Dual stop codon suppression in mammalian cells with genomically integrated genetic code expansion machinery

Birthe Meineke,1,2* Johannes Heimgärtner,1,2 Rozina Caridha,1,2 Matthias Block,1 Kyle J Kimler,1 Maria F Pires,1 Michael Landreh3 and Simon J Elsässer1,2,*

1Science for Life Laboratory, Karolinska Institutet, Department of Medical Biochemistry and Biophysics, Division of Genome Biology, Stockholm 17165, Sweden
2Ming Wai Lau Centre for Reparative Medicine, Stockholm node, Karolinska Institutet, Stockholm 17165, Sweden
3Department of Microbiology, Tumor and Cell Biology, Science for Life Laboratory, Karolinska Institutet, Stockholm 17165, Sweden
*Correspondence: birthe.meineke@scilifelab.se; simon.elsasser@scilifelab.se

ABSTRACT
Genetic code expansion via stop codon suppression is a powerful strategy to engineer proteins. Suppressor tRNAs are aminoacylated with noncanonical amino acids (ncAAs) by dedicated aminoacyl-tRNA synthetases (aaRS) and direct ncAA incorporation site-specifically during translation. These pairs of tRNA/aaRS must be orthogonal to the host’s tRNAs, aaRS and natural amino acids. Pyrrolysyl-tRNA (PylT)/PylRS pairs from methanogenic archaea, as well as engineered tRNA/aaRS pairs derived from bacteria, are used for genetic code expansion in mammalian cells. Amber suppression is routinely achieved by transient introduction of the components leading to short-term and heterogeneous expression. Here, we demonstrate that stable integration of tRNA/aaRS genes allows for efficient, genetically encoded ncAA incorporation in diverse mammalian cell lines. We extend a general plasmid design and PiggyBac (PB) integration strategy developed for the Methanosarcina mazei PylT/PylRS pair to genomic integration of two tRNA/aaRS pairs of bacterial origin. We further explore suppression of ochre and opal stop codons and parallel incorporation of two distinct ncAAs, both accessible for click chemistry, by dual suppression in stable cell lines. Clonal selection allows for isolation of cells with high dual suppression efficiency and dual site-specific fluorescent labeling of a cell surface receptor using bioorthogonal click chemistries on live mammalian cells.
INTRODUCTION
The genetic code assigns all possible triplet codons to one of the canonical amino acids (sense codons) or to a translation stop signal (nonsense codons). Expansion of the genetic code, that is genetic encoding of noncanonical amino acids (ncAAs), requires reprogramming of a codon. Most commonly the amber (TAG) stop codon is repurposed, using an engineered tRNA/aminocyl-tRNA synthetase (aaRS) pair to deliver an ncAA-charged tRNA to the ribosome. The amber codon is the least abundant (23% in the human genome) and most efficiently suppressed stop codon, but the other two stop codons UUA (ochre) and UGA (opal) have also been exploited for ncAA incorporation in mammalian cells and enabled dual ncAA incorporation via dual suppression (Xiao et al. 2013; Zheng et al. 2017) (Meineke et al. 2018, 2020; Serfling et al. 2018; Beránek et al. 2019; Osgood et al. 2022). Quadruplet codon suppression has been proven possible in worms and mammalian cells as well (Niu et al. 2013; Xi et al. 2022). The pyrrolysyl-tRNA (PylT) and PylT synthetase (PylRS) pair from methanogenic archaea, is widely used for amber suppression in mammalian cells. In the Methanosarcinaceae family of archaea, specific amber codons in mRNAs coding for enzymes of the methanogenic pathway are suppressed by PylT to encode pyrrolysine (Pyl) (Rother and Krzycki 2010). PylT and PylRS have characteristics that make them ideal for genetic code expansion in model organisms: the conserved fold of PylT allows specific interaction with PylRS independent of anticodon identity (Srinivasan et al. 2002) (Ambrogelly et al. 2007) and the relatively promiscuous amino acid binding pocket of PylRS readily accommodates structurally diverse ncAAs. This broad substrate specificity has been further increased by protein engineering, over 100 ncAAs can be incorporated by PylRS variants (reviewed in (Wan et al. 2014)). Importantly, the PylT/RS pair is orthogonal to prokaryotic and eukaryotic tRNAs or aaRS (Suzuki et al. 2017). These features have made the pairs from Methanosarcina mazei (Mma) and Methanosarcina bakeri (Mba) widely used tools for site-specific integration of a plethora of noncanonical amino acids (ncAAs) in bacterial, eukaryotic, mammalian host systems and animals (Chin 2014).

Another approach to expand the genetic code by amber suppression relies on engineering of canonical tRNA/aaRS pairs from one organism in a host of a different clade of life. This has been achieved for both tyrosyl-tRNA synthetase (TyrRS) and leucyl-tRNA synthetase (LeuRS) from E. coli by changing the anticodon of the cognate tRNA (TyrT and LeuT, respectively) to CUA and sequential rounds of selection for orthogonality and efficient amber suppression. An
engineered EcoTyrRS mutant, AzFRS, has been shown to effectively incorporate p-azido-phenylalanine (AzF) with either EcoTyrT^CUA or B. stearothermophilus TyrT^CUA (BstTyr^CUA) (Sakamoto et al. 2002) (Chin et al. 2003) (Liu et al. 2007). EcoLeuRS-derived Anap-2C (AnapRS) paired with EcoLeuT^CUA, allows introduction of the minimal fluorescent ncAA 3-(6-acetylnapthalene-2-ylamino)-2-amino-propanoic acid (Anap) in mammalian cells (Lee et al. 2009; Chatterjee et al. 2013).

In mammalian cell culture, components of the amber suppression machinery, i.e. the tRNA/aaRS pair and reporter or gene of interest with amber codon, are commonly delivered by transient transfection. Lipofection methods allow easy delivery of high copy numbers of the expression vectors into the cell, but the resulting populations are heterogeneous in expression levels and endogenous transcripts are outcompeted by the transgene (Muerdter et al. 2018). For prolonged and homogenous protein production, extended expression experiments, selection of homologous expression within a cell population and other applications, amber suppression cell lines with all components of the genetic code expansion machinery integrated in their genome, are desirable. We have used PiggyBac (PB) transgenesis to achieve genetic code expansion from stable integrants previously (Elsässer et al. 2016; Elsässer 2018; van Husen et al. 2021; Lee et al. 2009). PB transposition utilizes a deconstructed transposon cassette: cotransfection of a PB transposase (PBT) expression plasmid with a second plasmid containing an integration cassette flanked by inverted repeats recognised by PBT leads to random genomic integration of the cassette at TTAA sites in the host genome. Integrants are selected by antibiotic resistance markers included in the inserted cassette. While PB transgenesis is known to allow multiple integration events per cell (Ding et al. 2005), expression of the protein of interest is expected to be much lower and closer to high-expressing endogenous genes as compared to a transient transfection experiment. Therefore, stable genetic code expansion is commonly viewed as inefficient. Lower, genomic levels of the gene of interest may increase the competition with off-target suppression at endogenous amber codons, raising the question how selectivity towards the transgene’s stop codon can be achieved in an endogenous expression context (Bartoschek et al. 2021). It is thus important to comprehensively characterize stable suppression systems for understanding the opportunities and limitations of transgenic cell lines. In the present study, we aimed to create a panel of representative cell lines and systematically characterize efficiency and selectivity of amber suppression using PB transgenesis. We demonstrate that stable genetic code expansion greatly varies in efficiency by cell line, but can
rival transient expression in specific settings and enable efficient single and dual stop codon suppression.

**RESULTS**

**Integrating amber suppression machinery into the genome of human cell lines by PB transgenesis**

We have previously developed a plasmid system for generation of stable amber suppression cell lines (pAS plasmids) using *Mma* PyIT/RS (PyIT/RS hereafter) (Figure 1A) (Elsässer et al. 2016; Elsässer 2018): One plasmid for PyIRS expression and a second encoding a sfGFP$^{150\text{TAG}}$ (GFP$^{150\text{TAG}}$) reporter gene, each controlled by the EF1α promoter. Both plasmids carry four tandem repeats of 7SK PolIII-controlled PyIT genes. A resistance marker is expressed from an internal ribosome entry site (IRES) in the PyIS or GFP transcripts to allow selection of integrants. The expression cassette is flanked on both ends by insulator sequences and inverted repeats for PB transposition. This plasmid design allowed for highest suppression efficiency in transient expression compared to other systems (Zhou et al. 2020) and we have successfully employed pAS plasmids for both transient transfection and stable integration in the past (Meineke et al. 2018, 2020; Lafranchi et al. 2020; van Husen et al. 2021; Elsässer et al. 2016).

Common immortalized animal cell lines differ in their characteristics with regard to transgene expression and ease of transfection. We compare PB-mediated integration of our amber suppression system, consisting of PyIT, PyIS and the GFP$^{150\text{TAG}}$ reporter, across a panel of commonly used cell lines. The two plasmids were co-transfected with the PBT plasmid into human embryonic kidney derived cell lines HEK293 and HEK293T, human colon cancer cell line HCT116, human melanoma A375 and osteosarcoma cell line U-2 OS as well as primate cell line COS-7. The resulting stable polyclonal pools were grown in the presence or absence of N-ε-[2-(2-methyl-2-cyclopropene-1-yl)-methoxy] carbonyl-L-lysine (CpK) for 24 h. GFP was only expressed in presence of CpK as evident from anti-GFP immunoblotting (Figure 1B, even numbered lanes). We used strain-promoted inverse electron-demand Diels-Alder cycloaddition (SPIEDAC) between CpK and tetrazine-silicon rhodamine (tetSiR) in lysates to visualize CpK incorporation (Elliott et al. 2014). A single band corresponding to GFP$^{150\text{CpK-tetSiR}}$ showed selective ncAA incorporation in all cell lines generated (Figure 1B). Faint additional bands indicated low level incorporation of ncAA into various endogenous proteins. The amount of GFP$^{150\text{CpK}}$ produced greatly differed between cell lines. HCT116-derived cells produced the
strongest GFP signal, followed by HEK293T and HEK293 cells. Curiously, PyIRS protein levels were increased or decreased upon ncAA addition in some cell lines and PyIRS expression levels did not correlate with production of full-length GFP. We did not further investigate this phenomenon. While lysates were normalized by total protein content before gel separation, GAPDH levels varied, presumably because of cell type-dependent expression patterns.

**PyIT gene copy number is limiting in HEK293T stable amber suppression cell lines**

To determine which components are limiting amber suppression, we transfected a relatively inefficient HEK293T amber suppression cell line with different pAS constructs to increase gene copy number of PyIT, PyIRS or both. A control pAS plasmid did not change GFP\(^{150\text{CpK}}\) or PyIRS expression (lanes 1-3, Figure 1C, D). Transfection with pAS expressing PyIRS lead to strongly increased FLAG-PyIRS protein level, but no increased amber suppression efficiency, as GFP signal did not increase (lane 4, Figure 1C, D). Addition of pAS with either only 4xPyIT or 4xPyIT/PyIRS increased GFP\(^{150\text{CpK}}\) expression (lanes 5 and 6, Figure 1C, D). The GFP fluorescence increased 2.5-fold with transient transfection of pAS 4xPyIT and tripled with pAS with 4xPyIT/PyIRS. Amber suppression did not improve by further increasing the PyIT gene dosage with pAS 8xPyIT, or by using reported optimized PyIT variants M15 or A2-1 (Serfling et al. 2018; Jewel et al. 2023). Like wild-type PyIRS, chimeric *Mba/Mma*PyIRS IPYE mutant had no effect on amber suppression, despite its reported higher catalytic rate (Bryson et al. 2017) (Figure S1A). This indicates that PyIT levels can limit amber suppression when they are exclusively expressed from randomly integrated pAS transgenes. The comparably high levels of amber suppression achieved in transient transfection are therefore linked to high PyIT gene copy number, while PyIRS expression and activity are not rate-limiting.

**Comparison of transient and stable amber suppression**

We further compared amber suppression efficiency of stable polyclonal HEK293, HEK293T and HCT116 cell lines with their transient transfected counterparts. After 24 h incubation with CpK, transiently transfected HEK293T cells express GFP\(^{150\text{CpK}}\) very heterogeneously: flow cytometry shows 60% GFP-positive cells, but the level of GFP fluorescence varies greatly within that population. Stable HEK293T cells produced a less extreme but more homogeneous level of GFP reporter fluorescence between 10\(^2\)-10\(^3\) arbitrary units (Figure 1E). While HEK293 and HCT116 cells showed much reduced transfection efficiency, stable cell lines exhibit efficient suppression across the entire population. Comparing mean GFP fluorescence per cell
(indicative of the total yield of GFP), stable HCT116 cells produced almost as much GFP$^{150CpK}$ as transiently transfected HEK293T cells, suggesting that transgenic HCT116 are an excellent expression hosts for producing ncAA-containing proteins in mammalian cells (Figure S1B). We further confirmed this finding using a different ncAA and PylRS variant. The chimeric Mma/Mba PylRS variant AbKRS-chlPYE (Bryson et al. 2017; Ai et al. 2011; Gautier et al. 2010), engineered for incorporation of the photocrosslinking ncAA 3’-azibutyl-N-carbamoyl-lysine (AbK), also produced an efficiently amber-suppressing polyclonal population of HCT116 cells (Figure S1C, D).

**PB transgenesis with non-PylT/RS amber suppression pairs in HCT116 cells**

We generated pAS constructs for AzF and Anap incorporation analogous to our PylT/RS two-plasmid system (Figure S2A). AzFRS was combined with BstTyr$^{CUA}$ (Tyr$^{CUA}$/AzFRS) and AnapRS was combined with EcoLeu$^{CUA}$ (Leu$^{CUA}$/AnapRS), both pAS were designed with a geneticin (neomycin) selectable marker. The cognate tRNA/GFP$^{150TAG}$ reporters were also generated.

We compared amber suppression efficiency of transient transfection and PB transgenesis for AzFRS-mediated amber suppression in HCT116 cells, measuring GFP fluorescence by flow cytometry after 24 h with AzF. Transient transfection of HCT116 cells was again inefficient. Selection with geneticin and blasticidin for PB mediated integration of Tyr$^{CUA}$/AzFRS and Tyr$^{CUA}$/GFP$^{150TAG}$ produced a heterogeneous pool, but the fraction of GFP positive cells in presence of AzF increased from 4.7% in transient transfection to 39.2% (Figure 2A). Fluorescence-assisted cell sorting (FACS) of the polyclonal pool for single, strongly GFP-expressing cells, isolated clonal populations with very high amber suppression efficiency (Figure S2B). Evidently, the Tyr$^{CUA}$/AzFRS pair can support extremely efficient amber suppression from integrated transgenes.

We similarly compared the polyclonal pool of stable Leu$^{CUA}$/AnapRS and Leu$^{CUA}$/GFP$^{150CpK}$ integrants with transiently transfected cells. Here, we assessed both GFP and Anap fluorescence by flow cytometry (Figure 2B). The polyclonal pool of stable cell lines showed strong GFP and Anap fluorescence when incubated with the ncAA for 24 h, indicating expression of GFP$^{150Anap}$ (Figure 2B, S2C). A fluorogenic ncAA like Anap ideally allows direct detection of the protein of interest, however unincorporated and misincorporated ncAA can contribute substantially to signal intensity. To assess the contribution of misincorporated Anap in endogenous proteins, we generated a HCT116 cell line with integrated Leu$^{CUA}$/AnapRS without
amber suppression reporter. When cultured in presence of Anap for 24 h this AnapRS-only integrant cell line produced no GFP fluorescence, but Anap fluorescence was clearly detectable by flow cytometry despite extended ncAA-washout before measurement (Figure S2C-E). Evidently, Anap was retained in the cell in aminoacyl-tRNA complexes or incorporated at off-target amber codons even in the absence of a dedicated target protein. Total GFP yield as determined by western blot was marked higher in stable AzFRS and AnapRS amber suppression HCT116 cell lines (Figure 2C). We confirmed AzFRS expression via a C-terminal FLAG tag after stable integration, while AnapRS, expressed in its originally described tagless form ((Lee et al. 2009; Chatterjee et al. 2013)), could not be probed.

Efficient amber-, ochre- and opal-suppressing clonal populations generated by PB transgenesis in HCT116 cells

The amber codon is most widely used in genetic code expansion in mammalian cells, because it is the least abundant stop codon and amber suppression efficiency is superior to suppression of either of the other two stop codons, ochre (UAA) or opal (UAG) in transient transfections (Zheng et al. 2017; Meineke et al. 2018, 2020). We sought to explore if genomic integration would allow for selection and isolation of efficient ochre and opal suppressor populations, a prerequisite for incorporating more than one ncAA site-specifically in response to two different stop codons in a stable dual suppression cell line. We generated 4x7SK PylT cassettes for expression of ochre and opal suppressor PylTs, PylT_UAA and PylT_UCA, and combined them with PylRS and their cognate GFP reporter GFP^{150TA} or GFP^{150TG}, respectively. In transient transfection, neither ochre nor opal suppression were efficient in HCT116 cells (Figure 3A, left column). Selection for stable transfectants (with the same antibiotic concentration for all three stop codon constructs) yielded polyclonal pools of HCT116 cell lines that express GFP^{150CpK}. GFP yield of the ochre and opal suppression polyclonal populations was however more than an order of magnitude lower (mean fluorescence 81 and 109 arbitrary units, respectively) than amber suppression (mean at 1687 arbitrary units) (Figure 3A, center column; Figure S3A). We isolated clones from the polyclonal population by flow cytometry, selecting for single cells with strong GFP signal after 24 h incubation with CpK. Six amber-suppressing clones, nine ochre- and one opal-suppressing clones were recovered and compared for GFP expression by flow cytometry (Figure S3A, B). For each clone, addition of CpK resulted in homogenous GFP fluorescence at higher intensity than observed for the polyclonal pool. The most efficiently suppressing clones for each stop codon are shown in Figure 3A (right column). Despite the clonal selection, maximal ochre and
opal suppression remained roughly one order of magnitude lower than amber suppression (Figure S3B). This confirmed the observation from transient transcription experiments that ochre and opal codons are more difficult to suppress.

Efficient suppression of the target stop codon is typically considered the most important variable for optimization, but the selectivity for ncAA labeling applications also crucially depends on low misincorporation at endogenous stop codons (Bartoschek et al. 2021). In order to compare the on- versus off-target incorporation of ncAA, we compared the best clones for each stop codon after extended CpK incorporation. Cell lysates were labeled by SPIEDAC with tetSiR during lysis (Figure 3B). The most prominent band corresponds to GFP\textsuperscript{150CpK-tetSiR} for each of the three stop codon-suppressing cell lines. The ochre suppression clonal cell lysate also featured a distinct band, suggesting a particularly strongly labeled endogenous protein. It is interesting to note, that endogenous opal codons do not appear to be more efficiently suppressed by Pyl\textsubscript{T} than the other two stop codons, despite the opal codon being known to show highest leakiness, i.e. readthrough with endogenous tRNAs (Floquet et al. 2012; Schmied et al. 2014; Palomar-Siles et al. 2022). In summary, productive ochre and opal suppression is possible to achieve in stable settings. Yet, as expected, amber suppression of the reporter GFP is markedly more efficient while exhibiting similarly low levels of incorporation into endogenous proteins.

**Dual suppression clonal cell line for amber and ochre co-suppression**

In the past years, several approaches have successfully implemented dual suppression of two different stop codons in transiently transfected mammalian cells, allowing for site-specific incorporation of two distinct ncAAs (Zheng et al. 2017; Xiao et al. 2013; Meineke et al. 2018, 2020; Beránek et al. 2019; Serfling et al. 2018). Encouraged by selection of ochre and opal suppressor cell lines expressing high levels of GFP\textsuperscript{150CpK}, we decided to investigate if dual suppression can be implemented in a stable setting. The compatibility of Tyr\textsubscript{T}/RS and Leu\textsubscript{T}/RS pairs from *E. coli* in combinations with archaeal Pyl\textsubscript{T}/PylRS has been explored in detail using transient transfection (Zheng et al. 2017). Choosing the two most efficiently suppressed stop codons (Figure 3A), we explored the combination of the ochre suppressor Pyl\textsubscript{T\textsuperscript{UUA}}/PylRS pair with a compatible orthogonal amber suppressor tRNA/aaRS pair for dual integration. We integrated Pyl\textsubscript{T\textsuperscript{UUA}}/RS, Tyr\textsubscript{T\textsuperscript{CUA}}/AzFRS and a GFP\textsuperscript{102TAG150TA} dual suppression reporter into HCT116 cells (Figure 4A). After co-selection with puromycin, geneticin and blasticidin, a small subpopulation of the polyclonal pool showed increased GFP fluorescence after 48 h with CpK and AzF. We isolated single cells with high GFP fluorescence by FACS and recovered clonal
populations with a varied range of dual suppression efficiency (Figure S4A). The most efficient clone showed robust, CpK- and AzF-dependent expression of GFP\textsuperscript{102AzF150CpK} (Figure 4B, C and S4B, C). No cross-incorporation of ncAA was observed as judged from the lack of GFP-positive cells after incubation with only one of the ncAAs (Figure 4B, C). We purified GFP expressed in presence of AzF and CpK from the clonal cell line (yield 80 ng/ml culture) and confirmed incorporation of both ncAAs by intact mass spectrometry (Figure 4D). In addition to CpK, GFP\textsuperscript{102AzF150CpK} contains AzF, which can also be used for bioorthogonal labeling. The azido group can react with alkynes in Cu(I) catalysed azide-alkyne cycloaddition (CuAAC) or with strained alkynes in copper independent strain promoted azide-alkyne cycloaddition (SPAAC). We visualized purified GFP\textsuperscript{102AzF150CpK} by SPIEDAC and SPAAC with tetSiIR and dibenzylcyclooctyne-PEG4-5/6-tetramethyl-rhodamine (DBCO-TAMRA), respectively (Figure 4E). In summary, stable dual suppression cell lines are useful for protein production.

**Amber suppression cell lines for cell surface receptor expression**

We have so far described PB cell line generation with an easily detectable, well expressed and stable GFP reporter. To demonstrate the robustness of our approach, we chose a range of other proteins involved in different cellular processes. We generated stable amber suppression cell lines in HCT116 cells for expression of a number of intracellular proteins of interest, either with premature TAG or STELLA-N-ter tag and C-terminal HA-tag (Lafranchi et al. 2020): The cytoskeletal protein tubulin, heterochromatin associated DAXX, the histone variant CENP-A and the microproteins PIGBOS and SARS2-CoV-M (Figure S5A). We used the PyIRS “AF” variant in these cell lines to incorporate axial trans-cyclooct-2-ene-L-lysine (TC0*K) (Nikić et al. 2014; Yanagisawa et al. 2008). Polyclonal stable populations were selected and all proteins were detectable by immunoblotting for their HA-tag, with expected variation in expression levels. SPIEDAC labeling with tetSiIR or tetrazine-AF488 (tet488) in cell lysates shows specific bands for the TCO*K-labeled proteins according to their respective size, albeit endogenous proteins appear to be labeled with similar or even higher intensity (Figure S5B, C). Fluorescence microscopy of fixed and permeabilized cells labeled with methyl-tetrazine-BDP-FL (metetBDPFL) and counterstained for HA further showed that bioorthogonal labeling created a strong ambient background signal across cytosol and nucleus while a specific signal from the amber suppressed protein was only marginally brighter (Figure S5D). This is most evident for the Golgi-restricted CoV-M protein (Lafranchi et al. 2020). Therefore, the strong expression and selective suppression of our GFP reporter must be considered an outlier compared to many
biologically relevant proteins that are desirable to label using amber suppression. Evidently, the
known large plasmid copy number, ten to hundred thousands copies per cell (Fliedl et al. 2015),
after transient transfection provides a much larger window for selective suppression of the target
amber codon than expression from stable PB integrants. Nevertheless, we considered that
mislabeled of endogenous C-terminal amber stop codons may generate predominantly
intracellular background and using a cell-impermeable dye may allow selective labeling of target
proteins on the cell surface.

To test this hypothesis, we used two cell surface receptors for generation of stable amber
suppression HCT116 cell lines: First, the class B GPCR corticotropin-releasing factor type 1
receptor (CRFR1). CRFR1 A95TAG has previously been used in amber suppression studies for
ncAA mediated bioorthogonal labeling on the surface of live cells (Coin et al. 2013; Serfling et
al. 2018; Meineke et al. 2020, 2021). Second, a minimal synthetic Notch receptor (SynNotch).
An amber codon at N1703 (in the Notch1 gene, position N442 in our construct) places the ncAA
on the cell surface close to the membrane spanning helix of the C-terminal fragment generated
by proteolysis during maturation of the receptor (Meineke et al. 2020).

Polyclonal stable integrants of each receptor were cultured with TCO*K for 48 h and live cells
were SPIEDAC-labeled with tetazine-Cy3 (tetCy3) and subsequently fixed for microscopy.
CRFR1^{95TCO*K} and SynNotch^{442TCO*K}-expressing cells showed a strong, TCO*K-dependent
tetCy3 signal on the surface of cells (Figure 5A, B). Expression levels were much more
homogeneous across the population compared to transient transfection (Meineke et al. 2020).
The same CRFR1^{95TAG} and SynNotch^{442TAG} cell lines were labeled live with tet488 and lysates
were labeled with tetSiR for comparison (Figure 5C). While the tet488 signal was highly specific
for the cell surface receptor and overlapped with the HA-immunostaining signal, tetSiR in
lysates preferentially labeled endogenous proteins. In summary, stable amber suppression cell
lines provide limited opportunity for selective bioorthogonal labeling of intracellular proteins, but
enable efficient suppression and labeling of extracellular receptors, offering an ideal platform for
biophysical studies of membrane receptors in their native environment.

Dual site-specific ncAA incorporation and dual color labeling by combined amber and
ochre suppression in transgenic HCT116 cells

Building on our promising results for combining amber and ochre suppression with PyrT^{UA}/RS
and TyrT^{UA}/AzFRS for the GFP reporter, we attempted generation of stable dual suppression
cell lines in HCT116 cells for SynNotch. The SynNotch^{204TA442TAG} amber and ochre double
mutant can incorporate two distinct ncAAs for dual bioorthogonal derivatization when transiently cotransfected with appropriate, orthogonal tRNA/aaRS pairs. The incorporation sites are designed to place the two ncAAs in close spatial proximity to each other, but on separate parts of the mature, proteolytically processed SynNotch receptor (Meineke et al. 2020). Initial attempts to generated PB transgenic cell lines for SynNotch204TA442TAG-HA failed, even when we exchanged the ochre suppressor PytT for the M15 mutant (Serfling et al. 2018), which we had found to enhance ochre suppression efficiency before (Meineke et al. 2020). We recovered PytRS and AzFRS expressing populations after selection, but were unable to detect ncAAs or HA-tag (data not shown). To facilitate selection of integrants with SynNotch receptor expression by fluorescence-activated cell sorting (FACS), we exchanged the C-terminal HA-tag for a GFP-tag. We transfected HCT116 cells with M15UUA/PytRS, TyrCUA/AzFRS, M15UUA/SynNotch204TA442TAG-GFP and PBT. After selection with geneticin, puromycin and blasticidin, GFP fluorescence could be detected in a small number of cells after culturing in presence of AzF and CpK for 3 days. From this population, we isolated single cells by FACS for highest GFP fluorescence (Figure S6A). One clonal population showed increased GFP fluorescence with both ncAAs over a 12 to 72 h time course (Figure S6B). Flowcytometry and imaging of this clone confirmed that GFP fluorescence was not the result of cross-incorporation, readthrough with endogenous amino acids or a secondary translation start, as it localized to the cell membrane and depended on addition of both CpK and AzF (Figure 6A, S6C).

We validated the functionality of the two ncAAs for dual bioorthogonal labeling with tetCy3 and DBCO-AF647 (DBCO647) by SPIEDAC and SPAAC, respectively, on SynNotch204CpK442AzF-GFP purified from the clonal population (Figure 6B). In SDS-PAGE the two proteolytic fragments of SynNotch are separated. The Cy3 and 647 fluorescence signals are isolated which demonstrates that each receptor domain is selectively labeled and confirms site-specific incorporation of the two ncAAs. As expected, the DBCO647 signal from the AzF labeled C-terminal transmembrane domain overlapped with anti-GFP immunostaining signal. We observe two diffuse bands for the glycosylated N-terminal domain, due to partial proteolytic cleavage at the TEV site (Meineke et al. 2020). A small proportion of unprocessed SynNotch204CpK442AzF-GFP was also captured, visible as a faint high molecular weight band in both fluorescence channels and immunoblotting. We validated assignment of the bands by transient transfection of SynNotch204TA442TAG-GFP with wild type PytT/RS and SPIEDAC labeling on live cells, which resulted in the same band pattern (Figure S6D).
Live-cell dual color bioorthogonal labeling of a receptor in stable cell lines.

Next, we wondered if SynNotch<sup>204CpK442AzF</sup>-GFP, produced by dual suppression in our stable clonal cell line, could also be labeled on live cells. Cells were grown in presence of both ncAAs for 72 h, labeled with tetCy3 and alkyne-AF647 (alk647) and imaged for fluorescence of both dyes (Figure 6C). We detected overlapping, ncAA dependent cell surface signals for all three fluorophores associated with SynNotch<sup>204CpK442AzF</sup>-GFP. The Cy3, 647 and GFP fluorences were dependent on addition of both ncAA. When CuAAC was performed on live cells, both GFP and Cy3 signals were decreased; highlighting the potential of copper to damage cellular proteins even at low concentration (Figure S6E). Using a copper-free SPAAC reaction with DBCO-467 as as alternative for labeling AzF, however, was not successful, as it leads to strong unspecific labeling of endogenous proteins (Figure S6F) (van Geel et al. 2012).

Clonal SynNotch<sup>204CpK442AzF</sup>-GFP expressing cells labeled by SPIEDAC and CuAAC under live conditions yield the same band pattern observed for affinity purified SynNotch<sup>204CpK442AzF</sup>-GFP when separated by SDS-PAGE (Figure 6D). Two ncAA-dependent bands with tet488-fluorescence are visible for the N-terminal SynNotch domain, due to partial proteolysis at the TEV-site. An additional higher molecular weight band represented an artifact of tet488 surface labeling since it was present in all dye-treated samples, even in absence of ncAAs (Figure 6D). After CuAAC labeling a single fluorescent band is visible, corresponding to C-terminal TMD-GFP, comprising N442AzF specifically labeled by alk647 (Figure 6D). SPAAC labeling of SynNotch<sup>204CpK442AzF</sup>-GFP with DBCO647 lead to a diffuse signal across the entire lane, owing to DBCO reactivity with protein residues other than the incorporated azide ncAA (Figure S6G). Taken together our results demonstrate that dual suppression and bioorthogonal derivatization of two distinct ncAAs with two mutually orthogonal reactions is possible in a stable cell line with integrated suppression machinery. The resulting fluorescence signal strength was limited by the low efficiency of dual stop codon suppression and by the limited choice of mutual orthogonal reactions.

DISCUSSION

Systematic evaluation of suppression systems and mammalian host cell lines enabled us to generate stable cell lines with expanded genetic codes for suppression of all three stop codons using PB transgenesis. We established a universal two-plasmid design for generating mammalian cell lines with the choice of three distinct orthogonal tRNA/Synthetase pairs. This versatility should enable incorporation of virtually any genetically encoded ncAA described to
date in stable cell lines. While stop codon suppression, especially for ochre and opal codons, can be initially low after drug selection, highly efficient clones can be isolated. In the present study. We did not systematically assess DNA copy number and resulting relative expression levels of PylRS, PylT and the target protein in a larger number of clones. Consequently, a universal recipe for the optimal dosage of tRNA, aaRS and gene of interest cannot be derived from our data, and the best condition may be highly context specific with respect to cell line, target protein, and ncAA used. Nevertheless, we find that maximizing tRNA and gene of interest expression are determinants of ncAA incorporation efficiency and selectivity. In addition, position and context of the stop codon have been shown to greatly influence suppression efficiency and hence are important parameters to optimize (Bartoschek et al. 2021).

Our approach benefits from the efficiency of PB transgenesis. Optimal clones can be isolated by selection without a priori knowledge of the ideal copy number configuration. As an alternative to genomic integration, an episomal self-replicating plasmid has been used to long-term express the amber suppression machinery in human hematopoietic stem cells (Shao et al. 2020). This approach, however, will likely require continuous selection to maintain stable transgene copy number and efficient amber suppression. Stable integration and clonal selection, including by PB transposition, has been also used to generate amber suppression CHO cell lines efficiently expressing ncAA-modified antibodies for drug conjugation (Oller-Salvia et al. 2018; Axup et al. 2012).

Building on the toolbox created in this study, we proceeded to demonstrate site-specific incorporation and chemical labeling of two distinct ncAAs in a stable mammalian cell line. Our study highlights opportunities and limitations of dual suppression systems, requiring two selective, efficient and mutual orthogonal incorporation machineries. The relatively lower efficiency of ochre and opal stop codon suppression limits overall yield of dually suppressed protein (Zheng et al. 2017; Meineke et al. 2018, 2020; Osgood et al. 2022). We chose AzF for encoding the CuAAC handle because of its efficient incorporation (Figure 2A, S2B), but this likely limited the fluorescent labeling efficiency on live cells. We have previously shown that a genetically encoded picolyl-azide provide superior CuAAC reactivity with azide dyes (Meineke et al. 2021), but orthogonal machineries to introduce picolyl-azide and TCO*K or CpK do not exist to date.

It is an important goal of the field to achieve selective labeling of amber codons in target proteins expressed at endogenous level or even endogenously edited genes. Our experiments with the GFP reporter demonstrate that, under most optimal conditions, endogenously
expressed proteins can be ncAA labeled with high efficiency and good selectivity. However, our
data also underscores that such an ideal scenario is unlikely to apply to the average cellular
protein. Our methodology provides most attractive applications for labeling proteins on the
plasma membrane. Site-specifically incorporated ncAAs on the cell surface face little
competition from misincorporated ncAAs at endogenous stop codons due to the fact that
C-termini of plasma membrane proteins are mostly intracellular or posttranslationally processed.
Our stable genetic code expansion technology, together with fluorescent bioorthogonal labeling
on the cell surface, provides a valuable toolbox to investigate the localization, movement and
clustering of membrane receptors on the cell surface. Dual-color labeling will enable new
applications for studying conformational dynamics and multimeric assemblies of cell surface
proteins.
Figure 1 - PB transgenesis integrates PylT/RS amber suppression machinery into the host cell genome.

A) Schematic representation of pAS (Amber Suppression) PB transgenesis plasmids. N-terminally FLAG-tagged PylRS (PylRS) is expressed under control of the EF1α promoter. An internal ribosome entry site (IRES), followed by a puromycin resistance selectable marker (PuroR), is positioned immediately downstream of PylRS. A cassette of four tandem repeats of PylT each controlled by h7SK PolIII promoter are positioned upstream of the EF1α promoter. The entire PylT/RS cassette is flanked by insulator sequences (Ins) and 3' and 5' inverted repeats for genomic integration by PB transposase (pink triangle). The GFP\textsuperscript{150TAG} amber suppression reporter pAS construct has the same architecture, but contains a blasticidin resistance marker (Bsd) instead.

B) Immunoblot comparing GFP\textsuperscript{150Cpk} reporter expression in different mammalian PB integrant cell lines. The chemical structure of Cpk is shown at the top of the panel. HEK293T (lanes 1 and 2), HEK293 (lanes 3 and 4), HCT116 (lanes 5 and 6), A375 (lanes 7 and 8), COS-7 (lanes 9 and 10) and U-2 OS (lanes 11 and 12) PB amber suppression integrants were grown with 0.5 mM Cpk (+, even numbered lanes) or without nCAA (−, odd numbered lanes) for 24 h. Soluble lysates were SPIEDAC labeled with 1 μM tetSiR and normalized for total protein content. SiR signal was visualized at 630 nm in-gel after SDS-PAGE. Immunostaining for GFP expression levels, FLAG-PylRS and GAPDH.

C) Quantification of GFP fluorescence in lysates of a low selected (2 μg/ml puromycin, 500 μg/ml blasticidin), stable PylT/RS and GFP\textsuperscript{150TAG} integrant, HEK293T cell line. Where indicated (yellow shading), the cell line was transiently transfected with different pAS plasmids bearing 4xPylT and/or PylRS variants. Transfections were performed in quadruplicate with 0.2 mM Cpk addition for 24 h in triplicate. GFP fluorescence signal is normalized to untransfected cells cultured in presence of 0.2 mM Cpk for 24 h, background fluorescence (−Cpk) values were subtracted. Error bars indicate standard deviation. Plasmids transfected by lanes: 2 no transfection, lane 3: control transfection with a pAS control plasmid, lane 4: transfection with pAS (4xPylT), lane 5: transfection with pAS_MmaPylRS (PylRS), lane 6: transfection with pAS_4xMmaPylT/RS (4xPylT and 4xPylRS).

D) Immunoblot for PylRS in lysates shown in C). Lanes 1 and 2: no transfection without (−) and with (+) Cpk, others lanes as in C). Soluble lysates were separated by SDS-PAGE. Immunostaining for FLAG-PylRS and β-actin.
E) Flow cytometry comparing amber suppression in HEK293T (top), HEK293 (middle) and HCT116 (bottom) after transient transfection (left) and PB transgenesis (right). Cells were transfected transiently with PyIT/RS and the PyIT/GFP<sup>150TAG</sup> reporter. Alternatively, the PyIT/RS and the PyIT/GFP<sup>150TAG</sup> reporter were integrated by PB transgenesis using the same plasmids. GFP fluorescence from the sfGFP<sup>150TAG</sup> reporter was measured after 24 h without ncAA (dashed black line) or in presence of 0.2 mM CpK (green line).
**Figure 2**

**A**

Tyr\textsuperscript{CUA}/AzFRS

- ncAA
- +AzF

HCT116 integration transient

**B**

Leu\textsuperscript{CUA}/AnapRS

- ncAA
- +Anap

Anap fluorescence [AU]

**C**

<table>
<thead>
<tr>
<th>Tyr\textsuperscript{CUA}/AzFRS</th>
<th>Leu\textsuperscript{CUA}/AnapRS</th>
</tr>
</thead>
<tbody>
<tr>
<td>AzF - integration - transient</td>
<td>AzF - integration - transient</td>
</tr>
<tr>
<td>AzF + integration - transient</td>
<td>AzF + integration - transient</td>
</tr>
<tr>
<td>AzF - integration + transient</td>
<td>AzF + integration + transient</td>
</tr>
</tbody>
</table>

GFP\textsuperscript{150ncAA}

α-GFP

AzFRS

α-FLAG

actin

α-actin

*CC-BY-NC-ND 4.0 International license available under a (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made The copyright holder for this preprint this version posted March 27, 2023. ; https://doi.org/10.1101/2023.03.26.534279 doi: bioRxiv preprint*
Figure 2 - PB integration of TyrT\(^{CUA}\)/AzFRS and LeuT\(^{CUA}\)/AnapRS in HCT116 cells.

A) Flow cytometry of amber suppression cell line generated with E. coli TyrRS-derived AzFRS. The chemical structure of AzF is shown. HCT116 cells were transiently transfected with TyrT\(^{CUA}\)/AzFRS and corresponding TyrT\(^{CUA}\)/GFP\(^{150TAG}\) reporter (transient, top). The same plasmids were integrated by PB transgenesis (integration, bottom). GFP fluorescence from the sfGFP\(^{150TAG}\) reporter was measured after 24 h without ncAA (dashed black line) or in presence of 0.5 mM AzF (green line, chemical structures shown).

B) Flow cytometry of amber suppression cell line generated with E. coli LeuRS-derived AnapRS. The chemical structure of Anap is shown. HCT116 cells were transiently transfected with LeuT\(^{CUA}\)/AnapRS and cognate LeuT\(^{CUA}\)/GFP\(^{150TAG}\) reporter (transient, top). The same plasmids were integrated by PB transgenesis (integration, bottom). GFP and Anap fluorescence were measured without ncAA (black) or after 24 h incubation with 0.05 mM Anap and a 2 h chase with ncAA-free medium (cyan) before measurement.

C) Immunoblotting of lysates prepared from transiently transfected and PB integrant HCT116 cells cultured in presence of ncAA as in A) and B). Samples were separated by SDS-PAGE and immunostained for GFP\(^{150ncAA}\), FLAG-AzFRS and an actin loading control.
Figure 3

(A) Graphs showing the normalized GFP fluorescence levels for transient, polyclonal, and clonal samples with and without CpK. The fluorescence is normalized to the mode for each condition.

(B) Western blot analysis of GFP150* samples with different tagging sequences (TAG, TAA, TGA) and presence (+) or absence (-) of CpK. The blots show the expression of GFP150CpK-tetSiR, α-GFP, and γ-FLAG proteins.
Figure 3 - Generation of efficient ochre and opal suppressor clonal HCT116 populations by PB transgenesis.

A) Flow cytometry comparing amber (UAG, top), ochre (UAA, center) and opal (UGA, bottom) suppression in HCT116 cells after transient transfection (left) and PB transgenesis (center and right). GFP fluorescence was measured without ncAA (dashed black line) or 24 h after addition of 0.2 mM CpK (green line). Left: Cells were transiently transfected with PylT/RS and the PylT/GFP^{150stop} reporter, suppressor PylTs have AUC, AUU, ACU anticodons to match the respective reporter’s UAG, UAA and UGA stop codons. Center: using the same plasmids, both PylT/RS and PylT/GFP^{150stop} reporter pAS constructs were integrated into the genome by PB transgenesis. Right: Clonal populations, isolated from the polyclonal pools (middle) by FACS, after 24 h incubation with 0.2 mM CpK. Best clones are shown for amber, ochre and opal suppression, respectively. Other clones recovered are compared in Figure S3.

B) Immunoblot of lysates from clonal stop codon suppressor HCT116-derived cell lines shown in (A). The asterisk (*) is a placeholder for the stop codon indicated. Clonal populations were grown in presence of 0.5 mM CpK for 72 h. GFP^{150CpK} was SPIEDAC labeled with 1 μM tetSiR in lysate followed by SDS-PAGE and immunostaining for GFP expression and FLAG-aaRS expression.
Figure 4

A. Diagram showing the interaction of TyrCUA/AzFRS and PyITUA/RS with GFP mRNA. The mRNA is represented as a bar chart with AUC and AUU/UAU sites. The spliceosome is depicted as a purple and orange structure, with CpK and AzF acting as regulators.

B. Graph showing normalized GFP fluorescence (fluorescence [AU]) plotted against GFP fluorescence levels. The graph includes data for -ncAA, +CpK, +AzF, and ++ conditions.

C. Table showing the effect of CpK and AzF on GFP fluorescence. The table includes columns for -CpK/-AzF, +CpK/-AzF, -CpK/+AzF, and +CpK/+AzF.

D. Mass spectrometry (MS) spectrum showing the calculated mass of GFP102AzF150CpK as 27098.5 Da. The spectrum includes peaks at 2710.76 and 3011.9 Da.

E. Western blot analysis showing the expression of SPAAC, TAMRA stain, and TAMRA merge. The blot also includes a control for GFP and α-FLAG.
Figure 4 - Dual suppression clonal cell line for GFP<sup>102AzF150Cpk</sup> expression in HCT116 cells.

A) Scheme illustrating amber and ochre co-suppression in HCT116 cells with PylT<sub>UAA</sub>/PylRS ochre suppressor pair, TyrT<sub>UAA</sub>/AzFRS amber suppressor pair and a GFP<sup>102TAG150TAA</sup> reporter integrated by PB transgenesis. Addition of AzF and Cpk allows GFP<sup>102AzF150Cpk</sup> production.

B) Flow cytometry of the HCT116 PylT<sub>UAA</sub>/PylRS, TyrT<sub>UAA</sub>/AzFRS and GFP<sup>102TAG150TAA</sup> integrant clone F6 grown without ncAA (–ncAA, black), with 0.2 mM Cpk (+Cpk, orange), with 0.5 mM AzF (+AzF, violet) or both ncAAs (++, green) for 48 h.

C) Immunoblot for lysates prepared from the GFP<sup>102TAG150TAA</sup> integrant clone F6 grown without ncAA, with 0.2 mM Cpk, with 0.5 mM AzF or both ncAAs for 48 h. Lysates were separated by SDS-PAGE followed by immunostaining for GFP expression and FLAG-aaRS expression.

D) Intact mass determination for purified GFP<sup>102AzF150Cpk</sup> expressed in the dual suppressor clonal cells. Calculated theoretical mass and determined mass are indicated.

E) GFP<sup>102AzF150Cpk</sup> purified from the PB dual suppressor HCT116 clone by SPAAC and SPIEDAC with 1 μM DBCO-TAMRA and 1 μM tetSiR, respectively. Fluorescence was detected at 520 nm (TAMRA) and 630 nm (SiR) after separation by SDS-PAGE, fluorescence merge and coomassie-stained gel are shown.
Figure 5

A

CRFR\textsuperscript{95TCO*K-tetCy3}  

Hoechst  Cy3  merge

-ncAA

+TCO*K

B

SynNotch\textsuperscript{442TCO*K-tetCy3}  

Hoechst  Cy3  merge

-ncAA

+TCO*K

C

\begin{tabular}{c}
\textbf{CRFR}^\text{95TCO*K} & SynNotch^\text{442TAG} \\
\text{tetzine} & 488 & SIR & 488 & SIR \\
\text{CRFR}^\text{195TCO*K} & SynNotch^\text{442TCO*K} & TMD \\
\text{CRFR}^\text{195TCO*K-tet488} & SynNotch^\text{442TCO*K-tet488} & TMD \\
\end{tabular}
Figure 5 - Amber suppression cell lines for bioorthogonal labeling of cell surface receptors.

Fluorescence microscopy imaging of HCT116 PB integrant cell lines for cell surface receptor expression by amber suppression with (A) PylT/RS AF and PylT/CRFR1$^{95TAG}$ or (B) PylT/RS AF and PylT/SynNotch$^{442TAG}$ integrated by PB transposition. Integrant populations recovered after selection were grown in presence of 0.1 mM TCO*K for 48 h, chased in ncAA free medium for 2 h prior to SPIEDAC labeling with 1 μM tetCy3. Cells were fixed in PFA and nuclei were counterstained with Hoechst 33342 prior to imaging on a Nikon Ti2. White scale bars indicate 25 μm.

C) Immunoblot for the same cell lines as in A) and B) expressing CRFR1$^{95TCO*K}$ or SynNotch$^{442TCO*K}$. Cells were grown in presence of 0.1 mM TCO*K for 48 h. The cell surface receptors were either labeled by SPIEDAC with 2 μM membrane impermeable tet488 before lysis or with 1 μM tet-SiR in cell lysate. Cleared lysate aliquots were separated by SDS-PAGE and tet488 surface fluorescence and tetSiR lysate labeling visualized at 460 nm (488, cyan) and 630 nm (SiR, magenta), respectively. Immunostaining for the C-terminal HA-tags confirms specific SPIEDAC surface labeling of CRFR1$^{95TCO*K}$ and SynNotch$^{442TCO*K}$. Immunostaining for FLAG-PylRS AF is shown. The SynNotch illustration was simplified by omitting the N-terminal LaG17 fusion.
Figure 6

A

SynNotch^204TAA442TAG-GFP

Normalized to Mode

GFP fluorescence [AU]

B

tetCy3

DBCO 647

α-GFP

merge

Cy3

647

D

A2F+CpK

tet488

alk647

α-GFP

merge

488

647

merge

+Hoechst

647

Cy3

GFP

+AzF

+CpK

+AzF

++
Figure 6 - Dual suppression HCT116 clonal cell line for SynNotch^{204 CpK442 AzF} expression and live-cell dual bioorthogonal labeling.

A) The dual suppressor clonal HCT cell line with M15^{UJA}/PyLRS ochre suppressor pair, Tyr^{CUA}/AzFRS amber suppressor pair and SynNotch^{204TAA442TAG}-GFP integrated via PB transgenesis was grown 72 h without ncAA (--ncAA, black), 0.2 mM CpK (+CpK, orange), 0.5 mM AzF (+AzF, violet) or both ncAAs (++, green) and analyzed by flow cytometry for GFP fluorescence. The positions of both ncAAs in the domains of the mature receptor are illustrated in the schematic. All SynNotch illustrations were simplified by omitting the N-terminal LaG17 fusion.

B) SynNotch^{204 CpK442 AzF}-GFP purified from the dual suppression HCT116 integrant clone D8. Bioorthogonal labeling by SPAAC and SPIEDAC with 2 µM DBCO647 and 2 µM tetCy3, respectively. Fluorescence of the bead-captured protein after separation by SDS-PAGE at 520 nm (Cy3), 630 nm (647) and merged signals for in-gel fluorescence, as well as immunoblotting against the C-terminal GFP are shown. The different domains and processing states of SynNotch are indicated - white arrow: 81 kDa pre-processed SynNotch^{204 CpK442 AzF}; yellow arrow: 43 kDa N-terminal domain with 204CpK labeled by tetCy3; magenta arrow: 38 kDa C-terminal TMD-GFP fusion with 442AzF labeled by DBCO647; yellow asterisk: additional SynNotch N-terminal fragment after proteolytic cleavage of LaG17 (Meineke et al. 2020).

C) Fluorescence microscopy for SynNotch^{204 CpK442 AzF}-GFP expression in the clonal dual suppressor cell line. Cells were grown in presence of 0.2 mM CpK and 0.5 mM AzF for 72 h, labeled with 2 µM tetCy3 by SPIEDAC and 2 µM alk647 by CuAAC, respectively. Prior to imaging on a Nikon Ti2, nuclei were counterstained with Hoechst 33342. White scale bar indicates 25 µm. See also Figure S6E.

D) SPIEDAC and CuAAC labeling with 2 µM tet488 and 2 µM alk647 on live HCT116 clonal SynNotch^{204 CpK442 AzF}-GFP cells grown in presence of 0.2 mM CpK and 0.5 mM AzF for 72 h. Lysates were separated by SDS-PAGE. In-gel fluorescence captured at 460 nm (488) and 630 nm (647) and immunoblotting against the C-terminal GFP are shown. Arrows indicate the different processing states of SynNotch: Cyan arrow indicates the 43 kDa N-terminal domain which contains 204CpK-tet488 after SPIEDAC, magenta arrow indicates 38 kDa C-terminal TMD-GFP fusion which comprising 442AzF-alk647, asterisk indicates an additional SynNotch N-terminal fragment without LaG17 (Meineke et al. 2020).
MATERIAL AND METHODS

DNA constructs

The *M. maezi* Pyt/RS expression and superfolder GFP<sup>150TAG</sup> (sfGFP<sup>150TAG</sup> and sfGFP, referred to as GFP<sup>150TAG</sup> and GFP throughout) reporter constructs with four tandem h7SK PytT repeats have been described previously (Meineke et al. 2018). Analogous constructs for AnapRS and AzFRS with the respective cognate LeuT<sup>CUA</sup> or TyrT<sup>CUA</sup> were generated for this study. The plasmids share a common architecture and are here collectively referred to as “pAS” (Amber Suppression) plasmids: The aaRS, reporter or gene of interest are controlled by EF1α promoter and followed by a internal ribosome entry site (IRES) that allows expression of a downstream selection marker. A cassette with four tandem repeats of the tRNA gene, controlled by the human 7SK PolIII promoter, are placed upstream of the EF1α promoter in anti-sense orientation. A schematic of this architecture is shown in Figure 1A. All DNA constructs were verified by Sanger sequencing and most plasmids are available through Addgene.

Cell Culture and Transfection

HEK293, HEK293T, HCT116, A375, COS-7 and U-2 OS cells were maintained in Dulbecco’s modified Eagle’s medium (DMEM, GlutaMAX, Thermo) supplemented with 10 % (v/v) FBS (Sigma) at 37 °C and 5% CO<sub>2</sub> atmosphere. For transfection (1.5-2.0) x 10<sup>5</sup> cells/ml were seeded 24 h before transient transfection with 1 μg plasmid DNA/ml using TransIT-LT1 (Mirus) according to the manufacturer’s instructions. In transient transfection experiments, ncAAs were added at the time of transfection and cells were harvested after 24, 48 or 72 h, as indicated.

Generation of PB-mediated stable integration cell lines

Cell lines with PiggyBac (PB) integrated transgenes were generated as described (Elsässer et al. 2016; Elsässer 2018): Parental cells were seeded one day prior to transfection and transfected with desired pAS plasmids and PBT in a 4:1 ratio. Earliest 48 h after transfection, cells were split and selected with different concentrations of appropriate antibiotics for 7 days (ranges: 2-5 μg/ml puromycin (VWR), 500-4000 μg/ml blasticidin (Invivogen), 2000-3000 μg/ml geneticin (G418, Sigma)). Stable integrant cell populations were recovered from the highest submissible selection condition in DMEM + 10% FBS (v/v). Polyclonal populations were grown in presence of ncAA for 24, 48 or 72 h, as indicated, and characterized by flow cytometry for reporter fluorescence and immunoblotting for reporter or protein of interest and PytRS. A list of
cell lines generated in this study can be found on Mendeley data (DOI: 10.17632/tgc7mbv5xp.1).

Noncanonical amino acids
Working stocks of ncAAs were prepared in 100 mM NaOH with 15% DMSO (v/v) to 100 mM. Anap working stock was prepared to 20 mM in the same buffer. Chemical structures of N-ε-[2-methyl-2-cyclopropene-1-yl]-methoxy] carbonyl-L-lysine (CpK), p-azido-phenylalanine (AzF), 3-(6-acetylnaphthalen-2-ylmimo)-2-amino-propanoic acid (Anap) and axial trans-cyclooct-2-ene-L-lysine (TCO*K) are shown in Figures 1, 2 and 5, respectively. Refer to the Key Resource Table for supplier and CAS numbers.

SPIEDAC lysate labeling, SDS-PAGE and immunoblotting
Aliquots of cell lysate were separated on 4-20 % Tris-glycine gels (BioRad) and transferred to nitrocellulose membranes (BioRad). Cells were lysed in RIPA (for intracellular proteins) or PBS with 0.1% triton X-100 (v/v) (cell membrane proteins) supplemented with 1x Complete proteinase inhibitor (Roche). The insoluble fraction was removed by centrifugation. To compare amber suppression of PB transgenic integrants generated from different human cells, the soluble fraction of whole cell lysates were normalized for total protein content (BCA assay, Pierce). For SPIEDAC labeling in lysate, 1 μM tetrazine-Silicon rhodamine (tetSiR) (Spirochrome) was added to the lysis buffer. Lysates cleared by centrifugation were mixed with 6x Laemmli buffer and denatured for 10 min at 37°C, or at 95°C to completly denature GFP. After separation, gels were exposed at 460 nm, 520 nm and 630 nm in a GE Al600 imager for in-gel fluorescence. Expression of GFP reporter, HA- or GFP-fusion proteins and FLAG-aARS was confirmed by immunoblotting with antibodies against GFP (Santa Cruz, sc-9996), HA-HRP (Roche, 12013819001), FLAG-HRP (Sigma, A8592), β-actin (cell signaling, 4970), GAPDH (Millipore, AB2302) and corresponding secondary HRP-conjugated antibodies when needed (BioRad). Annotated, preprocessed immunoblotting data can be found on Mendeley data (DOI: 10.17632/tgc7mbv5xp.1).

Quantification of GFP expression
HEK293T-derived cells, selected for pAS_4xMmaPylT/RS and pAS_4xMmaPylT/GFP150TAG PB transposition (low antibiotic concentration: 2 μg/ml puromycin, 500 μg/ml blasticidin), were
transiently transfected with an additional pAS plasmid in quadruplicate. Transfected cells were grown in absence or presence (triplicate) of 0.2 mM CpK for 24 h. Cells were lysed in RIPA buffer with 1x complete protease inhibitor (Roche), the insoluble fraction was removed by centrifugation. GFP bottom fluorescence of an aliquot was measured in Tecan Infinite M200 pro plate reader (excitation 485 nm, emission 518 nm). Fluorescence measurements were normalized to total protein content of each sample as determined by Pierce BCA assay kit (Fisher Scientific) on the same sample. Average GFP fluorescence was calculated from the technical triplicates, the GFP fluorescence measured in absence of CpK was subtracted as background fluorescence and values were normalized to average GFP fluorescence measured for the untransfected amber suppression cell line.

Flow cytometry and fluorescence-assisted cell sorting (FACS)
Cells were trypsinized after the indicated time of culturing in presence of ncAAs and resuspended in PBS + 10% FBS (v/v). Flow cytometry was carried out on a Beckmann NAVIOS flow cytometer and analyzed in FlowJo™ software (BD Life Science, version 10.6.2). Polyclonal populations selected with antibiotics after PB transgenesis were grown in the presence of the cognate ncAAs for 24 h (GFP150TAG, GFP150TAA, GFP150TGA), 48 h (GFP102TAG150TAA) or 72 h (SynNotch204TAA442TAG-GFP). Single cells were isolated by gating for top 0.4-5% GFP fluorescence signal on a SONY SH800 cell sorter. Clonal populations were expanded in DMEM + 10% FBS and analyzed by flow cytometry and immunoblotting for GFP and aaRS expression.

Intact mass spectrometry
The dual suppression GFP102TAG150TAA clonal integrant cell line was cultured in the presence of 0.2 mM and 0.5 mM AzF for 72 h. Cells were lysed in RIPA buffer supplemented with 1x complete protease inhibitor (Roche). The insoluble fraction was removed by centrifugation. Expressed GFP was captured on GFP-Trap_MA magnetic beads (ChromoTEK), washed with RIPA buffer, PBS + 500 mM NaCl and PBS prior to elution in 1 % (v/v) acetic acid. Purified GFP102AzF150CpK samples were desalted and rebuffered into 100 mM ammonium acetate pH 7.5 using ZebaSpin columns with a 7 kDa cut-off (Thermo). Samples were directly infused into an Orbitrap Fusion Tribrid mass spectrometer equipped with an offline nanospray source using borosilicate capillaries (Thermo). The capillary voltage was 1.5 kV and the pressure in the ion-routing multipole was maintained at 0.11 torr. Spectra were acquired in the Orbitrap mass
analyzer operated in high mass mode at a resolution of 60,000 between 800-3000 m/z. Data were analyzed using Excalibur (Thermo).

**Bioorthogonal labeling on beads**

GFP^{102}AzF^{150}CpK and SynNotch^{204}CpK^{442}AzF-GFP were captured on GFP-Trap_MA magnetic beads (ChromoTEK) from cell lysate of the respective PB stable clonal cell lines cultured in presence of 0.5 mM AzF and 0.2 mM CpK for 72 h. Cells were lysed and the insoluble fraction was removed by centrifugation. Beads were washed with RIPA buffer, PBS + 500 mM NaCl and PBS. SPIEDAC and SPAAC were carried out on bead-bound proteins simultaneously by adding 1 μM tetSiR (spirochrome) or 1 μM tetCy3 (Click Chemistry Tools) and 1 μM DBCO-TAMRA (Jena Bioscience) or 1 μM DBCO647 (Jena Bioscience) for 10 min on ice. Excess dye was washed off and bead-bound proteins were eluted with 1 % (v/v) acetic acid for SDS-PAGE separation. Equal amounts of GFP purified from different samples were separated on 4-20% Tris-glycine gels (BioRad) and exposed for in-gel fluorescence at 520 nm and 630 nm in a GE Al600 imager. The gel was stained with InstantBlue (Expedeon) to visualize GFP^{102}AzF^{150}CpK bands or transferred to a nitrocellulose membrane for immunostaining against GFP to detect SynNotch^{204}CpK^{442}AzF-GFP.

**Live-cell surface SPIEDAC and CuAAC labeling for in-gel fluorescence and fluorescence microscopy**

Surface SPIEDAC labeling was performed with 1-2 μM tet488 (Jena Bioscience) or 1-2μM tetCy3 (Click Chemistry Tools) in DMEM + 10% FBS (v/v) for 30 min at 37 °C. Where indicated SPAAC was carried out with 2 μM DBCO-AF647 (DBC0647, Click Chemistry Tools) in DMEM +10% FBS (v/v) for 30 min at 37 °C. For subsequent CuAAC in the dual suppression cell line, cells were washed with PBS and labeling was performed with 50 μM CuSO₄, 250 μM THPTA, 2.5 mM sodium ascorbate and 2 μM alk647 (Jena Bioscience) for 10 min at 22 °C. Cells were washed and collected in PBS and lysed in PBS with 0.2% triton X-100 (v/v) and 1x complete protease inhibitor (Roche) on ice. Lysates cleared by centrifugation were mixed with 6x Laemmlli buffer and denatured for 10 min at 37°C, or at 95°C to completely denature GFP. Proteins were separated by SDS-PAGE and in-gel fluorescence measured as described above. Alternatively, for cells grown in poly-L-lysine coated 18-well glass imaging slides (Ibidi), cells were washed with PBS, fixed in paraformaldehyde (PFA) 4% (v/v) and counterstained with 2 μM hoechst 33342 prior to imaging in PBS on a Nikon Eclipse Ti2 inverted widefield microscope, using a 20x
(0.75 NA) objective and filter sets for DAPI, GFP, Cy3 and Cy5 fluorescence. Images were processed in Fiji (Schindelin et al. 2012).

**Table 1 - Reagents and resources**

<table>
<thead>
<tr>
<th>Antibodies</th>
<th>Supplier</th>
<th>Cat#/RRID</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mouse monoclonal anti-GFP</td>
<td>Santa Cruz</td>
<td>Cat#sc-9996; RRID:AB_627695</td>
</tr>
<tr>
<td>Rat monoclonal anti-HA HRP-coupled</td>
<td>Roche</td>
<td>Cat#12013819001; RRID:AB_390917</td>
</tr>
<tr>
<td>Mouse monoclonal anti-FLAG HRP-coupled</td>
<td>Sigma</td>
<td>Cat#A8592; RRID:AB_439702</td>
</tr>
<tr>
<td>Rabbit monoclonal anti-β-actin</td>
<td>Cell Signaling</td>
<td>Cat#4970; RRID:AB_2223172</td>
</tr>
<tr>
<td>Chicken polyclonal GAPDH</td>
<td>Millipore</td>
<td>Cat#AB2302; RRID:AB_10615768</td>
</tr>
<tr>
<td>Goat anti-chicken IgY (H+L) HRP</td>
<td>Invitrogen</td>
<td>Cat#A16054; RRID:AB_2534727</td>
</tr>
<tr>
<td>Goat anti-mouse HRP-coupled</td>
<td>BioRad</td>
<td>Cat#1721011; RRID:AB_11125936</td>
</tr>
<tr>
<td>Goat anti-rabbit HRP-coupled</td>
<td>BioRad</td>
<td>Cat#1721019; RRID:AB_11125143</td>
</tr>
</tbody>
</table>

**Chemicals, Peptides, and Recombinant Proteins**

<table>
<thead>
<tr>
<th>Chemical/Peptide</th>
<th>Supplier</th>
<th>Cat#/CAS</th>
</tr>
</thead>
<tbody>
<tr>
<td>CpK: N-ε-{[2-methyl-2-cyclopropene-1-yl]-methoxy] carbonyl-L-lysine</td>
<td>SiChem</td>
<td>Cat#SC-8017; CAS:1610703-09-7</td>
</tr>
<tr>
<td>ProK: N-propargyl-L-lysine</td>
<td>Iris Biotech</td>
<td>Cat#HAA2090; CAS:1428330-91-9</td>
</tr>
<tr>
<td>TCO*K: axial trans-cyclooct-2-ene-L-lysine</td>
<td>SiChem</td>
<td>Cat#SC-8008; CAS:1801936-26-4</td>
</tr>
<tr>
<td>Anap: (2S)-3-[(6-Acetynaphthalen-2-yl)amino]-2-aminopropanoic Acid</td>
<td>Cayman</td>
<td>Cat#15436; CAS 1313516-26-5</td>
</tr>
<tr>
<td>AzF: 4-Azido-L-phenylalanine</td>
<td>Santa Cruz</td>
<td>Cat#sc-289923; CAS: 33173-53-4</td>
</tr>
<tr>
<td>tetrazine-Siliconrhodamine (tet-SiR)</td>
<td>Spirochrom</td>
<td>Cat#SC008</td>
</tr>
<tr>
<td>tetrazine-AF488 (tet488)</td>
<td>Click Chemistry Tools</td>
<td>Cat#1361</td>
</tr>
<tr>
<td>Dibenzylcyclooctyne-PEG4-5/6-Tetramethyl rhodamine (DBCO-TAMRA)</td>
<td>Jena Bioscience</td>
<td>Cat#CLK-A131</td>
</tr>
<tr>
<td>Cy3-tetrazine (tetCy3)</td>
<td>Click Chemistry Tools</td>
<td>Cat#1204</td>
</tr>
<tr>
<td>alkyne-AF647 (alk647)</td>
<td>Jena Bioscience</td>
<td>Cat#CLK-1301A</td>
</tr>
<tr>
<td>Dibenzylcyclooctyne-AF647 (DBCOAF647)</td>
<td>Jena Bioscience</td>
<td>Cat#CLK-1302</td>
</tr>
</tbody>
</table>

**Critical Commercial Assays**

<table>
<thead>
<tr>
<th>Assay</th>
<th>Supplier</th>
<th>Cat#</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pierce BCA assay kit</td>
<td>Fisher Scientific</td>
<td>Cat#23227</td>
</tr>
<tr>
<td>GFP-Trap_MA magnetic beads</td>
<td>ChromoTek</td>
<td>Cat#gtma</td>
</tr>
</tbody>
</table>

**Deposited Data**

<table>
<thead>
<tr>
<th>Data</th>
<th>DOI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mendeley data</td>
<td>DOI: 10.17632/tgc7mbv5xp.1</td>
</tr>
</tbody>
</table>

**Experimental Models: Cell Lines**
<table>
<thead>
<tr>
<th>Human: HEK293T cells</th>
<th>N/A</th>
</tr>
</thead>
<tbody>
<tr>
<td>Human: HEK293 cells</td>
<td>N/A</td>
</tr>
<tr>
<td>Human: HCT116 cells</td>
<td>N/A</td>
</tr>
<tr>
<td>Human: A375</td>
<td>N/A</td>
</tr>
<tr>
<td>Human: U-2 OS</td>
<td>N/A</td>
</tr>
<tr>
<td>Simian: COS7</td>
<td>N/A</td>
</tr>
</tbody>
</table>

**Recombinant DNA**

<table>
<thead>
<tr>
<th>Piggy Bac Transposase expression plasmid</th>
<th>System Bioscience</th>
<th>Cat#PB5108-1</th>
</tr>
</thead>
<tbody>
<tr>
<td>pAS_MmaPyIT/PyIRS</td>
<td>(Meineke et al. 2018)</td>
<td>Addgene#140009</td>
</tr>
<tr>
<td>pAS_MmaPyIT/GFP150TAG</td>
<td>(Meineke et al. 2018)</td>
<td>Addgene#140015</td>
</tr>
<tr>
<td>pAS_MmaPyIT/GFP150TAG</td>
<td>(Meineke et al. 2020)</td>
<td>Addgene#154766</td>
</tr>
<tr>
<td>pAS_sfGFP102TAG133TGA150TAA</td>
<td>this study</td>
<td>N/A</td>
</tr>
<tr>
<td>pAS_MmaPyITRS</td>
<td>(Meineke et al. 2018)</td>
<td>Addgene#154762</td>
</tr>
<tr>
<td>pAS (MmaPyIT)</td>
<td>(Meineke et al. 2018)</td>
<td>Addgene#14008</td>
</tr>
<tr>
<td>pAS_MmaPyIT/AbKRS</td>
<td>this study</td>
<td>Addgene#140020</td>
</tr>
<tr>
<td>pAS_EcoLeuT/AnapRS</td>
<td>this study</td>
<td>Addgene#174894</td>
</tr>
<tr>
<td>pAS_EcoLeuT/GFP150TAG</td>
<td>this study</td>
<td>Addgene#174892</td>
</tr>
<tr>
<td>pAS_BstTyrT(UUA)/PyIRS</td>
<td>this study</td>
<td>Addgene#174896</td>
</tr>
<tr>
<td>pAS_BstTyrT(CUA)/EcoTyrRS</td>
<td>this study</td>
<td>Addgene#140018</td>
</tr>
<tr>
<td>pAS_BstTyrT(CUA)/sfGFP150TAG</td>
<td>this study</td>
<td>Addgene#174891</td>
</tr>
<tr>
<td>pAS_MmaPyIT/PyIRS AF</td>
<td>(Meineke et al. 2018)</td>
<td>Addgene#140023</td>
</tr>
<tr>
<td>pAS_BstTyrT(CUA)/sfGFP102TAG150TAA</td>
<td>this study</td>
<td>Addgene#174895</td>
</tr>
<tr>
<td>pAS_EcoLeuT(CUA)/AnapRS</td>
<td>this study</td>
<td>Addgene#140019</td>
</tr>
<tr>
<td>pAS_MmaPyIT(UUA)/GFP150TAA</td>
<td>this study</td>
<td>Addgene#174897</td>
</tr>
<tr>
<td>pAS_MmaPyIT(UCA)/GFP150TGA</td>
<td>this study</td>
<td>Addgene#174898</td>
</tr>
<tr>
<td>pAS_MmaPyIT(UCA)/PyIRS</td>
<td>this study</td>
<td>Addgene#174899</td>
</tr>
<tr>
<td>pAS_M15T(UUA)/MmaPyIRS</td>
<td>(Meineke et al. 2020)</td>
<td>Addgene#154774</td>
</tr>
<tr>
<td>pAS_MmaPyIT/CRFR1 95TAG</td>
<td>(Meineke et al. 2021)</td>
<td>Addgene#174900</td>
</tr>
<tr>
<td>pAS_MmaPyIT/LaG17-SynNotch442TAG</td>
<td>this study</td>
<td>N/A</td>
</tr>
<tr>
<td>pAS_M15PyIT(UUA)/LaG17-SynNotch39TAA277TAG</td>
<td>this study</td>
<td>N/A</td>
</tr>
</tbody>
</table>

**Software and Algorithms**

<table>
<thead>
<tr>
<th>Fiji Schindelin et al., 2012</th>
<th>(Schindelin et al. 2012)</th>
<th><a href="https://fiji.sc">https://fiji.sc</a>; RRID:SCR_002285</th>
</tr>
</thead>
<tbody>
<tr>
<td>FlowJoTM</td>
<td>BD Life Sciences</td>
<td>v10.6.2 Software</td>
</tr>
</tbody>
</table>
REFERENCES


to Figure 1

A

B

C

D

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]

mean GFP fluorescence [AU]
to Figure 2

A

pAS backbone

- pUC Ori
- AmpR

4x 75K TyrT<sup>CUA</sup>

4x 75K LeuT<sup>CUA</sup>

4x tRNA

Ef1<sup>＞</sup>

aaRS or reporter

IRES Res

P<sub>A</sub>

pAS

4x RNA/aaRS

4x 75K TyrT<sup>CUA</sup>

4x 75K LeuT<sup>CUA</sup>

4x tRNA

Ef1<sup>＞</sup>

aaRS or reporter

IRES Res

P<sub>A</sub>

pAS

4x RNA/GFP<sup>150TAG</sup>

pUC Ori

AmpR

B

Normalized to Mode

GFP fluorescence [AU]

clone 1

clone 3

C

Anap fluorescence [AU]

GFP fluorescence [AU]

HCT116

LeuT/AnapRS+GFP<sup>150TAG</sup>

-ncAA

+Anap 2h chase

+Anap no chase

D

Normalized to Mode

GFP fluorescence [AU]

Anap fluorescence [AU]

HCT116

LeuT/AnapRS+GFP<sup>150TAG</sup>

-ncAA

+Anap 2h chase
A. Mean and median GFP fluorescence for different clones and conditions.

B. Comparison of GFP fluorescence for different clones and conditions with and without CpK treatment.
to Figure 4

A

B

C

GFP fluorescence [AU]

mean GFP fluorescence [AU]

+ CpK + AzF

+ CpK + AzF

Normalized to Mode

Normalized to Mode

GFP fluorescence [AU]
to Figure 5

A

GOI Xpos*

internal suppression

4x Py/T

TAG

ncAA incorporation & labeling

POI

*GOI

STELLA

N-ter tag

4x Py/T

TAG

POI

B

Daxx48*

*Tub6

surface

lysate

250 kDa

150 kDa

75 kDa

50 kDa

α-FLAG

C

*cenpa

*nPACE30*

*pgk105

*cov ORF6

*covM

TCO*K

SIR

α-HA

PyIRS AF

D

Hoechst

metet-BDPFL

HA

merge

TCO*K

DaxxR48*

+TCO*K

-TCO*K

+TCO*K

-TCO*K

+TCO*K

-TCO*K

+TCO*K

-TCO*K

+TCO*K

-TCO*K

+TCO*K
Supplementary Figures

Figure S1, relates to Figure 1

A) Quantification of GFP fluorescence in lysates of low selected (2 μg/ml puromycin, 500 μg/ml blasticidin) stable integrant HEK293T cell line. Where indicated (yellow shading), the cell line was transiently transfected with different pAS plasmids bearing 4xPyIT and/or PylRS variants. For each condition, quadruplicate transfections were performed. For three of the four samples medium was supplemented with 0.2 mM CpK, all samples were harvested 24 h post transfection. GFP fluorescence signal is normalized to that of untransfected cells cultured in presence of 0.2 mM CpK for 24 h, background fluorescence (–CpK) values were subtracted. Error bars indicate standard deviation. Plasmids transfected by lanes: 1 no transfection, 2 control plasmid GFP102TAA133TGA150TAG, 3 PylRS (Addgene #154762), 4 4xPyIT (Addgene #140008), 5 4xPyIT/RS (Addgene #140009), 6 8xPyIT/RS (Addgene #140008), 7 4xM15T/PylRS (Addgene #174890), 8 4xA2-1T/PylRS (A2-1 (Kelemen et al. 2022)), 9 4xPyIT/IPYE_Mba/MmachPylRS ((Bryson et al. 2017)).

B) Quantification of mean GFP fluorescence in flow cytometry samples shown in Figure 1D.

C) Immunoblotting of amber suppression with PylRS variant AbKRS-chlIPYE (AbKRS). The chemical structure of N6-[2-(3-Methyl-3H-diazirin-3-yl)ethoxy]carbonyl]-L-lysine (AbK) is shown. HCT116 cells transiently transfected with PylT/AbKRS and PylT/GFP150TAG plasmids (transient, left), or polyclonal stable integrants with the same plasmids (integration, right) were grown without ncAA (–AbK) or presence of 0.5 mM AbK for 24 h. Immunostaining for GFP expression levels, FLAG-AbKRS and β-actin.

D) Flow cytometry of HCT116 cells transiently transfected with PylT/AbKRS and PylT/GFP150TAG plasmids (transient, top), or polyclonal stable integrants with the same plasmids (integration, bottom) were grown in absence of ncAA (black) or in presence of 0.5 mM AbK (green) for 24 h.

Figure S2, relates to Figure 2

A) Schematic representation of pAS constructs for AzF and Anap incorporation. AzFRS or AnapRS are expressed under EF1α promoter control. An IRES (internal ribosome entry site) followed by a neomycin resistance (NeoR) is positioned immediately downstream of the aARS. A cassette of four tandem copies of orthogonal tRNA (B. stearothermophilus TyrT_CUA (TyrT_CUA) or E. coli LeuT_CUA (LeuT_CUA) for 7SK PolIII driven expression is integrated upstream
of the EF1α promoter. We chose to add external 7SK PolIII promoters leading each tRNA, even for BstTyrT\textsuperscript{CUA}, which contains an internal B-box for PolIII transcription (Liu et al. 2007).

B) The entire 4xtRNA/aaRS cassette is flanked by insulator sequences (Ins) and 3’ and 5’ inverted repeats for genomic integration by PBT (pink triangle). Plasmids for GFP\textsuperscript{150TAG} amber suppression reporter with either TyrT\textsuperscript{CUA} or LeuT\textsuperscript{CUA} cassettes have the same architecture, but contain a blasticidin resistance gene (Bsd).

C) Flow cytometry of amber suppression clonal cell lines isolated from the AzFRS, cognate TyrT\textsuperscript{CUA} tRNA and GFP\textsuperscript{150TAG} reporter integrant polyclonal population shown in Figure 2. GFP fluorescence from the GFP\textsuperscript{150TAG} reporter was measured in the two clonal cell lines after 24 h without ncAA (dashed black line) or in presence of 0.5 mM AzF (green line).

D) Flow cytometry of amber suppression cell line generated with AnapRS, cognate LeuT\textsuperscript{CUA} tRNA and GFP\textsuperscript{150TAG} reporter compared to the parental HCT116 cells. Cells were grown in presence of 0.05 mM Anap for 24 h and either measured directly or after a 2 h chase with ncAA-free medium. The 2D plot for GFP and Anap fluorescence shows the wild type HCT116 populations without Anap (-ncAA, red), with Anap (+Anap, grey) and with Anap washed out (+Anap, chase, black). The LeuT\textsuperscript{CUA}/AnapRS+GFP\textsuperscript{150TAG} integrant stable cell line is shown in blue (-ncAA) and green (+Anap).

E) Same data as in C) presented as histogram plots separating GFP and Anap fluorescences.

F) Flow cytometry of amber suppression cell line generated with AnapRS, cognate LeuT\textsuperscript{CUA} tRNA and a GFP\textsuperscript{150TAG} reporter (green) compared to an integrant cell line with LeuT\textsuperscript{CUA}/AnapRS only (no reporter, blue). Cells were grown in presence of Anap for 24 h and measured after a 4 h chase with ncAA-free medium. The “no reporter” cell line retains Anap.

Figure S3, relates to Figure 3

A) Quantification of mean and median GFP fluorescence measured in Figure 3A.

B) Comparison of GFP fluorescence generated by amber, ochre or opal suppression in the dual suppression clones sorted from polyclonal pools of PylT/PylRS and GFP\textsuperscript{150} reporter HCT116 integrants generated by PB transgenesis. The clonal populations were grown in presence of 0.2 mM CpK for 24 h. The clones used in Figure 3A are highlighted in red. For comparison, shown in grey, a population grown from 100 cells is also shown as –ncAA control and approximation of the polyclonal pool.

Figure S5, relates to Figure 5
A) Comparison of GFP fluorescence generated by dual amber and ochre suppression in the dual suppression clones sorted from a common polyclonal pool of PylT^{UA}/PylRS ochre suppressor, TyrT^{CUA}/AzFRS amber suppressor pair and GFP^{102TAG150TA} reporter integrants generated by PB transgenesis. The clonal populations were grown in presence of 0.2 mM CpK and 0.5 mM AzF for 48 h.

B) Fluorescence microscopy images of dual suppressor clonal F6 population grown without ncAA supplement, with either CpK or AzF or both for 48 h. White scale bars indicate 250 μm.

C) Comparison of GFP fluorescence generated by dual amber and ochre suppression in the dual suppression cell polyclonal pool of PylT^{UA}/PylRS ochre suppressor, TyrT^{CUA}/AzFRS amber suppressor pair and GFP^{102TAG150TA} reporter, the derived F6 clonal population or HCT116 parental cells. The populations were grown in presence of 0.2 mM CpK and 0.5 mM AzF for 48 h before measuring GFP fluorescence by flow cytometry.

D) Time course of ncAA incorporation by the dual suppressor F6 clonal population, cells were grown without ncAAs or with 0.2 mM CpK and 0.5 mM AzF for 12 h, 24 h, 48 h or 72 h prior to measuring GFP fluorescence by flow cytometry.

Figure S5, relates to Figure 5

A) Scheme illustrating construct design and ncAA labeling approaches for internal TAG and STELLA amber suppression. An asterisk (*) is used to indicate position of the amber codon and ncAA incorporation site. The amber stop codon signaling for ncAA incorporation can either be placed within the coding sequence of a gene of interest (GOI), producing the protein of interest (POI) in an ncAA-dependent manner, alternatively the POI can be expressed with an N-terminal ubiquitin-tag separated by an amber codon, leading to POI with ncAA at the N-terminus after ubiquitin processing (STELLA N-terminal tag) (Lafranchi et al. 2020).

B) Amber suppression PB integrant cell lines with PylT/RS AF and either PylT/Daxx48TAG-HA (Daxx48*) or PylT/STELLA N-terminally TAGed Tub6-HA (*Tub6). Stable integrant populations recovered after 7 d selection were grown in presence of 0.1 mM TCO*K for 24 h and SPIEDAC-labeled on live cells with 4 μM tet488 or in lysate with 1 μM tet488 or 1 μM tetSiR. Lysate aliquots were separated by SDS-PAGE and fluorescence captured in-gel prior to membrane transfer and immunoblotting for the C-terminal HA-tag and the FLAG-PylRS AF.
C) In-gel fluorescence and immunoblotting for further PB integrant amber suppression cells lines generated in HCT116 cells with PylT/RS AF and STELLA N-terminally TAGed CENPA, pigbos, CovORF6 and CovM or murine CENPA Q30TAG (msCENPA30*). Stable integrant populations recovered after 7 d selection were grown in presence of 0.1 mM TCO*K for 24 h and SPIEDAC-labeled in lysate with 1 μM tetSiR. Lysate aliquots were separated by SDS-PAGE and in-gel fluorescence visualized prior to membrane transfer and western blot for the C-terminal HA-tag and the FLAG-PylIRS AF.

D) Fluorescence microscopy of amber suppressor HCT116 polyclonal pools encoding PylT/RS AF and a protein of interest: Daxx^{48TAG} or STELLA N-terminally tagged Tub6 or CovM. Cells were grown in presence of 0.1 mM TCO*K for 48 h. Cells were fixed with 4% PFA, permeabilized with 0.1% triton, blocked with 2% BSA and stained with 0.5 μM metetBDPFL, primary anti-HA and secondary 555-coupled antibodies. Prior to imaging on a Nikon Ti2, nuclei were counterstained with Hoechst 33342. White scale bars indicate 25 μm.

Figure S6, relates to Figure 6

A) Comparison of GFP fluorescence generated by dual amber and ochre suppression in the dual suppression cell polyclonal pool of M15^{UA}/PyIR ochre suppressor, TyrT^{CUA}/AzFRS amber suppressor pair and SynNotch^{204TAG442TAA}-GFP, the derived D8 clonal population or HCT116 parental cells. The populations were grown in presence of 0.2 mM CpK and 0.5 mM AzF for 72 h before measuring GFP fluorescence by flow cytometry.

B) Time course of ncAA incorporation by the dual suppressor D8 clonal population, cells were grown without ncAAs or with 0.2 mM CpK and 0.5 mM AzF for 12 h, 24 h, 48 h or 72 h prior to measuring GFP fluorescence by flow cytometry.

C) Fluorescence microscopy images of dual suppressor SynNotch^{204TAG442TAA}-GFP clonal D8 population grown without ncAA, with either CpK or AzF or both ncAAs for 48 h. GFP fluorescence (green) and Hoechst 33342 nuclear staining (blue) are shown.

D) Transient transfection of HEK293T cells with PylT/RS and M15^{UA}/SynNotch^{204TAA442TAG}-GFP plasmids (at 1:4 ratio) grown with and without 0.2 mM CpK. Live cells were surface labeled with 2 μM tet488 48 h post transfection. Aliquots of cell lysates were separated by SDS-PAGE, exposed at 460 nm (488) for in-gel fluorescence and immunostained for GFP and FLAG-RS after membrane transfer.

E) Live-cell imaging of SPIEDAC- and CuAAC-labeled HCT116-derived dual suppressor clonal population with M15^{UA}/RS, TyrT^{CUA}/AzFRS, and M15^{UA}/SynNotch^{204TAA442TAG}-GFP integrated by PB transposition. Cells were grown for 72 h in the absence (−ncAA) or the presence of
0.2 mM CpK and 0.5 mM AzF. Surface labeling by SPIEDAC with 2 μM tetCy3 and CuAAC with 2 μM alk647 was carried out prior to Hoechst 33342 counterstaining, fixation with PFA and imaging. White scale bars indicate 25 μm.

F) Conditions like E) but instead of CuAAC with alk647, AzF was labeled by SPAAC with 2 μM DBCO647. White scale bars indicate 25 μm.

G) SPIEDAC and SPAAC labeling on live HCT116 clonal SynNotch204CpK442AzF-GFP cells. grown in presence of 0.2 mM CpK and 0.5 mM AzF for 72 h. Cells were labeled with 2 μM tet488 and 2 μM DBCO647. Lysates were separated by SDS-PAGE, in-gel fluorescence at 460 nm (488) and 630 nm (647) and immunoblotting against the C-terminal GFP are shown.
Supplementary material and method

Noncanonical amino acids

N6-[[2-(3-Methyl-3H-diazirin-3-yl)ethoxy]carbonyl]-L-lysine (AbK, CAS: 1253643-88-7, SiChem, SC-8034) was prepared as described for the other ncAAs.

Intracellular SPIEDAC labeling and immunofluorescence microscopy

HCT116 PB integrant polyclonal pools with PyrT/PyrRS AF and PyrT/gene of interest (GOI) were seeded at 15 000 cells per well on poly-L-lysine coated 96 well imaging plates (Ibidi) in presence of 0.1 mM TCO*K for 48 h. Subsequently, the cells were washed with PBS, fixed in 4% PFA) for 10 min, permeabilized in PBS with 0.1% triton X-100 (v/v) (PBS-T), blocked in 2% BSA in PBS-T and stained with 0.5 µM methyltetrazine-BDP-FL (metetBDPFL) (Jena Bioscience) for 1 h at RT, incubated with primary anti-HA antibody (Santa Cruz, sc-7392) followed by incubation with secondary 555-coupled antibody (Life Technologies, A-31570) and 2 µM Hoechst 33342 (Life Technologies). The cells were washed and imaged in PBS on a Nikon Eclipse Ti2 inverted widefield microscope, using a 20x (0.75 NA) objective and filters sets for DAPI, AF488, Cy3 and Cy5 fluorescence.

Supplementary References


