100A and prey 100B. The identities of both the bait and prey were verified by identifying the targeted organelles for each of the YFP, CFP, and BFP channels. In the YFP channel, the plasma membrane is visible.

DOI: 10.1021/acssynbio.5b00060
ACS Synth. Biol. XXXX, XXX, XXX−XXX

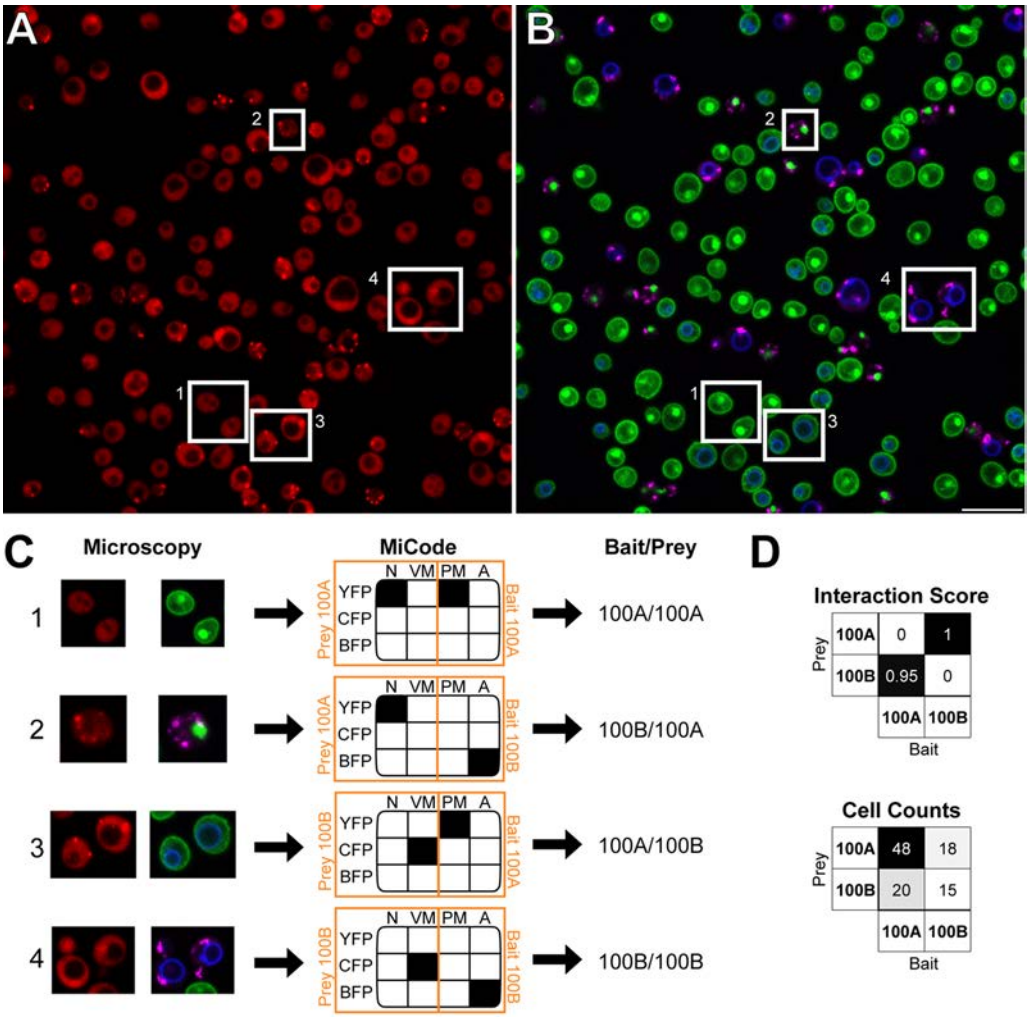


Figure 3. An example library that links the leucine zipper assay to preassigned MiCodes. (A) Leucine zipper assay with cells representative of the four MiCode phenotypes indicated. (B) Merged color channels used for MiCode identification. (C) Isolated examples of the four MiCode phenotypes possible in this small library. In the MiCode column, the half-MiCodes used for the prey and bait are boxed in orange. (D) Interaction scores and interaction counts are based solely on the images in panels A and B. The interaction scores across cells were computed by scoring all cells in a micograph on binary scale: 0 for a lack of a visible interaction and 1 for any visible punctae. The counts value corresponding to each bait/prey pairing represents the number of cells used to calculate the average binding score. The values were then averaged over the given number of cells to yield the final score. The micographs were contrast-adjusted for clarity.

In the CFP channel, the vacuolar membrane is visible, while no organelle is visible in the BFP channel.

Next, we extended the MiCoding strategy to the simultaneous testing of the entire two-by-two interaction matrix for the 100A/100B zipper pairing. The complete interaction matrix contained four experimental configurations: bait/prey 100A/100A (test for 100A homodimerization), 100B/100A (test for heterodimerization in orientation one), 100A/100B (test for heterodimerization in orientation two), and 100B/100B (test for 100B homodimerization). We first tested these four interaction pairings clonally without MiCoding to avoid any possible physiological effect of the MiCode on the interaction phenotype (R.C. unpublished data). Next, we tested all four pairings simultaneously by MiCoding each configuration and screening the resulting mixed population (Figure 3). Since our assembly strategy preassigned bait and prey proteins each with a specific half-MiCode (Figure 3C and SI Figure S2), the full MiCode phenotype allows us to extract the genetic identity of the interacting coiled-coil pair. Cells representative of the four experimental configurations from the two-by-two matrix are boxed in Figure 3A and 3B and highlighted in Figure 3C along with the MiCodes and the decoded bait/prey identities. As shown in Figure 3D, the interaction scores could be rapidly determined by surveying a large number of cells. We calculated the interaction score for each bait/prey configuration by averaging the binary output of the assay—interaction (1) or no interaction (0), over all the cells corresponding to that genotype in the field of view. In addition to reporting this numerical score, we also present the number (counts) of cells of each genotype present on the micrograph used for this calculation. From this small 2 × 2 matrix screen, we report that protein pair 100A/100B exhibits only heterodimerization, with no detectable homodimerization. Furthermore, we note that in this test the presence of the MiCode did not affect the phenotypic output of the interaction assay.

Encouraged by these results, we screened a larger set of novel designed coiled-coils with MiCoding. Ten new protein pairs (101A/B−110A/B) were designed to form a set of orthogonal heterodimers (see Methods, SI Figure S8, and SI Table S5).

DOI: 10.1021/acssynbio.5b00060
ACS Synth. Biol. XXXX, XXX, XXX−XXX

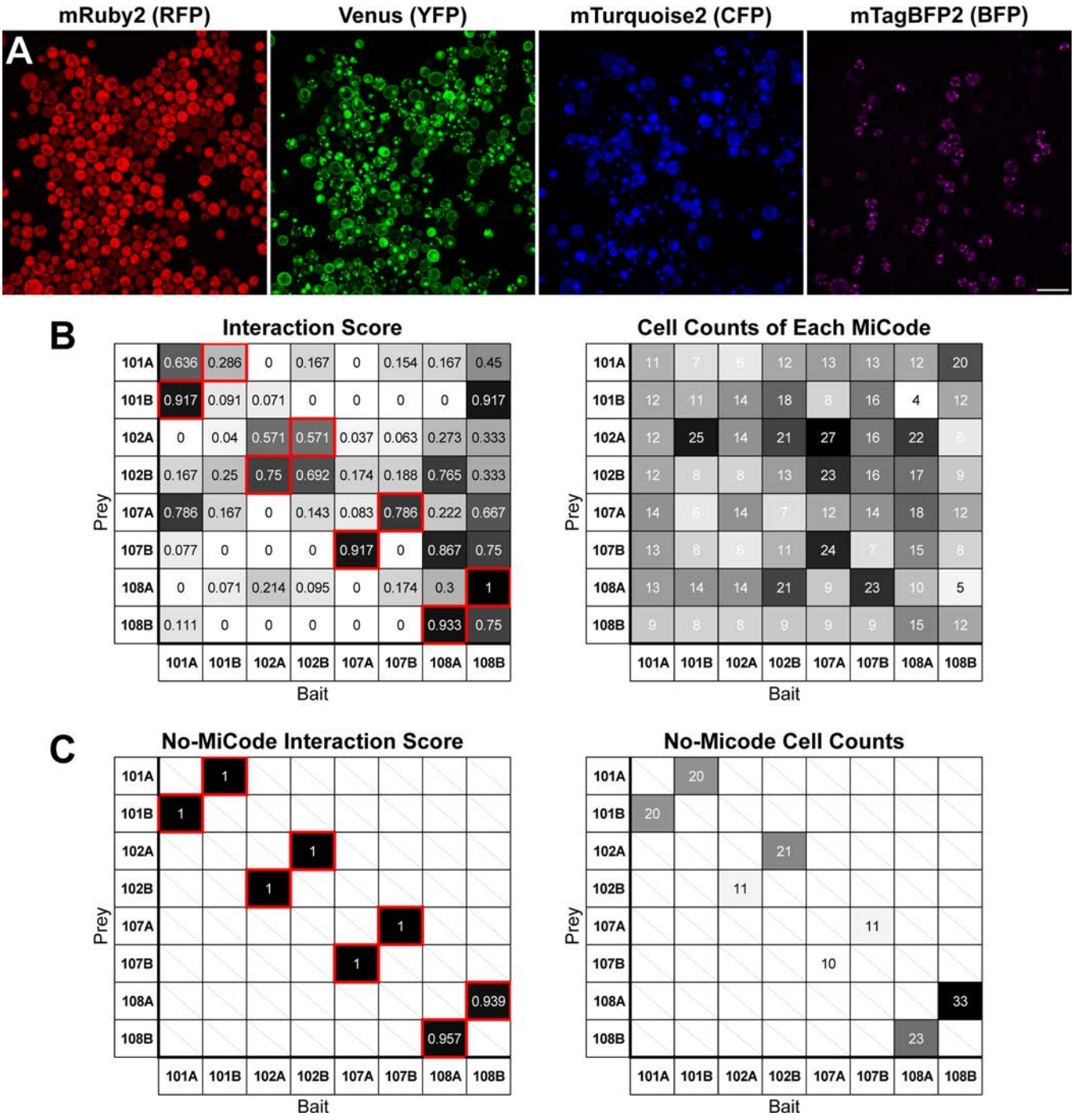


Figure 4. MiCodes applied to assaying an 8 × 8 array of coiled-coils. (A) A representative image of the 8 × 8 library that contains a mixed population of all bait−prey interactions. (B) Tabulated data for the two-hybrid fluorescence localization assay experiment showing interaction scores and interaction counts. These data are averaged across 36 micrographs. These results were supplemented with micrographs of single monoclonal strains for bait/prey 107A/107B, 107B/102B, 108A/101A, 108A/101B, 108A/102A, 108A/102B, 108A/107A, 108A/107B, 108A/108A, and 108A/108B, the results of which were added to the pool. Unsupplemented data are shown in SI Figure S10. (C) Tabulated data for the no-MiCode control strains that was used for validating the cognate hetero-oligomer interactions. Each of the sight cognate coiled-coil interactions was assayed using the two-hybrid fluorescence localization assay. The micographs were contrast-adjusted for clarity.

Each designed pair was pretested in a two-hybrid localization assay screen for heterodimerization. Eight of ten pairs showed clear evidence for interaction—a score greater than 0.4—with the intended partner (SI Figure S9). Three of 20 constructs severely inhibited cell growth, and three of the eight designed pairs gave asymmetrical interaction results with respect to the bait/prey configuration (SI Figure S9). We chose four of the remaining five well-behaved coiled-coil pairs for screening in a full 8 × 8 matrix of 64 potential interactions (Figure 4B).

To determine whether each designed protein heterodimerized exclusively with its cognate pair in the set (*i.e.*, exhibited no interactions with other proteins in the matrix), we MiCoded and tested all possible interactions between noncognate pairs, including potential homodimers. To assemble all possible

DOI: 10.1021/acssynbio.5b00060
ACS Synth. Biol. XXXX, XXX, XXX−XXX

interacting pairs, we first generated the eight half-MiCoded prey strains individually. Next, we pooled the prey strains in a one-pot reaction and transformed a cocktail consisting of the eight bait half-MiCode integration cassettes to generate the full library of MiCoded interacting partners. A representative series of images for this assay, run with all variants pooled, is shown in Figure 4A. Thirty-six images were taken in each of the four channels to acquire multiple cell counts of each of the 64 matrix members (Figure 4B). As expected, all of the cognate pairs showed interactions in a high fraction of the cells analyzed, with the lone exception of bait/prey 101B/101A. Many proteins also showed evidence of undesired homodimerization, notably proteins 101A, 102A, 102B, and 108B. Although there was a much lower incidence of noncognate and off-target interactions than cognate interactions, several were observed such as 101A/101A (noncognate) and most prey coiled-coils with bait 108B (off-target). In some cases like 108B, we observed asymmetric results: bait 108B showed some binding to all prey proteins, whereas prey 108B showed little or no binding to most bait proteins.

We again probed the possibility that the addition of a MiCode could perturb the assay or impede cell growth. First, we examined the interactions of designed cognate coiled-coil pairs with and without the MiCode tag (Figure 4C). Most (seven of eight) of the cognate heterodimer pairings exhibited a similar interaction phenotype both with and without the MiCode as expected. Notably, the interaction of bait 101B and prey 101A was much weaker in the MiCoded assay than in the un-MiCoded assay. Thus, with the exception of 101B/101A, expression of the MiCode did not seem to impact the ability to assay a positive interaction. Another consideration is the potential growth inhibition resulting from overtargeting certain organelles. For example, most strains with bait 108A were present in lower abundance (*i.e.*, had lower cell counts) in the library than expected. All strains with bait 108A had actin targeted by both Venus and mTurquoise2, suggesting that double-targeting actin with those fluorescent proteins in combination is mildly toxic to the respective host strains causing low cell counts (SI Figure S10). To acquire appreciable cell counts, we grew these problematic MiCoded strains individually using exactly the same MiCodes to measure candidate interactions reported in Figure 4B. We previously observed a similar toxicity when double-targeting the vacuole membrane using any two of YFP, CFP, or BFP, and therefore we removed all double-targeted vacuole membrane containing MiCodes (surprisingly, any combination of YFP, CFP, or BFP with RFP targeted to the vacuole membrane is not toxic, such as the strain shown in Figure 1B). This toxicity is particularly an issue when growing the library collectively in a single batch where these MiCode-containing clones can be outcompeted by other strains. Such growth biases have been known to be present in library-based approaches. For example, mutants in pooled, high-throughput transposon mutagenesis libraries can grow at different rates based on how the transposon insertion perturbs the genetic context of each variant,[26] thereby resulting in library members to be present at different abundances.

We present the results for two small MiCoded libraries, a 2 × 2 and an 8 × 8 matrix, that we used to screen newly designed coiled-coils that were computationally optimized for orthogonality. In the 2 × 2 matrix, we tested the four possible interactions—two cognate and two noncognate—with and without the MiCode and observed no significant differences in the results between the two methods. Importantly, MiCoding improved the ease of screening and allowed us to extract all the results from a single yeast transformation reaction and a small number of micrographs. Encouraged by the agreement between the MiCoded and the non-MiCoded data, we moved toward a larger set of 20 proteins (10 pairs), which we pared down to an 8 × 8 matrix screen based on knowledge gained from an initial non-MiCoded interaction screen of designed cognate pairs. We then performed a MiCoded screen of the 8 × 8 matrix to acquire data for all 64 possible interactions in order to screen for off-target and homodimer interactions. In this case, we report good, but not perfect, agreement between the MiCoded and non-MiCoded data.

Our interaction data demonstrate that achieving orthogonality in this cellular localization assay was difficult, even when proteins were designed with this explicit goal in mind. Regardless, we report two new coiled-coil pairs, 100A/100B from the 2 × 2 matrix screen and 107A/107B from the 8 × 8 matrix screen, that can be used for synthetic system designs where hetero-oligomerization without homo-oligomerization is desired. With respect to the unintended interactions, we emphasize that we have not definitively determined the sensitivity of the localization assay compared to that of other techniques. Although the overall agreement of the new assay with previously reported SYNZIP interaction data was good, it is also true that SYNZIP pairs not previously reported to interact by any of the four other assays were positive for our screen. Furthermore, it is important to note that the interaction geometry and stoichiometry of complexes formed by the designed proteins have not been determined. The computational design procedure considered only parallel coiled-coil heterodimer structures, yet small changes in sequence can lead to changes in coiled-coil assembly.[27] Thus, there is a possibility that the unintended complexes are higher-order homo- or hetero-oligomers, or dimers with helices aligned in an antiparallel orientation. Indeed, retrospective computational analysis indicated that pairs 101A/101A, 102A/102A, 102B/102B, 101B/108B, 102B/108B, 107A/108B might form strong interactions when the constituent alpha helices associate in an antiparallel mode. In such a mode, charged residues at the e and g helical positions in these pairs could form attractive interhelical Coulombic interactions. This finding highlights the importance of considering multiple competing states when designing coiled coils. Recently, an improved model for scoring antiparallel coiled-coil interactions was developed that could be used to augment the design protocol used in this work.[17]

While finding orthogonal coiled-coiled zipper pairs from a rationally designed matrix of interaction candidates proved difficult, the use of MiCodes greatly increased the speed of the screening process by allowing for data collection and analysis of pooled, heterogeneous samples. MiCoding allowed many cells to be analyzed simultaneously in a single micrograph, which enabled a larger matrix of proteins to be assayed while still achieving replicates for analysis. In conclusion, the two-hybrid fluorescence localization screen, when coupled with MiCodes, enabled quick screening of a large matrix of interactions through mixed-sample analysis and should facilitate faster iterations of the design—build—test—learn cycle in future studies.

V. Scaling the Size of a MiCoded Library. In this study, we used only a small subset of the potential MiCode library as required by the size of our experiment. However, it should be feasible to conduct experiments with considerably larger libraries where each individual member is uniquely MiCoded.

DOI: 10.1021/acssynbio.5b00060
ACS Synth. Biol. XXXX, XXX, XXX−XXX

As the size of the MiCode library is scaled, however, a few considerations should be made. First, care should be taken to choose organelle-targeting tags that both produce a phenotypically distinct MiCode and confer no physiological growth defect to the host. Expression of additional organelle-targeting tags will increase the risk of toxicity. Although we optimized expression level of the organelle targeting tags for a widefield fluorescence microscope, it is possible that the use of a more sensitive microscope or method such as confocal microscopy could achieve a higher signal-to-noise ratio and potentially allow the expression levels of these targeting tags to be lowered below the toxicity threshold, even with multiple targeting. Another consideration is the difficulty in correctly identifying organelles that can look similar. Clearly, this is the foremost consideration for incorporating additional organelles into a MiCode set. Nuclear and vacuolar membrane targeting is an example of an easily misdiagnosed targeting, although experience in identification considerably improves the accuracy of MiCode identification (SI Figure S3).

The size of MiCode libraries could be exponentially increased through the addition of channel states, organelles, or fluorescent channels, provided they are clearly discernible from the other organelles and fluorescent channels. For example, a strategy for increasing the number of channel states is to use a UV-photoactivatable (PA) fluorescent protein. All other fluorescent proteins could be photobleached followed by UV stimulation of the PA fluorescent proteins. Preliminary tests of PA fluorescent proteins were promising (R.C. unpublished data). If these PA proteins could be implemented without interfering with the ON/OFF states of other fluorescent proteins, this could give each fluorescent channel three states (ON/OFF/PA) instead of the two described above. Further expansion of the number of channel states could be achieved by precise modulation of fluorescence intensity, giving both STRONG and WEAK ON states.

Additional fluorescent channels could be incorporated through peptide-fused quantum dots as fluorescent protein analogues. Quantum dots are fluorescent nanoparticles with precise emission wavelengths that can be tuned to allow additional channels to be multiplexed.[28] When conjugated with organelle-targeting peptides, quantum dots have been shown to be taken up by live cells and targeted to organelles.[29] A scaffold molecule, protein, DNA, or RNA, could be expressed and targeted to organelles to recruit different patterns of quantum dots yielding unique MiCodes with linkages to known genotypes. The MiCoding platform can also potentially be coupled with other *in vivo* surface labeling probes to further increase barcode diversity.[5,30]

Increased numbers of organelles may be achievable, particularly if the confidence in accurate identification is improved by combining fluorescent microscopy with organelle staining methods using compatible fluorescent dyes or analytical approaches such as confocal microscopy that can improve signal over background. For example, there are additional organelles such as the mitochondria that we have not fully explored which could be compatibly targeted. It may even be feasible to subdivide an organelle, such as the nucleus and nucleolus.[31]

Returning to equation 1:

$$m = f(s, c, o) = (s\text{^}c)\text{^}o \quad (1)$$

where m is the number of MiCodes, s is the number of channel states, *c* is the number of fluorescent channels, and o is the number of organelles. Here, we leveraged the flexibility of MiCode library design and assembly with our prior knowledge of morphologically distinct fluorescent protein-targeting tag combinations to create a subset of easily identifiable MiCodes that satisfied our screening requirements. With this strategy, the following were created:

$$\begin{aligned} m &= \#(\text{bait MiCodes})*\#(\text{prey MiCodes}) \\ &= (S_{GFP,b}\text{^}O_{GFP,b}*S_{CFP,b}\text{^}O_{CFP,b}*S_{BFP,b}\text{^}O_{BFP,b}) \\ &\quad *(S_{GFP,p}\text{^}O_{GFP,p}*S_{CFP,p}\text{^}O_{CFP,p}*S_{BFP,p}\text{^}O_{BFP,p}) \\ &= (2\text{^}2*2\text{^}2*2\text{^}1)*(2\text{^}2*2\text{^}2*2\text{^}1) \\ &= 1024 \text{ possible MiCodes} \end{aligned}$$

where b = bait and p = prey

If all fluorescent channels are used for all organelles, then m = (s^c)^o = (2^4)^4 = 65536 possible MiCodes. If the system can be expanded using the techniques described above to increase the number of fluorescent channels by one, the MiCode library size can be expanded to (2^5)^4 = 1 048 576. Adding a third state to these fluorescent channels can increase this further: (3^5)^4 = 3 486 784 401.

The cloning of such large libraries will likely require one-pot assembly strategies rather than the individual construction of each library member. Using a scheme analogous to the Combinatorial Genetics En Masse (CombiGEM) approach for assembling higher-order combinations of DNA barcoded constructs, it should be possible to construct large-scale combinatorial libraries of MiCoded constructs (SI Figures S11 and S12).[32] These DNA-barcoded MiCode constructs could then be combined with a set of DNA-barcoded assay constructs (*i.e.*, bait−prey constructs for the leucine zipper assay) and the two barcodes can be associated using Illumina next-generation sequencing (SI Figure S13). In this manner, the MiCode and experimental phenotype can be visualized in parallel on the microscope and the unique identity of the library member can be extracted using the DNA barcodes associated with the MiCode and experimental genotype.

For analysis of larger libraries, it should also be feasible to design algorithms for automating MiCode image acquisition and identification. Publicly available cell segmentation software as well as custom scripts can be used to perform automated cell segmentation of MiCoded cells through edge detection and morphology-based filtering with up to 90% accuracy.[33] It should be feasible to automate MiCode identification (i.e., calling different organelles) by using morphology-based filtering thresholds derived from geometric analysis of a known MiCode data set. For example, the actin and nucleus organelles could be distinguished by their respective radii. Even so, more sophisticated algorithms applying supervised machine learning on preidentified MiCode micrographs will likely be necessary for robust image analysis identification and represent an area of future focus.

VI. Summary. Overall, MiCoding should prove useful for a variety of outputs measurable by microscopy but not amenable to high-throughput FACS. In this work, we first created tags for morphologically distinct organelles, then expanded the MiCode library size by adding several spectrally unique fluorescent channels. The diversity of a MiCoded library can be further scaled by increasing the number of channel states through

DOI: 10.1021/acssynbio.5b00060
ACS Synth. Biol. XXXX, XXX, XXX−XXX

photoactivatable proteins, subcellular tags through further subdivision, and fluorescent channels using quantum dots.

In addition, the MiCoding platform should prove generalizable and be applicable to other eukaryotic hosts. Possible future studies that could benefit from MiCoding are the exploration of genetic circuitry controlling neutrophil chemotaxis, the developmental tracking of cell lineages in small multicellular organisms such as *Caenorhabditis elegans*, and mechanobiological studies. For example, MiCoding could allow for the rapid surveying of genome-wide gene knockout or CRISPRi-mediated knockdown libraries at the single-cell level by enabling the identification of potentially undiscovered morphologies in pooled screens across a variety of conditions. This work represents a first step toward these exciting applications, allowing outputs clearly observable by microscopy to be analyzed in higher-throughput.

METHODS

MiCode Design and Construction. All cloning was performed using the Yeast Toolkit Golden Gate assembly strategy.[12,13] Single-gene plasmids containing a single organelle and fluorescence tag fusion or a single SYNZIP were transformed in TG1 chemically competent *E. coli* and were selected on LB containing ampicillin (100 mg/L). In subsequent reactions, the single-gene plasmids were assembled together into half-MiCode plasmids. As described earlier, we rationally designed the half-MiCode plasmids such that they linked the SYNZIP to preassigned MiCodes. These half-MiCode plasmids were transformed into TransforMax EPI300 electrocompetent *E. coli* and were selected on LB containing kanamycin (25 mg/L).

In all the constructs, any homologous regions that could potentially lead to unwanted recombination were eliminated. Each cassette within the system was expressed using different promoters of equivalent expression level and different terminators of equivalent termination efficiency. Additionally, an alternate version with different codon usage was generated for each of the MiCode tags and each of the fluorescent proteins. The codons were randomly chosen based on *S. cerevisiae* codon usage frequencies, and any rare codons were removed.

SYNZIP Design. We aimed to design several pairs of coiled-coil proteins (N pairs, 2N proteins) such that the first protein in a pair forms a strong interaction with the second protein from the same pair but forms only weak interactions with any other protein in the designed set. Interactions in sets of pairs that satisfy these criteria are called orthogonal interactions. To design orthogonal coiled coils, we used a scoring function previously developed for predicting bZIP coiled-coil interactions.[22] As in prior work, sequences composed of 7-residue fragments from known bZIP proteins were optimized with integer linear programming, with constraints imposed to favor desired patterns of binding and nonbinding. Specifically, we minimized the sum of the predicted binding scores for cognate interactions:

$$S = \sum_{i=1}^{N} S(A_i, B_i) \tag{2}$$

where N is the number of designed pairs and $S(A_i B_i)$ is the predicted binding score between protein A and protein B in the ith designed pair. Simultaneously, constraints were imposed to ensure that noncognate interactions were weak. In particular, to prevent formation of homodimers, we imposed constraints $S(A_i A_i) > S_{\text{cutoff}}$ and $S(B_i B_i) > S_{\text{cutoff}}$ for each designed pair i = 1...N, where S_{cutoff} was a predefined value chosen to ensure that homodimer interactions were at least 100-fold weaker compared to cognate interactions. Similar constraints were imposed for other noncognate interactions between sequences from different pairs (off-target interactions).

Simultaneously optimizing a large number of designed sequences and their interactions is a computationally demanding task even when using a simple scoring function. We were able to optimize four pairs of sequences at a time using a subset of the heptad library. To design a larger set of pairs, we repeated the design procedure many times, every time using a random subset of heptads for building designs. Following this, a set of 10 pairs was obtained by taking many four-pair sets and searching for larger sets of pairs with low potential for off-target interactions (SI Table S5, SI Figure S5). An initial set of 10 predicted orthogonal pairs was reduced to four pairs selected for further testing, based on an initial screen of the predicted heterodimers using the novel two-hybrid fluorescence localization assay (SI Figure S6). In this initial screen, toxic proteins that reduced cell growth, and asymmetric-strength protein pair interactions for which prey−bait and bait−prey cognate interactions differed in strength, were eliminated.

Strains and Media. All experiments were conducted in a BY4741 *S. cerevisiae* strain background (MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0). Multigene MiCode and zipper assay plasmids containing Leu2 and Ura3 auxotrophic markers were integrated into the respective chromosomal loci. Multigene plasmids were linearized with NotI then transformed using a standard lithium acetate transformation protocol.[34] Following heat shock, the cells were selected and grown on synthetic dropout media (6.7 g/L Difco Yeast Nitrogen Base without amino acids; 2 g/L US Biological Drop-out Mix Synthetic. Minus Leucine, Uracil, or Leucine and Uracil without Yeast Nitrogen Base; 20 g/L Dextrose). All *S. cerevisiae* strains were grown at 30 °C.

To form a complete MiCode, a bait and prey multigene plasmid were cointegrated to form a full MiCode strain. This process was used to create the strains shown in Figures 1 and 2. The full library (Figure 3) was created through a pooled cointegration with all the bait multigene plasmids. Each member the pooled set of bait plasmids was standardized to 1.5 pmol of DNA. Prior to the pooled library reaction, each individual prey MiCode strain was verified through fluorescence microscopy. In summary, to form the full library two successive rounds of integrations were performed, first selecting on leucine drop-out media then on uracil and leucine drop-out media.

I-SceI Integration. To increase the efficiency of these chromosomal integrations, as well as lower the percentage of multiple integration events, we first inserted an I-SceI homing endonuclease recognition site into the *URA3* locus that would later be targeted for integration. The I-SceI landing pad was integrated into the *URA3* locus by inducing double-stranded break with CRISPR/Cas9 and repairing the site with a PCR cassette containing the I-SceI recognition site. Transformation and expression of SceI is expected to create a double-stranded break in the chromosome and increase the efficiency of homologous recombination at this target site.[35] During the second transformation, the pooled multigene MiCode plasmids were cotransformed with a markerless plasmid to transiently express I-SceI. Testing determined that 40 fmoles of this

DOI: 10.1021/acssynbio.5b00060
ACS Synth. Biol. XXXX, XXX, XXX−XXX

plasmid per integration reaction was the optimum amount and that it increased integration efficiency by 9.7-fold.

Microscopy. Images were taken on either a Zeiss Axio Observer D1 bright field or Zeiss LSM710 confocal microscope. The Axio Observer D1 offered greater throughput but sacrificed the spatial resolution that the LSM710 could provide. Samples were prepared from a mid-log culture (grown for ∼4 h to 0.6 OD600). Cultures were pelleted in a microcentrifuge at 8000 rpm for 1 min, then resuspended in a tenth of the original volume in PBS. A coverslip was pre-treated by adding a 2 mg/mL solution of concanavalin A onto it, incubating for 10 min, and then washing with PBS. Then, 2.5 μL of the concentrated culture was added to a glass slide and covered with the concanavalin A-coated coverslip.

Axio Observer D1 images were taken at 1000 ms exposure for the mKate channel, 1000 ms for the Venus channel, 200 ms for the mTurquoise2 channel, 1000 ms for the mTagBFP2 channel, and 25 ms for the DIC image. Zeiss LSM710 images were taken at a scan speed of 4 and an averaging mode of 8. Gain for each channel was set to subsaturation values for each sample. Microscopy files were viewed and then exported to images in Fiji under an auto-exposure setting. Several images were taken per slide, and the fields of view were spaced to prevent photobleaching. Each micrograph was contrast-adjusted for clarity.

■ ASSOCIATED CONTENT

Supporting Information

Additional data from experiments described in the text; additional characterization of the coiled-coils; and additional illustrations for the Yeast Toolkit Golden Gate assembly method, tables that list sequences of fluorescent proteins, promoters, terminators, and coiled-coils used in the experiment, and sequence files in GenBank format for these sequences. The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acssynbio.5b00060.

■ AUTHOR INFORMATION

Corresponding Authors

*Tel: 510-643-4616. Fax: 510-642-9725. E-mail: jdueber@berkeley.edu.

*Tel: 510-394-5523. Fax: 510-642-9725. E-mail: wideloache@berkeley.edu.

Author Contributions

†R.C. and H.S.R. contributed equally to this work.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The authors thank S. Ruzin, D. Schichnes, and K. Thorn for assistance with microscopy. They also thank J. C. Anderson and members of the Dueber and Keating Laboratories for feedback and discussion. Early funding for the 2012 Berkeley iGEM team, consisting of R.C., H.S.R., M.R.Y., V.J.Y., T.C., C.L.C., A.T.J., and W.C.D., was provided by Agilent Technologies, IDT, SynBERC, and SBI. Further funding for R.C. was provided by the Haas Scholars Program. Funding for J.E.D. was provided by NSF CAREER award MCB-1151195. Funding to A.E.K for coiled-coil design was provided by NIH Award GM067681.

■ REFERENCES

(1) Arkin, A. P. (2013) A wise consistency: Engineering biology for conformity, reliability, predictability. *Curr. Opin. Chem. Biol. 17*, 893−901.

(2) Livet, J., Weissman, T. a, Kang, H., Draft, R. W., Lu, J., Bennis, R. a, Sanes, J. R., and Lichtman, J. W. (2007) Transgenic strategies for combinatorial expression of fluorescent proteins in the nervous system. *Nature 450*, 56−62.

(3) Weber, K., Thomaschewski, M., Warlich, M., Volz, T., Cornils, K., Niebuhr, B., Täger, M., Lütgehetmann, M., Pollok, J.-M., Stocking, C., Dandri, M., Benten, D., and Fehse, B. (2011) RGB marking facilitates multicolor clonal cell tracking. *Nat. Med. 17*, 504−509.

(4) Krutzik, P., and Nolan, G. (2006) Fluorescent cell barcoding in flow cytometry allows high-throughput drug screening and signaling profiling. *Nat. Methods 3*, 361.

(5) Mali, P., Aach, J., Lee, J.-H., Levner, D., Nip, L., and Church, G. M. (2013) Barcoding cells using cell-surface programmable DNA-binding domains. *Nat. Methods 10*, 403−6.

(6) Purcell, O., Peccoud, J., and Lu, T. K. (2014) Rule-Based Design of Synthetic Transcription Factors in Eukaryotes. *ACS Synth. Biol. 3*, 737.

(7) Božič, S., Doles, T., Gradišar, H., and Jerala, R. (2013) New designed protein assemblies. *Curr. Opin. Chem. Biol. 17*, 940−945.

(8) Selgrade, D. F., Lohmueller, J. J., Lienert, F., and Silver, P. a. (2013) Protein scaffold-activated protein trans-splicing in mammalian cells. *J. Am. Chem. Soc. 135*, 7713−7719.

(9) Zarrinpar, A., Park, S.-H., and Lim, W. a. (2003) Optimization of specificity in a cellular protein interaction network by negative selection. *Nature 426*, 676−680.

(10) Schreiber, G., and Keating, A. E. (2011) Protein binding specificity versus promiscuity. *Curr. Opin. Struct. Biol. 21*, 50−61.

(11) Reinke, A. W., Grant, R. a, and Keating, A. E. (2010) A synthetic coiled-coil interactome provides heterospecific modules for molecular engineering. *J. Am. Chem. Soc. 132*, 6025−31.

(12) Weber, E., Engler, C., Gruetzner, R., Werner, S., and Marillonnet, S. (2011) A modular cloning system for standardized assembly of multigene constructs. *PLoS One 6*, e16765.

(13) Lee, M. E., DeLoache, W. C., Cervantes, B., and Dueber, J. E. (2015) A Highly-characterized Yeast Toolkit for Modular, Multi-part Assembly. *ACS Synth. Biol.*, DOI: 10.1021/sb500366v.

(14) Shoemaker, B. a., and Panchenko, A. R. (2007) Deciphering protein-protein interactions. Part I. Experimental techniques and databases. *PLoS Comput. Biol. 3*, 0337−0344.

(15) Walton, P. a, Hill, P. E., and Subramani, S. (1995) Import of stably folded proteins into peroxisomes. *Mol. Biol. Cell 6*, 675−683.

(16) Havranek, J. J., and Harbury, P. B. (2003) Automated design of specificity in molecular recognition. *Nat. Struct. Biol. 10*, 45−52.

(17) Negron, C., and Keating, A. E. (2014) A Set of Computationally Designed Orthogonal Antiparallel Homodimers that Expands the Synthetic Coiled-Coil Toolkit. *J. Am. Chem. Soc. 136*, 16544−56.

(18) Nautiyal, S., Woolfson, D., King, D., and Alber, T. (1995) A designed heterotrimeric coiled coil. *Biochemistry 34*, 11645−51.

(19) Arndt, K. M., Müller, K. M., and Plückthun, a. (2001) Helix-stabilized Fv (hsFv) antibody fragments: substituting the constant domains of a Fab fragment for a heterodimeric coiled-coil domain. *J. Mol. Biol. 312*, 221−228.

(20) Mason, J. M. (2010) Design and development of peptides and peptide mimetics as antagonists for therapeutic intervention. *Future Med. Chem. 2*, 1813−1822.

(21) Thompson, K. E., Bashor, C. J., Lim, W. a, and Keating, A. E. (2012) SYNZIP Protein Interaction Toolbox: in Vitro and in Vivo Specifications of Heterospecific Coiled-Coil Interaction Domains. *ACS Synth. Biol. 1*, 118−129.

(22) Potapov, V., Kaplan, J. B., and Keating, A. E. (2015) Data-Driven Prediction and Design of bZIP Coiled-Coil Interactions. *PLoS Comput. Biol. 11*, e1004046.

(23) Reinke, A. W., Baek, J., Ashenberg, O., and Keating, A. E. (2013) Networks of bZIP protein-protein interactions diversified over a billion years of evolution. *Science 340*, 730−4.

DOI: 10.1021/acssynbio.5b00060
ACS Synth. Biol. XXXX, XXX, XXX−XXX

(24) Brückner, A., Polge, C., Lentze, N., Auerbach, D., and Schlattner, U. (2009) Yeast two-hybrid, a powerful tool for systems biology. *Int. J. Mol. Sci. 10*, 2763−2788.
(25) Fields, S. (2005) High-throughput two-hybrid analysis: The promise and the peril. *FEBS J. 272*, 5391−5399.
(26) Guthrie, C.; Fink, G. R. *Guide to Yeast Genetics and Molecular Biology*; Methods in Enzymology; Elsevier: 1991.
(27) Grigoryan, G., and Keating, A. E. (2008) Structural specificity in coiled-coil interactions. *Curr. Opin. Struct. Biol. 18*, 477.
(28) Mattheakis, L. C., Dias, J. M., Choi, Y.-J., Gong, J., Bruchez, M. P., Liu, J., and Wang, E. (2004) Optical coding of mammalian cells using semiconductor quantum dots. *Anal. Biochem. 327*, 200−8.
(29) Delehanty, J. B., Bradburne, C. E., Susumu, K., Boeneman, K., Mei, B. C., Farrell, D., Blanco-Canosa, J. B., Dawson, P. E., Mattoussi, H., and Medintz, I. L. (2011) Spatiotemporal multicolor labeling of individual cells using peptide-functionalized quantum dots and mixed delivery techniques. *J. Am. Chem. Soc. 133*, 10482−9.
(30) Lemon, C. M., Curtin, P. N., Somers, R. C., Greytak, A. B., Lanning, R. M., Jain, R. K., Bawendi, M. G., and Nocera, D. G. (2014) Metabolic tumor profiling with pH, oxygen, and glucose chemosensors on a quantum dot scaffold. *Inorg. Chem. 53*, 1900−1915.
(31) Saito, M., and Saitoh, H. (2012) Labeling of polyethylenimine with fluorescent dye to image nucleus, nucleolus, and chromosomes in digitonin-permeabilized HeLa cells. *Biosci., Biotechnol., Biochem. 76*, 1777−80.
(32) Cheng, a. a., Ding, H., and Lu, T. K. (2014) Enhanced killing of antibiotic-resistant bacteria enabled by massively parallel combinatorial genetics. *Proc. Natl. Acad. Sci. U. S. A. 111*, 12462.
(33) Bredies, K., and Wolinski, H. (2011) An active-contour based algorithm for the automated segmentation of dense yeast populations on transmission microscopy images. *Comput. Visualization Sci. 14*, 341−352.
(34) Gietz, R. D., and Woods, R. A. (2006) Yeast transformation by the LiAc/SS Carrier DNA/PEG method. *Methods Mol. Biol. 313*, 107−120.
(35) Wingler, L. M., and Cornish, V. W. (2011) Reiterative Recombination for the in vivo assembly of libraries of multigene pathways. *Proc. Natl. Acad. Sci. U. S. A. 108*, 15135−40.

DOI: 10.1021/acssynbio.5b00060
ACS Synth. Biol. XXXX, XXX, XXX−XXX