Reporter system architecture affects measurements of noncanonical amino acid incorporation efficiency and fidelity

Journal:	*Molecular Systems Design & Engineering*
Manuscript ID	ME-ART-08-2019-000107.R2
Article Type:	Paper
Date Submitted by the Author:	09-Jan-2020

Complete List of Authors:	Potts, Kelly; Tufts University, Chemical and Biological Engineering Stieglitz, Jessica; Tufts University, Chemical and Biological Engineering Lei, Ming; Tufts University, Chemical and Biological Engineering Van Deventer, James; Tufts University, Chemical and Biological Engineering; Tufts University, Biomedical Engineering

Design, System, Application Paragraph

On earth, the genetic code provides nearly invariant instructions for generating the proteins present in all organisms using 20 primary amino acid building blocks. Scientists and engineers have long recognized the potential power of altering the genetic code to introduce amino acids that enhance the chemical versatility of proteins. Proteins containing such “noncanonical amino acids” (ncAAs) can be used to elucidate basic biological phenomena, discover new therapeutics, or engineer new materials. However, tools for measuring ncAA incorporation during protein translation (reporters) exhibit highly variable properties including intrinsic support of ncAA incorporation and reporter expression levels, limiting our ability to engineer improved ncAA incorporation systems. In this work, we sought to understand what properties of these reporters affect measurements of ncAA incorporation events. Using a series of ncAA incorporation systems in yeast, we evaluated reporter architecture, measurement techniques, and alternative data analysis methods. We identified key factors contributing to quantification of ncAA incorporation in all of these categories and demonstrated the immediate utility of our approach in identifying genomic knockouts that enhance ncAA incorporation efficiency. Our findings have important implications for how to evolve cells to better accommodate alternative genetic codes.

Title

Reporter system architecture affects measurements of noncanonical amino acid incorporation efficiency and fidelity

Authors

Potts, K.A.,† Stieglitz, J.T.,† Lei, M.,† and Van Deventer, J.A.†‡*

†Chemical and Biological Engineering Department, Tufts University, Medford, Massachusetts 02155, United States
‡Biomedical Engineering Department, Tufts University, Medford, Massachusetts 02155, United States

*Corresponding author

Abstract

The ability to genetically encode noncanonical amino acids (ncAAs) within proteins supports a growing number of applications ranging from fundamental biological studies to enhancing the properties of biological therapeutics. Currently, our quantitative understanding of ncAA incorporation systems is confounded by the diverse set of characterization and analysis approaches used to quantify ncAA incorporation events. While several effective reporter systems support such measurements, it is not clear how quantitative results from different reporters relate to one another, or which details influence measurements most strongly. Here, we evaluate the quantitative performance of single-fluorescent protein reporters, dual-fluorescent protein reporters, and cell surface-displayed protein reporters of ncAA insertion in response to the TAG (amber) codon in yeast. While different reporters support varying levels of apparent readthrough efficiencies, flow cytometry-based evaluations with dual reporters yielded measurements exhibiting consistent quantitative trends and precision across all evaluated conditions. Further investigations of dual-fluorescent protein reporter architecture revealed that quantitative outputs are influenced by stop codon location and N- and C-terminal fluorescent protein identity. Both dual-fluorescent protein reporters and a “drop-in” version of yeast display support quantification of ncAA incorporation in several single-gene knockout strains, revealing strains that enhance ncAA incorporation efficiency without compromising fidelity. Our studies reveal critical details regarding reporter system performance in yeast and how to effectively deploy such reporters. These findings have substantial implications for how to engineer ncAA incorporation systems—and protein translation apparatuses—to better accommodate alternative genetic codes for expanding the chemical diversity of biosynthesized proteins.

1 Introduction

Genetically encoding noncanonical amino acids (ncAAs; also referred to as unnatural amino acids (uAAs), nonstandard amino acids (nsAAs), or nonnatural amino acids (nAAs)) in proteins enables control over protein structure and function with atomic-level precision.[1-5] Effective exploitation of ncAAs enhances our understanding of basic biology[6-8] and provides opportunities for engineering new classes of materials[9] and biological therapeutics.[2, 10-12] Many of these applications require high efficiency, high fidelity ncAA incorporation and subsequent careful evaluation of such events. The use of mass spectrometry-based characterizations offers the highest level of rigor,[9, 13] but lacks the throughput needed for initial screening and characterization. Additionally, mass spectrometry methods are not suitable for direct monitoring of incorporation events during protein translation. Another method for evaluating ncAA incorporation utilizes protein reporter systems in cells and cell-free translation systems as tools for understanding ncAA incorporation events, although the deployment of these systems varies widely. Even basic fluorescent reporters, where fluorescence is observed if a noncognate codon is suppressed, can possess drastic architectural differences between studies. These include the fluorescent protein variant utilized, the position within the reporter at which the ncAA is encoded, and data collection, analysis, and reporting methods.[14, 15] Due in part to these structural variations, it remains unclear how differences in reporters affect quantitative characterizations of ncAA incorporation events. Using the protein translation apparatus to insert a ncAA into a protein sequence via codon suppression (stop codon,[16, 17] 4-base codon,[18] or codon containing unnatural bases[19]) is complex, and usually inefficient compared to wild-type protein translation. Numerous studies have shown that these inefficiencies can result from, but are not limited to, the activities of engineered aminoacyl-tRNA synthetase/tRNA pairs (i.e. orthogonal translation systems; OTSs),[9, 20, 21] intracellular expression levels of OTS components,[22, 23] activities of the ribosome,[24] activities of elongation and release factors,[25, 26] and the codon composition of the host genome.[27] Elucidation of how these factors interact with one another is highly desirable for engineering translation apparatuses to accommodate alternative genetic codes. Integration of these observations and comparisons across studies requires a full understanding of how reporter systems and data analysis practices affect quantitative measurements of ncAA incorporation events.

Several reporter strategies for evaluating ncAA incorporation efficiency and fidelity have been described in the literature.[28, 29] By far the most common approach is single-fluorescent protein reporters.[14, 15] The primary advantage of these reporters is the easy-to-read fluorescent output, which in most cases is strongly correlated to the level of ncAA incorporation. However, readouts from these systems can be confounded by variability in intracellular plasmid levels or other processes that change reporter expression levels without altering suppression efficiency.[29] In addition, it remains unclear how variations in the properties of these reporters, such as changing stop codon position, affects the quantitative evaluation of ncAA incorporation events. Recently described dual-fluorescent protein reporters, which consist of two fluorescent proteins with distinct spectral properties connected by a linker, have some inherent advantages over single-fluorescent protein reporters.[29] Because these constructs provide a means of detecting both the expression level of the reporter (N-terminal fluorescent protein prior to codon for suppression) and full-length protein (C-terminal fluorescent protein), variations in reporter system expression can be accounted for during analysis. Barrick and coworkers introduced the metrics “Relative Readthrough Efficiency” and “Maximum Misincorporation Frequency” for quantifying the efficiency and fidelity of ncAA incorporation, respectively, while normalizing for changes in reporter expression levels.[29] Both single- and dual-fluorescent reporters support moderate to high throughput measurements with microplate readers and flow cytometry. One potential weakness of single- and dual-fluorescent reporters is that the folding times of fluorescent proteins may confound accurate determination of codon suppression efficiency. The use of epitope tags or conjugation reactions eliminate the potential for fluorescent protein folding rates

to confound analysis. However, a primary drawback is the need to label the displayed constructs of interest with suitable detection reagents prior to quantitative evaluation. Recent reports have demonstrated the use of cell surface display systems for evaluating codon suppression events.[28, 30] We recently showed that detection of the N- and C-termini of yeast-displayed constructs facilitates the use of the rigorous relative readthrough efficiency and maximum misincorporation frequency metrics described above. In addition, the surface accessibility of the reporter enables the use of chemical modifications to confirm the presence of a ncAA containing a specified reactivity, reminiscent of earlier residue-specific ncAA incorporation engineering work.[31] Söll and coworkers implemented the use of an *E. coli* display system to screen for aminoacyl-tRNA synthetase (aaRS) variants that support ncAA incorporation based on full-length protein expression and selective chemical modification to identify the presence of a specific ncAA,[32] but did not report quantitative measures of incorporation with this system.

Enzyme reporters of codon suppression that enable colorimetric readouts of enzyme activities have also been implemented in yeast and *E. coli*.[17, 33] Like fluorescent reporters, these enzymatic reporters, such as β-galactosidase reporters, decouple codon suppression events from cell survival, and provide a means of evaluating relative levels of suppression activity. Coupling codon suppression events to cell survival has been utilized in both prokaryotic and eukaryotic cells.[17, 34] These life-or-death assay formats support positive and negative selections and the ability to tune selection stringencies. On the other hand, these assays do not support quantitative measurements of ncAA incorporation efficiency or fidelity. The varied properties of the reporters described above suggests that each system has a role to play in discovering and evaluating ncAA incorporation systems. However, it remains unclear how results from distinct systems can be compared with one another due to significant differences in reporter system design, data collection, and data analysis.

In this study, we investigated the performance of three types of reporter systems that support fluorescence-based measurements of ncAA incorporation efficiency and fidelity. Our work here is conducted in *S. cerevisiae*, the only organism in which quantitative measurements of ncAA incorporation efficiency and fidelity with single-fluorescent protein reporters,[15, 35, 36] dual-fluorescent protein reporters,[28] and display-based reporters[28] have all previously been performed (Fig. 1). We compared reporters constructed in these three formats using flow cytometry- and microplate-based measurements (when possible) to evaluate ncAA incorporation efficiency and fidelity. In our hands, flow cytometry-based measurements led to more precise measurements than microplate-based measurements across all systems tested. Examination of a series of ncAA incorporation events known to exhibit a range of efficiencies and fidelities yielded similar trends in each reporter format. However, observed levels of ncAA incorporation efficiency and fidelity varied as a both a function of the system used and the method of downstream analysis. Based on these results, we constructed a series of dual-fluorescent protein reporters to better understand the effects of varying the fluorescent proteins utilized, orientation of proteins within the reporter, and TAG codon location and number. We then investigated the utility of several reporters for assessing ncAA incorporation events in a series of yeast knockout strains harboring genomic deletions of nonessential genes known to affect protein translation. Multiple strains supported enhanced ncAA incorporation efficiency without apparent loss of fidelity. We also found that controlling for changes in wild-type reporter expression levels is critical to determining whether a genomic modification is attributable to changes in codon suppression efficiency. These findings highlight the utility of these reporters in evaluating how the protein translation apparatus can be engineered to better support the use of alternative genetic codes and should support genome engineering efforts to construct and evolve organisms that utilize such codes. Taken as a whole, our results provide important

insights into how to effectively deploy reporter systems in search of ncAA incorporation systems that expand the chemical versatility of proteins.

2 Materials and Methods

2.1 Materials

All restriction enzymes used for molecular biology were from New England Biolabs (NEB). Synthetic oligonucleotides for cloning and sequencing were purchased from Eurofins Genomics or GENEWIZ. All sequencing in this work was performed by Eurofins Genomics (Louisville, KY) or Quintara Biosciences (Cambridge, MA). Epoch Life Science GenCatch™ Plasmid DNA Mini-Prep Kits were used for plasmid DNA purification from *E. coli*. Yeast chemical competent cells and subsequent transformations were prepared using Zymo Research Frozen-EZ Yeast Transformation II kits. *O*-methyl-L-tyrosine and *p*-azido-L-phenylalanine were purchased from Chem-Impex International, Inc. (catalog numbers 06251 and 06162, respectively).

2.2 Media preparation and yeast strain construction

The preparation of liquid and solid media was performed as described previously.[28] The strain RJY100 was constructed using standard homologous recombination approaches and has been described in detail previously.[37] The strain BY4741 (YSC1048) was purchased from Dharmacon. The BY4741 knockout strains from the Yeast Knockout Collection were obtained from the laboratory of Stephen P. Fuchs at Tufts University. The BY4705 strains (shown in Fig. S12) were obtained from the laboratory of Catherine Freudenreich at Tufts University (stock numbers 483 and 646 in the Freudenreich Lab) and were originally purchased from ATCC (*Saccharomyces cerevisiae* ATCC 200869™).

2.3 Reporter plasmid construction

The pCTCON2-FAPB2.3.6, pCTCON2-FAPB2.3.6L1TAG, pCTCON2-RYG, and pCTCON2-RXG reporter constructs have been previously described.[28] The pCTCON2-BXG reporter was cloned by replacing the RFP segment in pCTCON2-RXG with a BFP gene amplified from pBAD-mTagBFP2, obtained from the laboratory of Nikhil U. Nair at Tufts University and originally purchased from Addgene (Addgene plasmid # 34632 ; http://n2t.net/addgene:34632 ; RRID:Addgene_34632), by digesting both pCTCON2-RXG and the PCR-amplified BFP gene with EcoRI-HF and BamHI-HF (NEB), then ligating with T4 DNA ligase (NEB) to insert BFP. The pCTCON2-BXG-2TAG and pCTCON2-BXG-altTAG constructs were cloned by Gibson assembly with EcoRI-HF- and PstI-HF-digested pCTCON2-BXG while the pCTCON2-BXG-altTAG2, pCTCON2-BXG-altTAG3, and pCTCON2-BXG-altTAG4 constructs were cloned by Gibson assembly with EcoRI-HF- and PstI-HF-digested pCTCON2-BYG. Primers for the altTAG reporter were designed to introduce the alternative TAG codon at the first serine residue in the linker sequence between the BFP and sfGFP genes and revert the TAG codon at the end of the linker back to the tyrosine residue in the wild-type linker sequence.[29] Primers for the additional three alternate TAG locations were designed to introduce the TAG codon at either the second serine residue (altTAG2), the third alanine residue (altTAG3) or the third serine residue (altTAG4) found within the linker sequence between the two fluorescent proteins (Fig. 4A).

Primers for the 2TAG reporter introduced the TAG codon at the first serine residue in the linker and maintained the second TAG codon at the original location in pCTCON2-BXG. The pCTCON2-GXB and pCTCON2-GYB reporter constructs were cloned by amplifying BFP and sfGFP from pCTCON2-BXG with primers designed to maintain the same linker sequence with and without the TAG codon positioned at the original location in the linker, then cloned into EcoRI-HF- and BglII-digested pCTCON2 via Gibson assembly. The first amino acid in sfGFP, which was an alanine in the pCTCON2-BXG/BYG constructs, was reverted back to methionine. The pCTCON2-GFP constructs were cloned by amplifying pCTCON2-BXG with primers to revert the first residue in sfGFP back to methionine, then cloned into EcoRI-HF- and BglII-digested pCTCON2 via Gibson assembly. For the pCTCON2-GFP-TAG construct, additional primers were used to introduce a stop codon in place of tyrosine at the 151st amino acid position of the construct. The PCR products corresponding to the two sfGFP fragments before and after the 151st amino acid were cloned into pCTCON2 via Gibson assembly. pCTCON2-Aga1p-FAPB2.3.6 and pCTCON2-Aga1p-FAPB2.3.6L1TAG were constructed in two steps. First, the Aga1p gene was amplified from YIP SHRPa-Aga1p and cloned via Gibson assembly into pCTCON2 digested with restriction enzymes AgeI and KpnI. YIP sHRPa-Aga1P was a gift from Alice Ting (Addgene plasmid # 73151; http://n2t.net/addgene:73151; RRID:Addgene_73151). The resulting plasmid, pCTCON2-Aga1p, was sequence verified. The FAPB2.3.6 and FAPB2.3.6L1TAG genes were amplified from pCTCON2-FAPB2.3.6 and pCTCON2-FAPB2.3.6L1TAG, respectively, and then cloned via Gibson assembly into the pCTCON2-Aga1p vector digested with restriction enzymes BamHI-HF and NheI-HF. The resulting plasmids, pCTCON2-Aga1p-FAPB2.3.6 and pCTCON2-Aga1p-FAPB2.3.6L1TAG, were sequence verified. pRS416-Aga1p-FAPB2.3.6 was constructed by amplifying the Aga1p-FAPB2.3.6 segment from pCTCON2-Aga1p-FAPB2.3.6 and using a Gibson assembly to insert the fragment into pRS416 digested with restriction enzymes XbaI and SalI-HF. The sequence verification revealed a point mutation in the Aga1p gene that was subsequently removed. The TAG version of the pRS416-Aga1p-FAPB2.3.6 reporter was made by cloning in a TAG codon at the first position of the light chain of the scFv in the same position as the other yeast display reporter plasmids. Resulting plasmids were sequence verified. The promoter and BXG or BYG DNA fragments were amplified from pCTCON2-BXG and pCTCON2-BYG and cloned into XbaI- and SalI-HF-digested pRS416 via Gibson assembly for the pRS416-BXG and pRS416-BYG plasmids, respectively. pRS416-BXG-altTAG was constructed by amplifying the BFP-GFP DNA fragment with the alternative TAG codon from pCTCON2-BXG-altTAG and inserting it via Gibson assembly into pRS416 double digested with XbaI and SalI-HF. Resulting plasmids were sequence verified. Sequences of all primers used for cloning described in this work are listed in Table S13.

2.4 Suppressor plasmid construction

Suppressor plasmids pRS315-OmeRS[30, 37] containing the tyrosyl OmeRS and pRS315-LeuOmeRS[28] containing the leucyl OmeRS have been previously reported and characterized in detail.

2.5 Preparing noncanonical amino acid liquid stocks

All ncAA stocks were prepared at a 50 mM concentration of the L-isomer. DI water was added to the solid ncAA to approximately 90% of the final volume, and 6.0 N NaOH was used to fully dissolve the ncAA powder in the water by vortexing. Water was added to the final volume and the solution was sterile filtered through a 0.2 micron filter. OmeY solutions were pH adjusted to

7 using HCl. Filtered solutions were stored at 4°C for up to one week for AzF and two weeks for OmeY prior to use.

2.6 Yeast transformations, propagation, and induction

Reporter construct plasmids containing either a TRP1 (pCTCON2) or URA3 (pRS416) marker and aaRS/tRNA suppression plasmids pRS315-OmeRS and pRS315-LeuOmeRS (LEU2 marker) were transformed simultaneously into Zymo competent *S. cerevisiae* strains RJY100, BY4705 483, BY4705 646, BY4741, and the BY4741 deletion strains, plated on solid SD-SCAA media (either −TRP −LEU, −LEU −URA, or −TRP depending on the combination of plasmids), and grown at 30°C until colonies appeared (3 days).

Biological triplicates were used for all quantitative measurement experiments. All cells were grown and induced in tubes regardless of whether readthrough was assessed on a flow cytometer or on a plate reader. All liquid cultures were supplemented with a 100X penicillin/streptomycin to a final concentration of 1X (Corning 100X Penicillin:Streptomycin solution) to decrease the probability of contamination. To propagate samples with biological replicates and prepare them for induction, three separate colonies from each transformation were inoculated in 5 mL selective media and allowed to grow to saturation at 30°C (2−3 days). For cases where liquid colonies were already available, samples from saturated cultures stored at 4°C were pelleted and resuspended to an OD_{600} of 0.5−1.0 in 5 mL fresh media and allowed to grow to saturation overnight. Following saturation, the cultures were diluted to an OD_{600} of 1 in fresh media and grown at 30°C until reaching mid log phase (OD 2−5; 4−8 h). Cells were pelleted (5 min at 2,400 rpm) and resuspended to an OD_{600} of 1 in induction media (cells containing reporter construct only: SG-SCAA (−TRP); cells containing both reporter constructs and suppression constructs: either SG-SCAA (−TRP −LEU), SG-SCAA (−LEU −URA), or SG-SCAA (−TRP −LEU −URA), depending on the combination of suppressor and reporter plasmids in each yeast strain). To enable site-specific incorporation of ncAAs, induction media was supplemented with 1 mM final concentration of the L-isomer of the following ncAAs: *O*-methyl-L-tyrosine (pH 7) and *p*-azido-L-phenylalanine, and then induced at 20°C for 16 h.

2.7 Flow cytometry data collection and analysis

Freshly induced samples were labeled in 1.7 mL microcentrifuge tubes or 96-well V-bottom plates. Flow cytometry was performed either on an Attune NxT flow cytometer (Life Technologies) at the Tufts University Science and Technology Center or on a BD™ LSR II (BD Biosciences) at the Tufts University Flow Cytometry Core in the Jaharis Building. Labeling of induced yeast cultures with antibodies for detection of the N- and C-terminal epitope tags has been previously described in detail[28] and was not modified for these experiments.

2.8 Plate reader data collection

To measure RFP and sfGFP levels for fluorescent protein reporters co-transformed with suppression constructs, 2 million cells per sample of freshly induced cells were pelleted (5 min at 2,400 rpm) and washed 3 times with 1X PBSA in 96-well V-bottom plates and then transferred to Corning 96-well clear bottom black-walled microplates for fluorescence measurements. Cultures containing pCTCON2-FAPB2.3.6 with no suppressor were used to measure the autofluorescence of the cells (cell blank), as they were not expected to exhibit expression of RFP, BFP, or sfGFP. All samples, including the cell blank, were run in biological

triplicate and resuspended in 200 µL room temperature PBSA before measurements were taken. Fluorescence and OD measurements were performed using a SpectraMax i3X microplate reader (Molecular Devices, LLC., San Jose, California). OD readings were taken as end point measurements at 600 nm. RFP, GFP, and BFP readings were taken as end point measurements with RFP excitation and emission wavelengths set at 550 nm and 675 nm, respectively. The GFP excitation and emission wavelengths were set to 480 nm and 525 nm, respectively. The BFP excitation and emission wavelengths were set to 399 nm and 456 nm, respectively.

2.9 Calculating RRE and MMF

Detailed methods for flow cytometry RRE and MMF data analyses for yeast-displayed reporter constructs, including error propagation, has been described previously.[28] Dual-fluorescent reporter RRE and MMF analyses were performed similarly, replacing HA and c-Myc detection with N-terminal and C-terminal fluorescent protein detection. Fraction of wild-type (Fraction WT) and misincorporation of the sfGFP reporter were calculated by replacing the dual detection with single detection of sfGFP and comparing the TAG-containing constructs to wild-type sfGFP expression under the same media conditions (i.e. in the presence or absence of ncAA) using the equations provided in Fig. 1B. Both this work and our previous report use the Microsoft Excel function “STDEV” to determine standard deviation of samples measured in biological triplicate.

Microplate reader data analysis was performed using Microsoft Excel. The fluorescence from each sample was normalized by the sample’s respective OD_{600} and then averaged across the biological triplicates. The average normalized fluorescence of the cell blank triplicates was then subtracted from the normalized fluorescence sample average to correct for yeast cell autofluorescence. For dual fluorescent protein reporters, both the N-terminal and C-terminal proteins were taken into account for RRE and MMF calculations, whereas the fraction WT and misincorporation of the sfGFP reporter were determined using single detection and the equations from Fig. 1B.

2.10 RRE error as a percent of the magnitude

To determine error as a percent of the magnitude, the error-propagated standard deviations were divided by the magnitude of the relative readthrough efficiencies. These fractions were then converted to percentages and reported in Table S2 and S3.

2.11 Alternative analysis methods

Alternate analyses of flow cytometry data were performed using FlowJo and Microsoft Excel. For each sample collected on the flow cytometer, the overall population was gated for live, single cell events to exclude doublet and triplet data from the downstream analysis. From these single cell populations we performed the first set of alternate efficiency calculations (“Single Cell Population”) by averaging the median fluorescence intensity (MFI) data from the C-terminal fluorescent protein and taking the standard deviation of the biological triplicate (Fig. S2, S14, S21, S28). The second alternate analysis, “Reporter Expressing Cells + Background Subtraction,” utilized MFI values for C-terminal detection in cells expressing the reporter (i.e. cells demonstrating above-background levels of N-terminal fluorescent protein detection in the case of the dual reporters) and MFI values in cells not expressing a reporter (nonexpressing

cells). We then subtracted the nonexpressing-cell MFI values from MFI values for the subset of cells with above-background levels of N-terminal fluorescent protein detection. For the single-fluorescent protein reporters, the population was gated into cells exhibiting above-background sfGFP fluorescence and background-level sfGFP fluorescence. The MFI of the background-level sfGFP fluorescence population was subtracted from the above-background sfGFP population to obtain background-subtracted MFIs. The averages and standard deviations of the biological triplicates were then used to calculate propagated standard error (Fig. S3, S15, S22, S29). Equations used to propagate error are as previously described.[28] The last alternate analysis ("Reporter Expressing Cells + BG Subtraction Normalized to WT Reporter") takes the values as described in the second analysis method ("Reporter Expressing Cells + Background Subtraction") and reports the TAG-containing constructs as a fraction of the respective wild-type construct (without a TAG codon and in the presence of no ncAAs) efficiency (Fig. S4, S16, S23, S30). All three of these alternate analyses were calculated using median fluorescence intensity as well as mean fluorescence intensity (Fig. S6 − S8, S18 − S20, S25 − S27, S32 − S34).

Microplate reader alternate analyses were performed using Microsoft Excel. For the first of the alternate microplate calculations (OD Normalized), the fluorescence intensity of the C-terminal fluorescent protein measured in each sample was normalized to the sample's respective OD_{600} and then averaged with the associated biological triplicates. In the case of the single-fluorescent protein reporter, the fluorescence intensity of sfGFP was measured in place of the C-terminal fluorescence measurement (Fig. S9). For the second microplate reader analysis (OD Normalized + Background Subtraction), the normalized average fluorescence of the cell blank triplicates was subtracted from the average normalized fluorescence of the sample to correct the fluorescence signals for yeast cell autofluorescence (Fig. S10). The last of the alternate microplate reader analyses, "OD Normalized + BG Subtraction Normalized to WT Reporter," reported the previous TAG-containing construct efficiencies as a fraction of the related wild-type construct (without a TAG codon and in the presence of no ncAAs) efficiency (Fig. S11).

Finally, data from the dual-reporters presented in Fig. 2 and Fig. 5 were subjected to another analysis (Reporter Expressing Cell Population) in which the median fluorescence intensity of the N-terminal signal (i.e. HA epitope detection or N-terminal fluorescent protein detection) was reported for the wild-type construct in the absence of any ncAA as well as for the TAG constructs in the presence of OmeY, AzF or in the absence of ncAA (Fig. S5, S17, S24, S31).

In the statistical analysis (see below), four additional samples are shown in Tables S4 – S12 (wild-type + OmeY and wild-type + AzF for both aaRSs) in order to investigate how the presence of ncAAs within the induction media can affect wild-type reporter expression. For all of the alternate analyses described above, we used wild-type reporter expression data in the absence of any ncAAs instead of wild-type data in the presence of either OmeY or AzF, since statistical analysis suggested no significant effects on wild-type reporter expression levels in the presence of either OmeY or AzF.

2.12 Statistical analysis

One-way ANOVA (analysis of variance) tests were performed followed by groupings via the Games-Howell method (Minitab 18) to identify statistically distinct groups of collected data. Analysis was based off of populations of n=3, where each value was derived from a distinct single colony from a transformation. One-way ANOVA and groupings via the Games-Howell method were performed initially on the N-terminal, background-subtracted, median fluorescence intensity levels of sets of reporter-synthetase-ncAA combinations of interest to determine whether reporter expression levels were indistinguishable. We then performed ANOVA with the C-terminal, background-subtracted, median fluorescence intensity levels to determine the

Games-Howell groupings of data as described below. Tables S4 – S12 summarize the calculated 95% confidence intervals and groupings (in this approach, datasets that are not part of the same letter group possess statistically different sets of means). For Fig. 2 and Fig. 3, ANOVA and grouping via the Games-Howell method was used for the 6 OTS-ncAA combinations (LeuOmeRS + no ncAA, LeuOmeRS + OmeY, LeuOmeRS + AzF, TyrOmeRS + no ncAA, TyrOmeRS + OmeY, and TyrOmeRS + AzF) for each reporter (wild-type and TAG-containing constructs were analyzed together) and either the N- or C-terminal detection (Table S4 – S8). The data for Fig. 4 and Fig. 5 were grouped such that each of the aforementioned OTS-ncAA combinations were analyzed individually with respect to either all seven BFP-GFP reporter variants (Fig. 4, N- and C-terminal statistics summarized in Table S9) or with respect to all seven strain variants (Fig. 5, N-terminal statistics summarized in Table S10 – S12).

3 Results and Discussion

3.1 Framework for evaluating the performance of reporters of noncanonical amino acid incorporation in yeast

We initiated our studies of the performance of ncAA reporter systems by implementing reporters with architectures previously described in the literature (Fig. 1A): a yeast display-based reporter, a dual-fluorescent protein reporter, and a single-fluorescent protein reporter. Both the yeast display-based reporter and dual RFP-GFP reporter are identical to the reporters that we have described previously,[28, 29] with a TAG codon between two epitopes or proteins that support fluorescence detection. Moreover, the sfGFP used in the single-fluorescent protein reporter is genetically identical to the sfGFP in RFP-GFP. The amber codon in sfGFP was introduced at Y151, which is a commonly used permissive site for amber suppression.[14, 23] All three of these reporters are under the control of the inducible Gal 1-10 promoter and allow for the modular introduction of aminoacyl-tRNA synthetase/tRNA pairs comprising the orthogonal translation system (OTS) machinery. These reporters were introduced into yeast along with each of two previously reported OTSs that were originally engineered to incorporate *O*-methyl-L-tyrosine (OmeY) in response to TAG codons in yeast: an *E. coli* tyrosyl-tRNA synthetase/$tRNA^{Tyr}_{CUA}$ pair (TyrOmeRS)[38] and *E. coli* leucyl-tRNA synthetase/$tRNA^{Leu}_{CUA}$ pair (LeuOmeRS)[39] where the LeuOmeRS variant was modified further to contain a T252A mutation.[28] In previous work, we found that TyrOmeRS supports moderate levels of ncAA incorporation in the presence of either OmeY or *p*-azido-L-phenylalanine (AzF), with low but detectable levels of TAG codon readthrough in the absence of ncAAs. LeuOmeRS supports high levels of ncAA incorporation with OmeY, very low levels of ncAA incorporation with AzF, and essentially undetectable levels of readthrough in the absence of ncAAs. Thus, induction of the reporters under these different conditions provides a large expected range of readthrough efficiencies and fidelities for evaluation of reporter system performance.

All measurements of ncAA incorporation efficiency were conducted using end point measurements collected in biological triplicate, where the term “biological triplicate” refers to samples prepared from three separate colonies that were propagated following transformation with plasmids of interest. Following collection of data for all samples and controls using a flow cytometer (all reporters) or a microplate reader (fluorescent protein reporters), we determined ncAA incorporation efficiency and fidelity (Fig. 1B; see Materials and Methods for further details). For dual reporters (yeast display, dual-fluorescent protein reporters), we report in the main text the relative readthrough efficiency (RRE) and maximum misincorporation frequency (MMF) metrics as previously introduced by Barrick and coworkers.[29] These metrics are normalized to wild-type control constructs while also accounting for possible perturbations from

the presence of the ncAA or changes in expression of the reporter during induction. For single-fluorescent protein reporters, which do not support the use of RRE and MMF, we determined the fraction of wild-type sfGFP (Fraction WT) expression (Fig. 1B). For all measurements, the data we collected enabled us to investigate the effects of using other analysis methods (for example, fraction of wild-type sfGFP expression in RFP-GFP). We also used statistical tests to evaluate whether variability in the reporters and data collection methods affected our ability to distinguish between distinct ncAA incorporation events. Taken together, these reporter architectures, OTSs, and analyses provide a rigorous framework for evaluating the properties of reporters and their effects on ncAA incorporation efficiency and fidelity.

3.2 Comparisons of conventional reporter architectures

Fig. 2 depicts side-by-side comparisons of yeast display reporter, dual-fluorescent protein reporter, and single-fluorescent protein reporter measurements of ncAA incorporation efficiency and fidelity using the framework described above. Qualitatively, data collected with all reporters show the expected trends for ncAA incorporation efficiency and fidelity: LeuOmeRS + AzF < TyrOmeRS + OmeY ≅ TyrOmeRS + AzF < LeuOmeRS + OmeY. In addition, measurements determined via flow cytometry generally exhibit lower propagated error in comparison with the propagated error determined from microplate reader-based measurements. Flow cytometry measurements led to propagated errors determined to be at or below 19% of the magnitude of the RRE, whereas propagated error from measurements on the microplate reader was determined to be equal to or greater than 19% (Table S2). For flow cytometry-based measurements, a similar level of error was observed for reporters based on either yeast-displayed proteins or fluorescent proteins. We conducted one-way analysis of variation (ANOVA) with grouping via the Games-Howell method on the detection of C-terminal tags/fluorescent proteins to further evaluate how readily individual reporters and data collection methods distinguish between varying types of ncAA incorporation events and corresponding wild-type controls (Tables S4 – S8; the extensive error propagation in determining RRE and MMF precluded direct statistical analysis of these calculated values via one-way ANOVA, which does not account for error propagation). These analyses indicate that measurements of C-terminal detection levels obtained on a flow cytometer for all three reporters can be grouped into distinct high-, medium-, and low-efficiency readthrough events (Tables S4 – S8; high: LeuOmeRS + OmeY; medium: TyrOmeRS + AzF/TyrOmeRS + OmeY; low: LeuOmeRS + AzF). Measured fluorescence levels of wild-type reporters are also usually identified as statistically distinct from levels observed during suppression events. However, measurements of C-terminal detection levels obtained on a plate reader did not provide the same level of discrimination as observed with flow cytometry-based measurements, with groupings failing to separate high- and medium-efficiency events or medium- and low efficiency events into well-defined groups. These data and analyses indicate that flow cytometry can provide relatively precise measurements ncAA incorporation efficiency and fidelity with yeast display, dual RFP-GFP, and single sfGFP reporters.

Because RRE and MMF were only first used as metrics of ncAA incorporation processes in 2017, we conducted a series of alternative data analyses in order to explore the effects of data analysis methodologies on quantitative outputs. Consistent with strategies reported in the literature, we considered only the C-terminal reporter signal for each construct and used this as the basis for determining the “fraction of wild-type behavior.” For flow cytometry data, we determined the median and mean of the C-terminal reporter signals on gated populations consisting of all single cells, or on only the population of cells exhibiting evidence of reporter expression (see Materials and Methods for further details on analysis). Fig. S2 – S11 depict the results of our alternative analyses of the data collected for Fig. 2. For all three reporter

architectures, trends in efficiency data are generally consistent with those determined in Fig. 2 across the 4 OTS-ncAA combinations considered here. However, there are some cases in which calculated values of efficiency differ from the results depicted in Fig. 2. As an extreme example, the incorporation efficiency of the LeuOmeRS-OmeY combination determined by yeast display ranges from approximately 30% to 50% depending on the analysis method, greater than a 1.5-fold variation. However, in many other cases, these variations are within calculated error.

Measurements made with single-fluorescent protein reporters make the implicit assumption that the expression levels of reporter constructs do not change significantly when evaluating different OTS-ncAA combinations. To evaluate this assumption, we plotted the values of the N-terminal signal levels detected in all dual-detection samples characterized in Fig. 2 using the Reporter-Expressing Cell Population approach (Fig. S5). Under these conditions, reporter expression levels are reasonably consistent, but can exhibit greater than 20% variability between samples in the same data series. One-way analysis of variation (ANOVA) with grouping via the Games-Howell method revealed that most, but not all, of the samples evaluated for N-terminal signal levels are statistically indistinguishable from one another (Table S4, S5). This level of variation seems to be tolerable, as with flow cytometry measurements all reporter architectures considered here are able to distinguish between high, medium, and low efficiency OTS-ncAA combinations considered in this work. However, the RRE measurement framework eliminates the risk of inadvertently neglecting changes in reporter system expression levels. While variations are low under the conditions evaluated in this section, changes in expression can become statistically significant when evaluating incorporation events in different strains (in some cases, greater than 2-fold variation in reporter expression levels; see Section 3.5), suggesting that some caution should be exercised when using single-fluorescent protein reporters. Therefore, for the remainder of this manuscript, we have elected to utilize RRE and MMF as metrics for stop codon suppression efficiency and fidelity, respectively (we have also performed statistical analyses on background-subtracted median fluorescence intensity data presented in subsequent figures; see below). In addition to controlling for changes in reporter system expression levels, RRE provides an explicit comparison to wild-type protein translation, enabling immediate evaluation of how drastically a stop codon suppression event affects protein translation. As previously discussed by Barrick and coworkers, MMF provides an extremely stringent evaluation of canonical amino acid misincorporation. While conservative, this metric provides a sensitive means of identifying lower fidelity orthogonal translation systems. Given the high precision of the conventional protein translation machinery (approximate error rates of one in 1,000 or less)[40], such sensitivity may help facilitate high fidelity genetic code expansion.

A final noteworthy observation is that for a given OTS-ncAA induction condition, the calculated values of ncAA incorporation efficiency and fidelity depend on the specific reporter system used. Since ncAA incorporation efficiency and fidelity determined with all three reporters follow the same trends across all conditions examined, this indicates that the specific architecture of the reporter system dictates the quantitative values determined in a given experiment. This observation motivated the design of reporters containing subtle variations in architecture to better understand how these variations affect quantitative metrics of ncAA incorporation efficiency and fidelity.

3.3 Variation of fluorescent proteins used in dual-fluorescent protein reporters

Given the comparable precision of display- and fluorescence-based reporters, we evaluated several dual-fluorescent protein reporters in which only the identities of the fluorescent proteins utilized were changed to investigate the role of fluorescent protein folding properties on reporter

performance. Previous work raises the possibility that the long half maturation time of RFP in the RFP-GFP system could confound accurate determination of RRE and MMF (RFP maturation half-time in solution is on the order of one hour).[41] To evaluate this possibility, we replaced RFP with blue fluorescent protein (BFP)[42] to create a BFP-GFP dual-fluorescent protein reporter analogous to the RFP-GFP reporter (Fig. 3A). BFP has previously been reported to exhibit a maturation half-time in solution of approximately 12 minutes.[42] To our surprise, the RRE and MMF data we obtained using the two reporter systems appeared to be within propagated error of one another despite the different folding times of the N-terminal fluorescent proteins (Fig. 3B). We performed one-way ANOVA with grouping via the Games-Howell method on the background-substracted, median fluorescence intensity levels of the C-terminal fluorescent protein of all ncAA incorporation events measured with these reporters. This analysis indicates that, on a flow cytometer, both the RFP-GFP and BFP-GFP reporters can distinguish between high-, medium- and low-level ncAA incorporation events, as evidenced by their corresponding statistically distinct, separate groups (Tables S6 – S8). These separations are always not perfect, though, as some groupings include data from both “high” and “medium” or “medium” and “low” OTS/ncAA measurements. We also note the important caveat that the N-terminal fluorescence levels of the BFP-GFP reporters exhibited higher variability than in other experiments (shown here by the wider set of groups and 95% confidence intervals than other measured N-terminal fluorescence levels with BFP-GFP reporter variants; compare to data in Table S9). This cautionary observation underscores the value of dual-detection reporter systems in evaluating ncAA incorporation events. For the BFP-GFP system, we conducted measurements of RRE and MMF on multiple flow cytometers and observed similar quantitative values of efficiency, fidelity, and propagated error (Fig. S1; only one flow cytometer with optics suitable for evaluating the RFP-GFP system was readily available). We also switched the order of the BFP-GFP system to place sfGFP, which has a reported folding half-time in solution of under 1 minute,[43] in the N-terminal position while maintaining the structure of the linker (Fig. 3A). Flow cytometry readouts with the resulting GFP-BFP reporter result in calculated RRE and MMF values with similar trends to those determined with RFP-GFP and BFP-GFP systems. However, propagated error was determined to be much higher. In the case of LeuOmeRS with OmeY we observed that the calculated error of GFP-BFP was 43% of the RRE magnitude, whereas the error in the RFP-GFP and BFP-GFP systems was only 5.6% and 15% of the magnitude, respectively (Fig. 3B; Table S3). This trend also persisted for the other OTS-ncAA combinations evaluated (Table S3). One-way ANOVA with grouping via the Games-Howell method on C-terminal detection levels determined with the GFP-BFP detection system resulted in only a single group in an initial experiment (Table S6). This experiment was repeated, and the statistical analysis on the data (Table S7) again indicated that fluorescence values from different OTS/ncAA combinations exhibited too much variability to enable separation into distinct groups of readthrough events. It is not immediately clear why relocating a protein with an extremely fast folding rate to the N-terminus of a dual-fluorescent reporter should be detrimental to the performance of the reporter. In any case, these observations suggest that the fluorescent proteins in the dual reporter format cannot be treated as completely modular entities.

Our flow cytometry measurements with RFP-GFP and BFP-GFP reporters allow for reliable differentiation between several LeuOmeRS-ncAA and TyrOmeRS-ncAA incorporation events (Table S6). On the other hand, the large propagated error for experiments performed on the microplate reader makes reliable determination of differences in performance between OTS-ncAA combinations challenging (Fig. 3C; see Materials and Methods for calculation details). This is also reflected in the one-way ANOVA with grouping via the Games-Howell method, where high-, medium-, and low-level ncAA incorporation efficiency events determined by the flow cytometer are not separated into statistically distinct groupings with plate reader

measurements. This observation is consistent with the data presented in Section 3.2, our own previous report,[28] and that of Barrick and coworkers,[29] where online measurements over the course of multiple hours with varying numbers of technological replicates per condition were used to reduce error during determination of RRE and MMF. For the remainder of our studies, we report only data derived from flow cytometry experiments due to the high-precision determination of ncAA incorporation efficiency and fidelity supported by this technique. Since we observed similar performance of RFP-GFP and BFP-GFP dual-reporters for the OTS-ncAA combinations tested here (Figure 3, Table S6), we chose to use only BFP-GFP reporter variants to investigate the effects of altering stop codon positioning in the following section (this decision was also motivated by our more ready access to a flow cytometer supporting measurements of BFP-GFP reporters than to a flow cytometer supporting measurements with RFP-GFP).

3.4 Variation of stop codon position and number in the BFP-GFP dual-fluorescent protein reporter

Studies of nonsense suppression events in cells from several organisms have revealed that the context of a TAG codon, that is the bases flanking the codon, can affect stop codon readthrough efficiency.[44-48] These observations and more direct studies of the role of the bases flanking TAG codons targeted for ncAA incorporation in *E. coli*[49] highlight the need to investigate this issue further. Here, we examined a series of BFP-GFP reporters containing TAG codons at different positions for several reasons. First, for the same OTS-ncAA combination, the BFP-GFP reporter appears to support higher levels of ncAA incorporation in comparison to the yeast display reporter (Fig. 2). Second, in a previous study, we used the yeast display reporter to investigate three permissive stop codon locations within our antibody-based reporter and found ncAA incorporation efficiency to be consistent across the three positions, but lower than the BFP-GFP reporter efficiency observed in this study.[28] Fig. 4A depicts variants of the BFP-GFP reporter in which the position of the stop codon has been moved to several locations within the linker (Alt-TAG, Alt-TAG2, Alt-TAG3 and Alt-TAG4) or in which two stop codons have been introduced into the vector simultaneously (2-TAG). We selected the first serine residue (TCC) for the first alternate TAG location as well as for the dual-TAG reporter to minimize the number of bases to mutate to facilitate TAG codon introduction and to avoid consecutive or near-consecutive TAG codons in the 2-TAG case. The additional “Alt-TAG” reporter constructs position the stop codon at one of several locations distributed throughout the linker sequence (Fig. 4A). The RRE and MMF results in Fig. 4B clearly demonstrate that moving the position of the TAG codon to any of the alternate positions (Alt-TAG, Alt-TAG2, Alt-TAG3, or Alt-TAG4) results in lower suppression efficiencies with the combination of LeuOmeRS and OmeY (additional OTS-ncAA statistics presented in Table S9). One-way ANOVA with grouping via the Games-Howell method on the background-subtracted readthrough data confirm that the original TAG codon location supports statistically distinct, higher levels of stop codon readthrough in comparison to reporters containing stop codons in any of the “Alt-TAG” locations for the LeuOmeRS + OmeY, LeuOmeRS + AzF, and TyrOmeRS + AzF combinations. Introduction of two stop codons into the linker (at “standard” and “alternative” positions) does not drastically lower readthrough efficiency in comparison to single-TAG reporters at any of the “Alt-TAG” positions. These observations highlight the need for further studies to fully understand the effects of TAG number and location on readthrough efficiency. Taken together with our previous observations,[28] these data suggest that, at least in yeast, the original TAG codon position in the BFP-GFP (and RFP-GFP) reporter is especially permissive with respect to ncAA incorporation for most OTS/ncAA combinations examined in this work. The differences between ncAA incorporation efficiencies observed here highlight the importance of characterizing and understanding reporter systems, as seemingly small changes in stop codon positioning can have a large impact on reporter output. A high level of “permissiveness” in a reporter may also obscure the detection of minute

enhancements in ncAA incorporation efficiencies caused by genetic modification. To gauge the validity of this conjecture further, we conducted studies using strains from a single-gene knockout collection and various reporter systems to determine which reporters would support the detection of altered ncAA incorporation efficiency.

3.5 Characterization of single-gene knockouts using reporter systems

Genomic modifications ranging from single-gene knockouts[26] to complete overhaul of the genome[27] can enhance ncAA incorporation via codon suppression events. Therefore, we wanted to determine whether the reporters characterized in this work would be suitable for evaluating ncAA incorporation efficiency and fidelity in yeast strains possessing several genotypes. We selected six strains from the haploid yeast knockout collection (YKO) containing a deletion of a nonessential gene known to be associated with efficiency of termination of protein translation.[50] These strains and the parent BY4741 strain were co-transformed with OTSs and reporter systems and evaluated for ncAA incorporation efficiency and fidelity (Fig. 5). Because the strains of the knockout collection contain deletions of LEU2 and URA3, but not TRP1, we moved the BXG and Alt-TAG reporter systems from a pCTCON2 plasmid backbone into a pRS416 (URA3 marker) backbone. In addition, we prepared a “drop-in” version of our yeast display reporter within the pRS416 backbone suitable for use in any yeast strain (in contrast to conventional Aga1p-Aga2p yeast display, where Aga1p is integrated into the genome while Aga2p is encoded on pCTCON2). Fig. S12 provides a comparison of drop-in versus conventional yeast display using pCTCON2-based reporter plasmids (the yeast display strain RJY100 contains a genomically integrated URA3 marker, preventing use of pRS416-based plasmids in this strain). In BY4741, we observed that the LeuOmeRS-OmeY OTS-ncAA combination, which consistently yielded the highest relative readthrough efficiency in the pCTCON2 backbone (Fig. 2 – 4), exhibits similar RRE values to the TyrOmeRS-OmeY and TyrOmeRS-AzF OTS-ncAA combinations in the pRS416 backbone for all three reporters (Fig. 5A, 5B, Fig. S13 – S34). While clearly reproducible, the reason for this apparent shift in observed ncAA incorporation efficiencies is not clear. This unexpected change in relative performance of OTS-ncAA combinations upon switching plasmid backbones further emphasizes the significant changes in reporter system performance that can result from seemingly inconsequential alterations to the reporters.

Keeping these shifts in mind, we evaluated ncAA incorporation efficiency and fidelity in the parent strain and each of the knockout strains using the BXG, Alt-TAG, and drop-in display reporters. Fig. 5A depicts the results of these experiments using the Alt-TAG reporter, which exhibits the least efficient ncAA incorporation of the three reporters used here. Trends in efficiency, misincorporation, and propagated error observed using the Alt-TAG reporter were similar to the trends observed using the other two reporters (see Fig. 5B for selected data; Fig. S13 depicts the complete RRE and MMF data for all reporters and all strains tested here). The largest increase in RRE we observed here was in a strain lacking PPQ1; this was consistent among all three reporters (Fig. 5B, Fig. S13). PPQ1 is a protein phosphatase that has been anecdotally noted to impact stop codon readthrough, but to the best of our knowledge, its molecular role(s) in stop codon readthrough remains unknown.[50] The degree to which we observed enhancement of ncAA incorporation depended on which reporter was used: more than a 2-fold gain in efficiency was obtained with the Alt-TAG reporter when PPQ1 was deleted, in contrast to the more modest gains (roughly 50–60%) observed with the other two reporters. This confirms the notion that higher “permissiveness” in a reporter tends to show reduced gains in efficiency when evaluating factors that enhance ncAA incorporation events. Interestingly, in strains in which incorporation efficiency appeared to increase, the gains in efficiency did not

appear to come at the expense of fidelity. Maximum misincorporation frequency measurements indicated similar or reduced values in comparison to MMF values in the parent strain.

To understand why we observed these increases we conducted the same analyses as described in Section 3.2 for the data depicted in Fig. 5 (Fig. S14 – S34; see Materials and Methods for full details of analysis). In addition, we performed one-way ANOVA with grouping via the Games-Howell method to evaluate the role of changes in reporter expression, based on N-terminal fluorescence levels, on calculated RRE and MMF values (Table S10 – S12). These analyses indicate the critical role that normalization to reporter expression level plays in determining ncAA incorporation efficiency in different strains. Increases in both N-terminal and C-terminal signal levels of wild-type and TAG-containing reporters are evident in several knockout strains. Interestingly, this includes cases in which the presence of ncAA in the induction media increases the N-terminal signal level in cells containing a reporter with an amber codon (for example, UPF2 and PPQ1 knockouts with the Alt-TAG reporter). Trends in the alteration of reporter levels are different between dual-fluorescent protein reporters and the drop-in yeast display reporter. One potential explanation for this change is that only the display reporter constructs traverse the secretory pathway, where strong folding quality control mechanisms are present.[28, 51] Changes in reporter expression may help to explain why we did not observe an increase in ncAA incorporation efficiency in a UPF1 deletion strain. Because previous reports demonstrated increased ncAA-containing protein yield in a yeast strain containing this knockout, we initially hypothesized that this could be due to increased ncAA incorporation efficiency.[35, 36] However, our experiments suggest that knocking out UPF1 generally increases reporter protein expression levels over that of the parent strain (Table S10 – S12). Reporter expression levels (as detected by the N-terminal signal levels) in the UPF1 knockout vary by over 2-fold, and these variations are not fully accounted for by evaluating changes in C-terminal signal levels only (see Fig. S17, S24, S31). Thus, when evaluating ncAA incorporation efficiency (as opposed to ncAA-containing protein yield), our data here indicate the importance of both controlling for changes in reporter expression levels and performing normalization to determine whether increases in fluorescent signal were the result of enhanced ncAA incorporation efficiency. The use of a dual-detection reporter architecture in combination with the RRE analysis framework accounted for both of these sources of variability in ways that would have been challenging using a single-fluorescent protein reporter. Taken together, these data demonstrate the utility of both dual-fluorescent reporters and drop-in yeast display reporters in evaluating the effects of genomic modifications on ncAA incorporation events.

4 Conclusions

As applications of genetically encoded noncanonical amino acids continue to mature, characterization and improvement of ncAA incorporation systems is critical to ensuring their effective deployment. This is especially true when envisioning engineered translation systems (in cells or *in vitro*) that translate alternative genetic codes with the same efficiencies as wild-type translation apparatuses and the standard genetic code. However, even in applications where efficiency requirements for ncAA incorporation are more modest, quantitative characterizations can help determine critical information such as how ncAA incorporation with a given OTS alters expression levels of a protein of interest. The work presented here provides quantitative demonstrations of how the selection of reporter type, detection methodology, and data analyses all affect the determination of ncAA incorporation efficiency and fidelity in yeast. The tools described in this work will enable the combinatorial evaluation of the effects of several factors (e.g., OTSs and their expression levels, cell strains, conditions of cell growth and induction, and host genome composition) on ncAA incorporation events. These findings also

suggest the potentially high value of conducting similar investigations in other cells and organisms. In particular, a substantial portion of genetic code expansion work is conducted in either *E. coli* or mammalian cells; our studies could provide a blueprint for evaluating reporter type, detection methodology, and data analysis in precisely determining ncAA incorporation efficiency and fidelity in these types of cells. Understanding how reporters influence measurements of ncAA insertion events is an important step in facilitating cross-study comparisons of quantitative ncAA incorporation efficiency and fidelity. Single-fluorescent protein reporters, dual-fluorescent protein reporters, and yeast display reporters all exhibit similar levels of performance and precision when measured via flow cytometry. However, we identified cases when the use of single-fluorescent protein reporters could not identify changes in reporter expression levels, rather than perceived changes in ncAA incorporation effiency, suggesting that these reporters should be used with caution.

Excitingly, in this work we validated the use of dual-fluorescent reporters and a drop-in version of a yeast display reporter for evaluating the effects of gene knockouts on ncAA incorporation and found some gene knockouts that appear to enhance incorporation efficiency. We also found that even subtle details of reporter design and implementation alter the performance of ncAA incorporation systems. Understanding the effects of reporters with varying levels of stop codon suppression permissivity may be beneficial in future work to engineer cells for enhanced ncAA incorporation. As shown in Figure 5, the low level of ncAA incorporation efficiency supported by the Alt-TAG reporter supported ready identification of strains exhibiting enhanced readthrough efficiencies, whereas changes observed in efficiency were lower when using the parent BXG reporter. We expect that these different reporters will be useful in high throughput screening settings, enabling careful control over the stringency of screens. Until changes in stop codon suppression efficiencies between distinct reporter systems become predictable *a priori*, cross-study comparisons of ncAA incorporation system performance may prove to be challenging. One straightforward way to begin to facilitate such comparisons would be to perform reference characterizations with known OTSs and multiple reporter systems so that quantitative properties of new OTSs relative to a known OTS could be determined more readily. This also applies to existing OTSs in new cell strains, existing OTSs used under the control of new promoter systems or expression conditions, or additional modifications. “Benchmarking” in this way would eliminate the need to immediately standardize reporter systems, although such standardization could be valuable in the longer term. Future validation of the reporter systems used in this study with highly sensitive mass spectrometry-based methods[9, 13] could provide additional insights into developing reliable metrics for quantifying ncAA incorporation. In any case, fully illuminating the current landscape of ncAA incorporation systems available in different organisms would enable identification of effective combinations of engineering strategies for enhancing alternative translations of the genetic code. As we enter an era in which genomes can be rapidly constructed,[27, 52-55] evolved,[56-58] or expanded to include additional nucleic acid bases,[19] quantitative reporter systems with known properties will be indispensable for pushing the generation of polypeptides with chemically diverse building blocks to new heights.

Conflicts of Interest

The authors declare no competing financial interests.

Acknowledgements

Research reported in this publication was supported by the National Institute Of General Medical Sciences of the National Institutes of Health under Award Number R35GM133471, the Army Research Office under Award Number W911NF-16-1-0175, and Tufts startup funds (to J.A.V.). J.T.S. was supported in part by an NSF Graduate Research Fellowship (ID: 2016231237). Ming Lei was supported by funds from a 2018 Tufts Collaborates grant, PI Yoon H. Kang. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health, the Army Research Office, or Tufts University.

References

1. J. W. Chin, *Nature*, 2017, **550**, 53-60.
2. A. Rezhdo, M. Islam, M. Huang and J. A. Van Deventer, *Curr Opin Biotechnol*, 2019, **60**, 168-178.
3. M. P. Ledbetter and F. E. Romesberg, *Curr Opin Chem Biol*, 2018, **46**, A1-A2.
4. L. Wang, *Acc Chem Res*, 2017, **50**, 2767-2775.
5. K. Y. Fang, S. A. Lieblich and D. A. Tirrell, *Methods Mol Biol*, 2018, **1798**, 173-186.
6. J. Wang, Y. Liu, Y. Liu, S. Zheng, X. Wang, J. Zhao, F. Yang, G. Zhang, C. Wang and P. R. Chen, *Nature*, 2019, **569**, 509-513.
7. J. T. Ngo and D. A. Tirrell, *Acc Chem Res*, 2011, **44**, 677-685.
8. K. W. Barber, P. Muir, R. V. Szeligowski, S. Rogulina, M. Gerstein, J. R. Sampson, F. J. Isaacs and J. Rinehart, *Nat Biotechnol*, 2018, **36**, 638-644.
9. M. Amiram, A. D. Haimovich, C. Fan, Y. S. Wang, H. R. Aerni, I. Ntai, D. W. Moonan, N. J. Ma, A. J. Rovner, S. H. Hong, N. L. Kelleher, A. L. Goodman, M. C. Jewett, D. Soll, J. Rinehart and F. J. Isaacs, *Nat Biotechnol*, 2015, **33**, 1272-1279.
10. Y. Huang, M. M. Wiedmann and H. Suga, *Chem Rev*, 2018, DOI: 10.1021/acs.chemrev.8b00430.
11. T. Passioura, W. Liu, D. Dunkelmann, T. Higuchi and H. Suga, *J Am Chem Soc*, 2018, **140**, 11551-11555.
12. S. A. Lieblich, K. Y. Fang, J. K. B. Cahn, J. Rawson, J. LeBon, H. T. Ku and D. A. Tirrell, *J Am Chem Soc*, 2017, **139**, 8384-8387.
13. K. Mohler, H. R. Aerni, B. Gassaway, J. Ling, M. Ibba and J. Rinehart, *Biochim Biophys Acta Gen Subj*, 2017, **1861**, 3081-3088.
14. T. S. Young, I. Ahmad, J. A. Yin and P. G. Schultz, *J Mol Biol*, 2010, **395**, 361-374.
15. B. Wiltschi, *Fungal Genet Biol*, 2016, **89**, 137-156.
16. L. Wang, A. Brock, B. Herberich and P. G. Schultz, *Science*, 2001, **292**, 498-500.
17. J. W. Chin, T. A. Cropp, J. C. Anderson, M. Mukherji, Z. Zhang and P. G. Schultz, *Science*, 2003, **301**, 964-967.
18. J. C. Anderson, N. Wu, S. W. Santoro, V. Lakshman, D. S. King and P. G. Schultz, *Proc Natl Acad Sci U S A*, 2004, **101**, 7566-7571.
19. Y. Zhang, J. L. Ptacin, E. C. Fischer, H. R. Aerni, C. E. Caffaro, K. San Jose, A. W. Feldman, C. R. Turner and F. E. Romesberg, *Nature*, 2017, **551**, 644-647.
20. P. O'Donoghue, J. Ling, Y. S. Wang and D. Soll, *Nat Chem Biol*, 2013, **9**, 594-598.

21. O. Vargas-Rodriguez, A. Sevostyanova, D. Soll and A. Crnkovic, *Curr Opin Chem Biol*, 2018, **46**, 115-122.
22. K. Wang, A. Sachdeva, D. J. Cox, N. M. Wilf, K. Lang, S. Wallace, R. A. Mehl and J. W. Chin, *Nat Chem*, 2014, **6**, 393-403.
23. R. Gan, J. G. Perez, E. D. Carlson, I. Ntai, F. J. Isaacs, N. L. Kelleher and M. C. Jewett, *Biotechnol Bioeng*, 2017, **114**, 1074-1086.
24. H. Neumann, K. Wang, L. Davis, M. Garcia-Alai and J. W. Chin, *Nature*, 2010, **464**, 441-444.
25. Y. Doi, T. Ohtsuki, Y. Shimizu, T. Ueda and M. Sisido, *J Am Chem Soc*, 2007, **129**, 14458-14462.
26. D. B. Johnson, C. Wang, J. Xu, M. D. Schultz, R. J. Schmitz, J. R. Ecker and L. Wang, *ACS Chem Biol*, 2012, **7**, 1337-1344.
27. M. J. Lajoie, A. J. Rovner, D. B. Goodman, H. R. Aerni, A. D. Haimovich, G. Kuznetsov, J. A. Mercer, H. H. Wang, P. A. Carr, J. A. Mosberg, N. Rohland, P. G. Schultz, J. M. Jacobson, J. Rinehart, G. M. Church and F. J. Isaacs, *Science*, 2013, **342**, 357-360.
28. J. T. Stieglitz, H. P. Kehoe, M. Lei and J. A. Van Deventer, *ACS Synth Biol*, 2018, **7**, 2256-2269.
29. J. W. Monk, S. P. Leonard, C. W. Brown, M. J. Hammerling, C. Mortensen, A. E. Gutierrez, N. Y. Shin, E. Watkins, D. M. Mishler and J. E. Barrick, *ACS Synth Biol*, 2017, **6**, 45-54.
30. J. A. Van Deventer, D. N. Le, J. Zhao, H. P. Kehoe and R. L. Kelly, *Protein Eng Des Sel*, 2016, **29**, 485-494.
31. I. C. Tanrikulu, E. Schmitt, Y. Mechulam, W. A. Goddard, 3rd and D. A. Tirrell, *Proc Natl Acad Sci U S A*, 2009, **106**, 15285-15290.
32. H. S. Kwok, O. Vargas-Rodriguez, S. V. Melnikov and D. Soll, *ACS Chem Biol*, 2019, **14**, 603-612.
33. A. E. Owens, K. T. Grasso, C. A. Ziegler and R. Fasan, *Chembiochem*, 2017, **18**, 1109-1116.
34. L. Wang and P. G. Schultz, *Angew Chem Int Ed Engl*, 2004, **44**, 34-66.
35. Q. Wang and L. Wang, *J Am Chem Soc*, 2008, **130**, 6066-6067.
36. H. W. Ai, W. Shen, E. Brustad and P. G. Schultz, *Angew Chem Int Ed Engl*, 2010, **49**, 935-937.
37. J. A. Van Deventer, R. L. Kelly, S. Rajan, K. D. Wittrup and S. S. Sidhu, *Protein Eng Des Sel*, 2015, **28**, 317-325.
38. J. W. Chin, T. A. Cropp, S. Chu, E. Meggers and P. G. Schultz, *Chem Biol*, 2003, **10**, 511-519.
39. N. Wu, A. Deiters, T. A. Cropp, D. King and P. G. Schultz, *J Am Chem Soc*, 2004, **126**, 14306-14307.
40. D. A. Drummond and C. O. Wilke, *Nat Rev Genet*, 2009, **10**, 715-724.
41. R. E. Campbell, O. Tour, A. E. Palmer, P. A. Steinbach, G. S. Baird, D. A. Zacharias and R. Y. Tsien, *Proc Natl Acad Sci U S A*, 2002, **99**, 7877-7882.
42. O. M. Subach, P. J. Cranfill, M. W. Davidson and V. V. Verkhusha, *PLoS One*, 2011, **6**, e28674.
43. J. D. Pedelacq, S. Cabantous, T. Tran, T. C. Terwilliger and G. S. Waldo, *Nat Biotechnol*, 2006, **24**, 79-88.
44. S. I. Feinstein and S. Altman, *Genetics*, 1978, **88**, 201-219.

45. M. K. Phillips-Jones, F. J. Watson and R. Martin, *J Mol Biol*, 1993, **133**, 1-6.
46. M. K. Phillips-Jones, L. S. J. Hill, J. Atkinson and R. Martin, *Mol Cell Biol*, 1995, **15**, 6593-6600.
47. O. Namy, I. Hatin and J. P. Rousset, *EMBO Rep*, 2001, **2**, 787-798.
48. D. G. Schwark, M. A. Schmitt and J. D. Fisk, *Genes (Basel)*, 2018, **9**.
49. M. Pott, M. J. Schmidt and D. Summerer, *ACS Chem Biol*, 2014, **9**, 2815-2822.
50. T. von der Haar and M. F. Tuite, *Trends Microbiol*, 2007, **15**, 78-86.
51. E. V. Shusta, M. C. Kieke, E. Parke, D. M. Kranz and K. D. Wittrup, *Journal of Molecular Biology*, 1999, **292**, 949-956.
52. J. Fredens, K. Wang, D. de la Torre, L. F. H. Funke, W. E. Robertson, Y. Christova, T. Chia, W. H. Schmied, D. L. Dunkelmann, V. Beranek, C. Uttamapinant, A. G. Llamazares, T. S. Elliott and J. W. Chin, *Nature*, 2019, **569**, 514-518.
53. S. M. Richardson, L. A. Mitchell, G. Stracquadanio, K. Yang, J. S. Dymond, J. E. DiCarlo, D. Lee, C. L. Huang, S. Chandrasegaran, Y. Cai, J. D. Boeke and J. S. Bader, *Science*, 2017, **355**, 1040-1044.
54. Y. H. Lau, F. Stirling, J. Kuo, M. A. P. Karrenbelt, Y. A. Chan, A. Riesselman, C. A. Horton, E. Schafer, D. Lips, M. T. Weinstock, D. G. Gibson, J. C. Way and P. A. Silver, *Nucleic Acids Res*, 2017, **45**, 6971-6980.
55. C. A. Hutchison, 3rd, R. Y. Chuang, V. N. Noskov, N. Assad-Garcia, T. J. Deerinck, M. H. Ellisman, J. Gill, K. Kannan, B. J. Karas, L. Ma, J. F. Pelletier, Z. Q. Qi, R. A. Richter, E. A. Strychalski, L. Sun, Y. Suzuki, B. Tsvetanova, K. S. Wise, H. O. Smith, J. I. Glass, C. Merryman, D. G. Gibson and J. C. Venter, *Science*, 2016, **351**, aad6253.
56. A. J. Simon, S. d'Oelsnitz and A. D. Ellington, *Nat Biotechnol*, 2019, **37**, 730-743.
57. J. Steensels, A. Gorkovskiy and K. J. Verstrepen, *Nat Commun*, 2018, **9**, 1937.
58. Y. Shen, G. Stracquadanio, Y. Wang, K. Yang, L. A. Mitchell, Y. Xue, Y. Cai, T. Chen, J. S. Dymond, K. Kang, J. Gong, X. Zeng, Y. Zhang, Y. Li, Q. Feng, X. Xu, J. Wang, J. Wang, H. Yang, J. D. Boeke and J. S. Bader, *Genome Res*, 2016, **26**, 36-49.

Figures and Captions

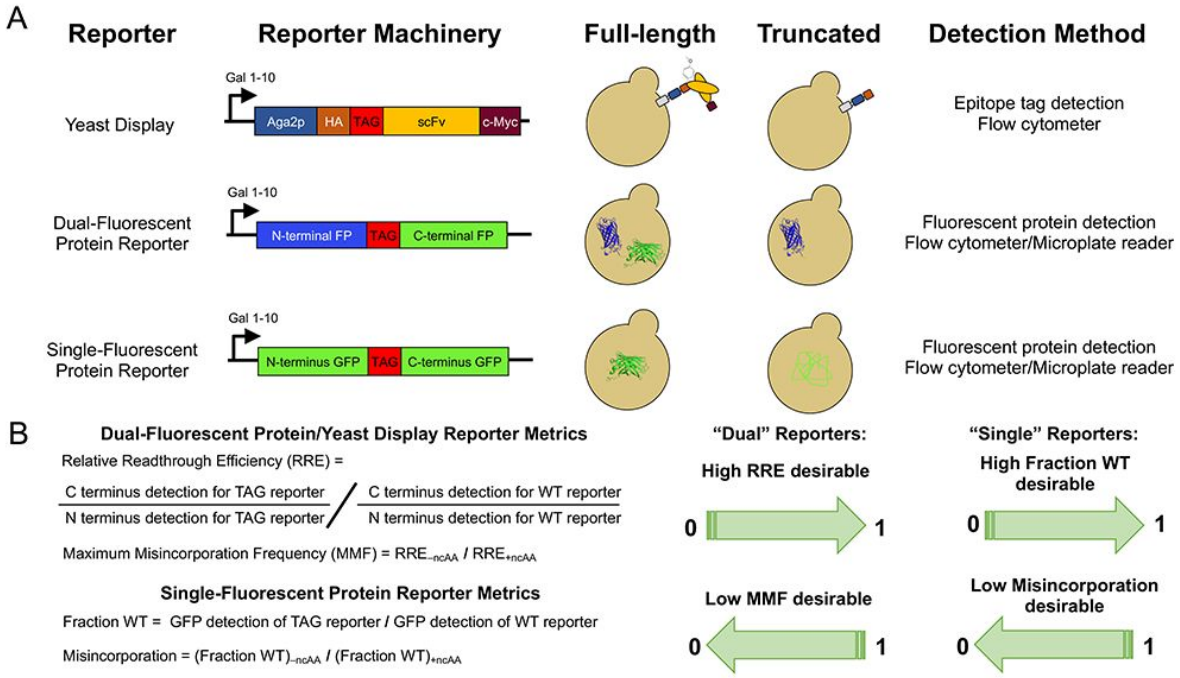

Fig. 1. Reporter architectures, detection methods, and analysis methods. (A) Architectures of the three major types of reporters used in this work, expected behaviors in yeast, and detection methods for quantifying ncAA incorporation events. (B) Metrics used to determine ncAA incorporation efficiency and fidelity based on N- and C-terminal detection in "dual" reporters (yeast display, dual-fluorescent protein reporter) and full-length protein detection in "single" reporters (single-fluorescent protein reporter).

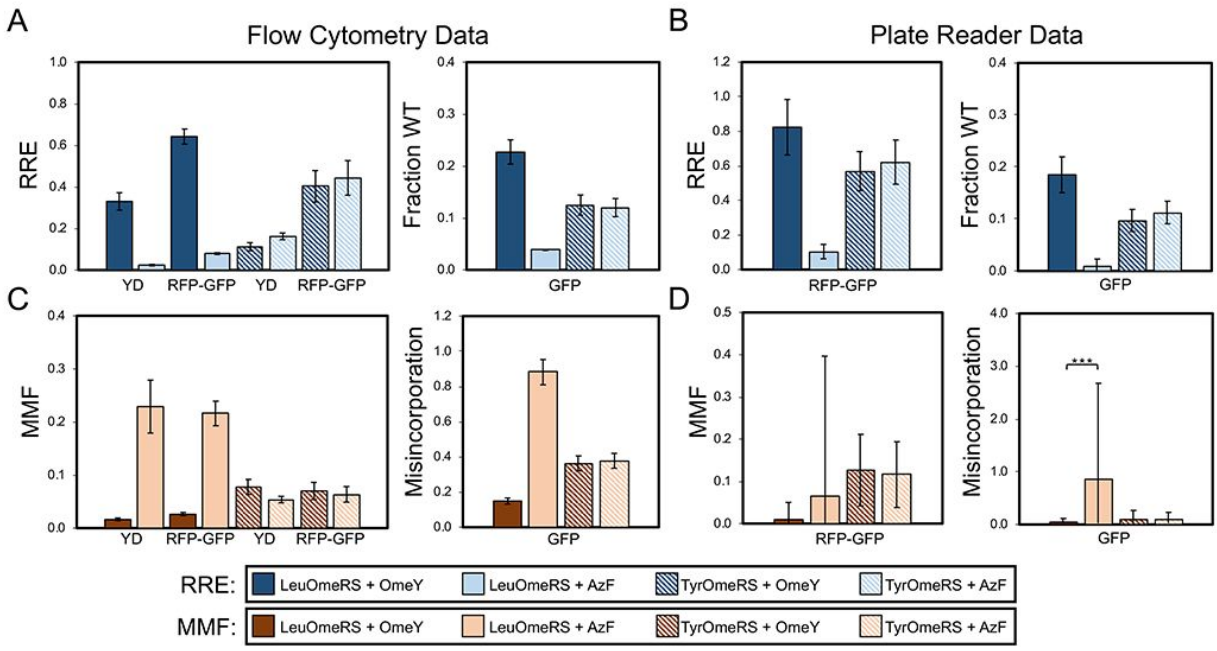

Fig. 2. Quantification of ncAA incorporation efficiency and fidelity for four orthogonal translation system-ncAA combinations. (A) Flow cytometry data quantifying relative readthrough efficiency (RRE) for "dual" reporter systems (yeast display (YD), RFP-GFP dual-fluorescent protein reporter) and fraction of wild-type reporter (Fraction WT) for sfGFP single-fluorescent protein reporter (see Materials and Methods for descriptions of calculations). (B) Microplate reader data quantifying RRE for RFP-GFP reporter and Fraction WT for sfGFP reporter. (C) Flow cytometry data quantifying maximum misincorporation frequency (MMF) or misincorporation for RRE and Fraction WT measurements, respectively, reported in (A). (D) Microplate reader data quantifying MMF or misincorporation for RRE and Fraction WT measurements, respectively, reported in (B). All conditions were evaluated using end point measurements in biological triplicate. Error bars represent the propagated error from performing RRE and MMF calculations. Statistical evaluation of the N- and C-terminal data corresponding to the data shown in this figure can be found in Table S4 – S5.

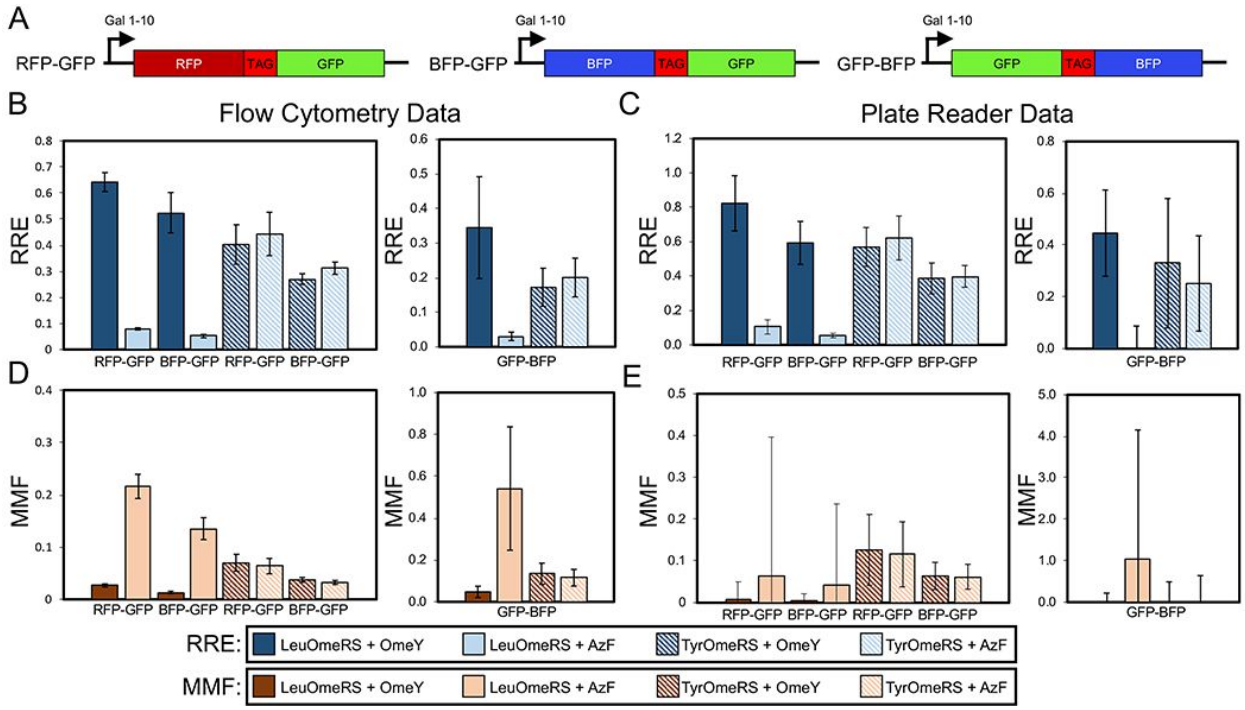

Fig. 3. Evaluation of dual-fluorescent protein reporter performance as a function of fluorescent protein identity and orientation. (A) Architectures of RFP-GFP, BFP-GFP, and GFP-BFP reporters. The reporters contain identical linkers. (B) Relative readthrough efficiency determined via flow cytometry measurements with 4 OTS-ncAA combinations (see Materials and Methods for details on calculations). Data for RFP-GFP and BFP-GFP were collected on an LSR-II flow cytometer, and data for GFP-BFP were collected with an Attune NxT flow cytometer (Fig. S1 indicates that data collected on these two instruments are equivalent). Note that the RFP-GFP data shown in Fig. 2 are being shown again to enable a direct comparison with the other dual-fluorescent protein constructs evaluated here. (C) RRE determined via microplate reader measurements for the same series of OTS-ncAA combinations as in (B). (D) Maximum misincorporation frequency determined via flow cytometry for RRE measurements reported in (B). (E) MMF determined via microplate reader for RRE measurements reported in (C). All conditions were evaluated using end point measurements in biological triplicate. Error bars represent the propagated error from performing RRE and MMF calculations. Statistical evaluation of the N- and C-terminal data corresponding to the data shown in this figure can be found in Table S6 – S8.

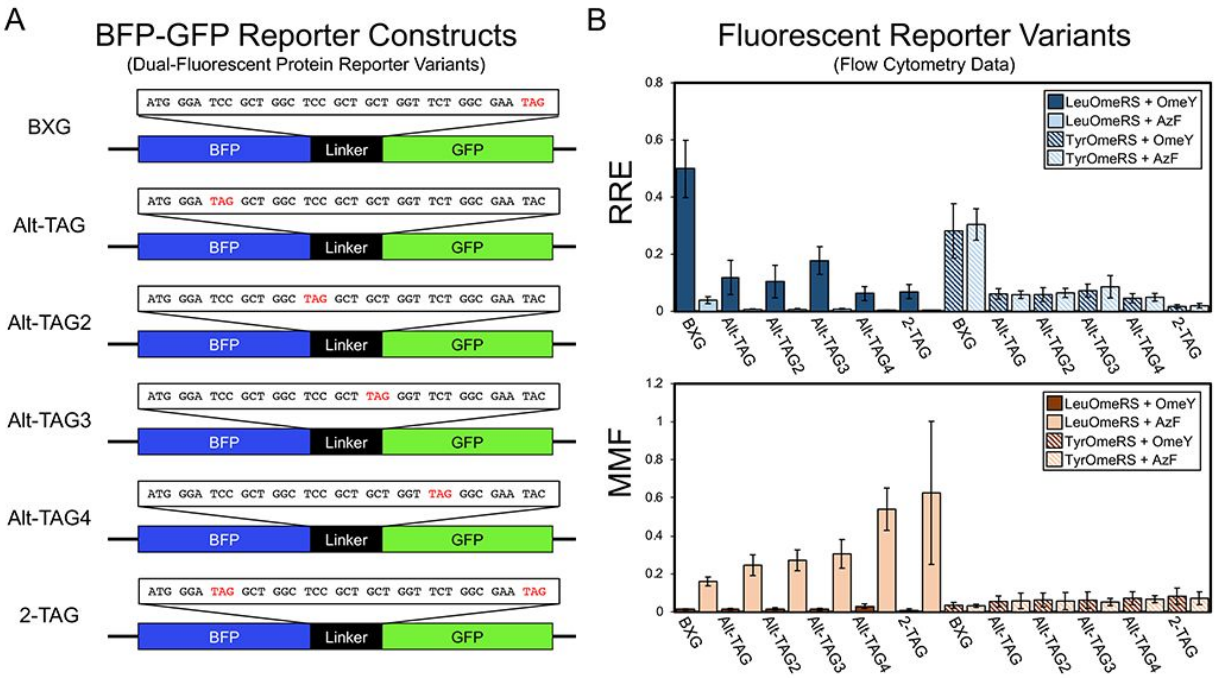

Fig. 4. Evaluation of dual-fluorescent protein reporter performance as a function of stop codon position within linker. (A) Structures of BFP-GFP linker variants. (B) Relative readthrough efficiency and maximum misincorporation frequency determined via flow cytometry measurements with a series of four OTS-ncAA combinations. All conditions were evaluated using end point measurements in biological triplicate. Error bars represent the propagated error from performing RRE and MMF calculations. Statistical evaluation of the N- and C-terminal data corresponding to the data shown in this figure can be found in Table S9.

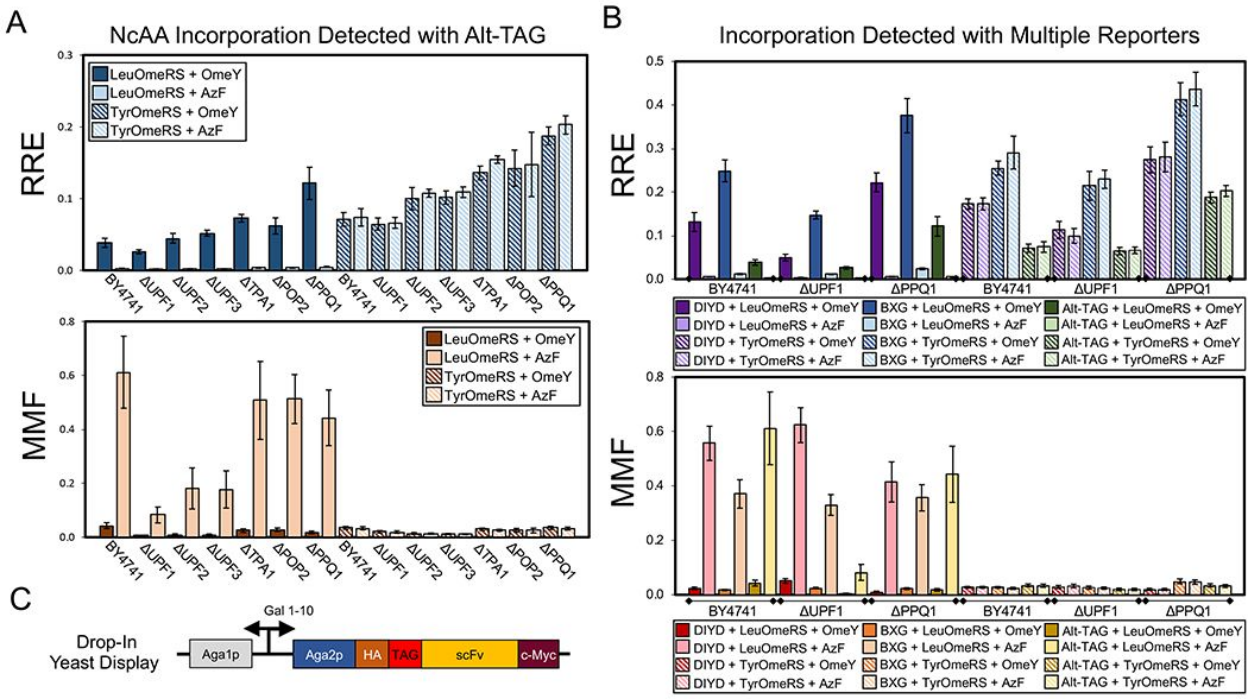

Fig. 5. Evaluation of ncAA incorporation events in a series of single-gene knockout strains using two dual-fluorescent protein reporters and a drop-in yeast display (DIYD) reporter. (A) Measurements of ncAA incorporation efficiency and fidelity in BY4741 and six single-gene knockouts of BY4741 using the Alt-TAG BFP-GFP reporter in a pRS416 (URA3 marker) plasmid backbone (see Materials and Methods for calculation details). (B) Selected measurements of ncAA incorporation efficiency and fidelity using the Alt-TAG BFP-GFP reporter, BXG BFP-GFP reporter, and drop-in yeast display reporter in pRS416 (URA3 marker) plasmid backbones. The complete set of measurements using all three reporters for each of the six knockout strains strains is available in Fig. S13. (C) Architecture of drop-in yeast display system. A quantitative comparison of drop-in and conventional yeast display is available in Fig. S12. All conditions were evaluated using end point measurements in biological triplicate. Error bars represent the propagated error from performing RRE and MMF calculations. Statistical evaluation of the N- and C-terminal data corresponding to the data shown in this figure can be found in Table S10 – S12.