Formation of Nitrogenase NifDK Tetramers in the Mitochondria of *Saccharomyces cerevisiae*

Stefan Burén,[†] Eric M. Young,[‡] Elizabeth A. Sweeny,[‡] Gema Lopez-Torrejón,[†] Marcel Veldhuizen,[†] Christopher A. Voigt,*,[‡] and Luis M. Rubio*,[†]

[†]Centro de Biotecnología y Genómica de Plantas, Universidad Politécnica de Madrid (UPM) - Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria (INIA), Campus Montegancedo UPM, 28223, Pozuelo de Alarcón, Madrid, Spain

[‡]Synthetic Biology Center, Department of Biological Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, United States

S *Supporting Information*

ABSTRACT: Transferring the prokaryotic enzyme nitrogenase into a eukaryotic host with the final aim of developing N_2 fixing cereal crops would revolutionize agricultural systems worldwide. Targeting it to mitochondria has potential advantages because of the organelle's high O_2 consumption and the presence of bacterial-type iron−sulfur cluster biosynthetic machinery. In this study, we constructed 96 strains of *Saccharomyces cerevisiae* in which transcriptional units comprising nine *Azotobacter vinelandii* *nif* genes (*nif HDKUSMBEN*) were integrated into the genome. Two combinatorial libraries of *nif* gene clusters were constructed:

a library of mitochondrial leading sequences consisting of 24 clusters within four subsets of *nif* gene expression strength, and an expression library of 72 clusters with fixed mitochondrial leading sequences and *nif* expression levels assigned according to factorial design. In total, 29 promoters and 18 terminators were combined to adjust *nif* gene expression levels. Expression and mitochondrial targeting was confirmed at the protein level as immunoblot analysis showed that Nif proteins could be efficiently accumulated in mitochondria. NifDK tetramer formation, an essential step of nitrogenase assembly, was experimentally proven both in cell-free extracts and in purified NifDK preparations. This work represents a first step toward obtaining functional nitrogenase in the mitochondria of a eukaryotic cell.

KEYWORDS: *NifDK, nif genes, mitochondria, nitrogen fixation, yeast, Azotobacter vinelandii*
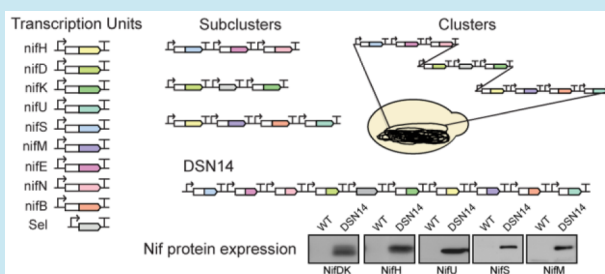
Nitrogen fixation, that is, the reduction of N_2 to NH_3, is a prokaryotic process catalyzed by a family of enzymes called nitrogenases, the most ecologically relevant and abundant one being the Mo-nitrogenase.[1] The Mo-nitrogenase is a two-component metalloenzyme consisting of a molybdenum−iron (MoFe) protein that catalyzes N_2 reduction and an iron (Fe) protein that acts as specific electron donor to the MoFe protein.[2] The Fe protein is a homodimer of the *nif H* gene product carrying a single $[Fe_4S_4]$ cluster between its subunits, whereas the MoFe protein is a heterotetramer of the *nif D* and *nif K* gene products carrying two pairs of a $[Fe_8S_7]$ cluster and a $[Fe_7S_9MoC\text{-}homocitrate]$, designated as the P-cluster and FeMo-co, respectively.[3,4] In addition to the nitrogenase structural genes, a number of accessory nitrogen fixation (*nif*) gene products are required for the maturation of the structural components to their catalytically active forms. These include proteins that interact in complex biosynthetic pathways for the assembly of nitrogenase metal clusters and their insertion into target apo-proteins.[5]

Engineering N_2-fixation activity could be achieved by transferring the appropriate bacterial *nif* genes to the plant.[6,7] However, the biochemical and genetic complexity of nitrogenase assembly and regulation poses a major challenge.

Although the *nif B*, *nif E*, *nif N*, and *nif H* gene products have been shown to carry out all essential biochemical reactions in FeMo-co biosynthesis *in vitro*,[8] the exact number of *nif* genes required for the genetic transfer of nitrogen fixation among bacterial species depends on the nature of both the nitrogenase source and the engineered host.[9−11] To add more complexity, Nif protein stoichiometry has been shown to be dynamic and strongly regulated.[12]

The plant cell offers several possibilities for *nif* gene expression due to its three genomes and multiple compartments. As Nif proteins are bacterial proteins, it can be envisioned that plant cell compartments of prokaryotic origin (plastids and mitochondria) would be superior for accumulation of active Nif proteins, each of the organelles offering its pros and cons.[7] Plastids are favorable because of the highly efficient homologous DNA recombination and the support for prokaryotic-type transcription and translation machineries, allowing multiple genes to be expressed from operons.[13] However, the extreme sensitivity of nitrogenase to O_2, which

destroys its metal clusters, poses a serious obstacle, particularly in chloroplasts where O_2 is generated during photosynthesis. This problem has been highlighted in a recent work, in which the Fe protein from transplastomic *Nicotiana tabacum* leaf extracts showed some *in vitro* activity only when plants had been previously incubated at subambient O_2 levels.[14] Therefore, mitochondria could offer an advantage due to the respiratory high O_2 consumption. Mitochondria also harbors a [Fe−S] cluster assembly machinery highly similar to the NifUS system,[15] in addition to the plentiful ATP and reducing power that is generated by respiration.[7] A minimum of 16 molecules of ATP is required to reduce one molecule of N_2. ATP is required for electron transfer from the NifH [Fe_4S_4] cluster to the P-cluster of the NifDK component during N_2 reduction, and possibly also for P-cluster maturation and FeMo-co synthesis.[2,5] While plastid proteins can be expressed either via transformation of the plastid genome, or imported post-translationally following nuclear expression, transformation of the mitochondrial genome is more difficult.[16]

Recently, it has been shown that mitochondria of aerobically grown recombinant *Saccharomyces cerevisiae* cells expressing *nif H* and *nif M* accumulate active Fe protein.[17] Yeast cells are convenient hosts to test eukaryotic expression of Nif proteins as they can be grown at varied O_2 levels, to reduce the negative effects from O_2 on Nif protein maturation. They are also easy to transform and maintain, enabling larger and more complex libraries to be constructed and analyzed. Its mitochondrial bacterial-like [Fe−S] cluster assembly machinery is also the best understood among eukaryotes.[18] Finally, yeast cultures can easily and economically be scaled up, which is required to generate sufficient biomass for protein purification and subsequent activity assays to prove Nif protein functionality. In this regard, *S. cerevisiae* will serve as an important eukaryotic model system when learning to engineer (more important) eukaryotic organisms, such as plants.

The entire prokaryotic *nif* gene cluster cannot be directly transferred to yeast due to differences in the molecular biology of gene expression. Thus, refactoring the *nif* cluster and generation of *S. cerevisiae* gene expression parts is critical.[19,20] Refactoring is an engineering approach to clean up genetic systems.[21,22] There are many examples of characterized yeast promoter[23−26] and terminator collections.[27,28] However, how much the promoter or terminator contributes to final gene expression strength, or whether there are emergent interactions with particular combinations of parts, have not been explored. Furthermore, models of gene expression in yeast, and eukaryotes in general, lag behind those used to construct powerful design tools in prokaryotes.[29] Large genetic systems in yeast have been constructed by repeating both the promoter and terminator across transcription units.[30] One approach to reducing part reuse is placing the 2A viral self-cleaving peptide between open reading frames, enabling multicistronic expression in yeast.[31] However, a monocistronic gene expression approach, where expression of each gene is independent from the others, is best suited to combinatorial design, assembly, and optimization. Thus, more yeast gene expression parts are needed to build stable constructs of varying expression strength.

To this end, a set of over 1200 transcription units composed of varying promoters and terminators were individually characterized for GFP expression in yeast (Eric M. Young, Johannes A. Roubos, D. Benjamin Gordon, and Christopher A. Voigt. *Composability and design of Saccharomyces cerevisiae gene expression* 2017, in preparation). From this set, part combinations can be selected to achieve varying expression levels, while preventing part reuse. This selection can be done manually or algorithmically with Double Dutch,[32] to aim for a suitable stoichiometric ratio of *nif* protein production but vary across the dynamic range of monocistronic gene expression strength in yeast.

In this study, we aimed to simultaneously optimize protein expression and mitochondria targeting for a 9-gene *nif* cluster (*nif HDKUSMBEN*) from the model diazotroph *Azotobacter vinelandii* in the yeast *S. cerevisiae*.[7,33] *A. vinelandii* was chosen as source of *nif* sequences due to its strictly aerobic life style, in which nitrogenase is protected from O_2 damage mainly by maintaining high respiratory rates. To refactor the 9 *A. vinelandii nif* genes, each gene was codon optimized for expression in *S. cerevisiae*. With expression ratio intuition derived from native expression of the cluster in *A. vinelandii*,[12] and little *a priori* knowledge of appropriate absolute expression levels or mitochondrial tags, combinatorial design and assembly was chosen in order to optimize the cluster in yeast.[20] Gene products were targeted to the yeast mitochondria using three different mitochondria leading sequences. These two variables were simultaneously optimized in the combinatorial design. In total, 96 strains of *S. cerevisiae* were generated with *nif HDKUSMBEN* genes integrated into the yeast genome. Expression and mitochondria targeting of properly processed Nif proteins was confirmed at the protein level and NifDK tetramer formation, an essential step of nitrogenase assembly, was experimentally proven. However, the purified NifDK protein was inactive, and did not appear competent for activation by FeMo-co insertion. To our knowledge, this is the first attempt undertaken in a eukaryotic cell to express sufficient *nif* genes necessary for a functional nitrogenase.

■ RESULTS AND DISCUSSION

Library Strategy and Design. Ninety-six strains of *S. cerevisiae* were constructed with genomic integration of yeast-optimized coding sequences for mitochondria-targeted versions of the *A. vinelandii nif HDKUSMBEN* genes. As stoichiometric levels of Nif proteins are essential for maturation and functionality,[12,19] 29 different promoters were combined with 18 distinct terminators to adjust the expression level of each *nif* gene. Similarly, combinations of three distinct mitochondria targeting signals were used to ascertain efficient Nif protein targeting and leader sequence processing upon mitochondria import (Supporting Information Tables S1−S3). In this study, DNA sequences encoding a duplicated mitochondria targeting signal for the *S. cerevisiae* Mn-superoxide dismutase[34] (here referred to as SOD2), as well as the leader sequences for subunit 9 of the *Neurospora crassa* F_0-ATPase[35] (SU9) or the *Arabidopsis thaliana* INDH1[36] (INDH) were fused to the 5′ end of *nif* genes. SOD2 and SU9 have been already shown to successfully deliver Nif proteins to yeast mitochondrial matrix.[17]

Two different combinatorial libraries of nitrogen fixation clusters were constructed. The purpose of the first was to determine the appropriate mitochondrial tags for *nif E*, *nif N*, *nif S*, *nif H*, and *nif M*. The purpose of the second was to vary expression, testing previous assumptions of gene expression ratios. To achieve the desired expression for each set, promoter and terminator combinations previously characterized by GFP expression were used. The sequences for each part are listed in
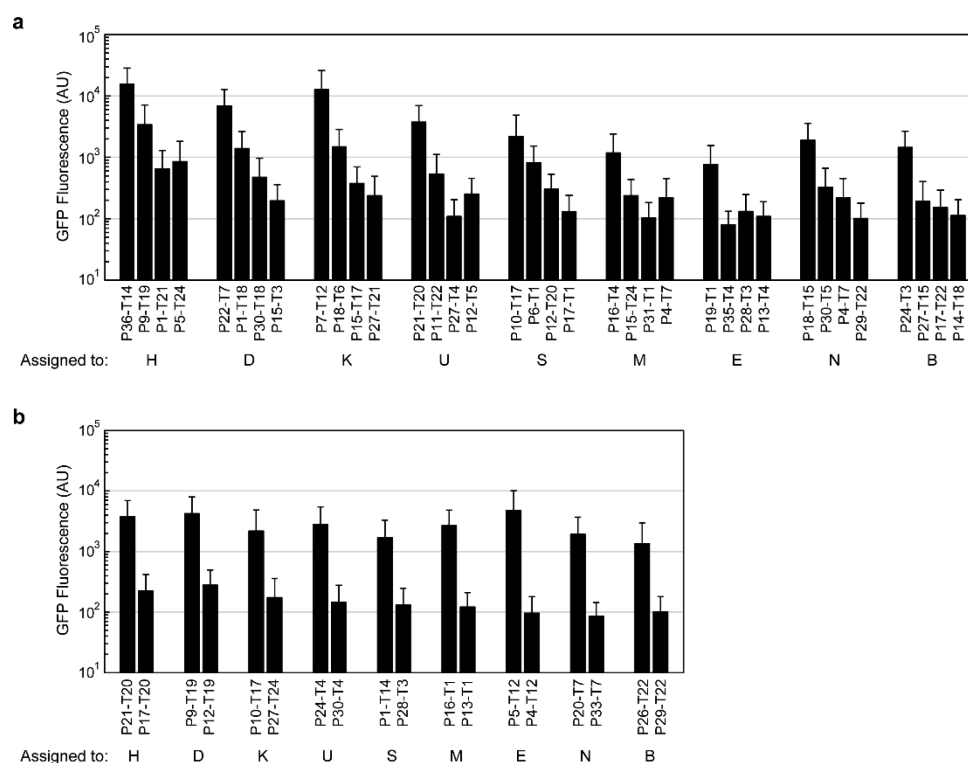
Figure 1. Measured strengths of transcription units. The geometric mean and standard deviation of biological replicates is plotted in arbitrary units for each promoter-terminator combination. (a) The four expression levels for each of the nine genes in the designed set. (b) The two expression levels for each of the nine genes in the factorial set. The targeted ratios, and how close each promoter-terminator pair approximates that ratio with GFP is shown in Table S2.

Table S1 and the expression strength of each combination is listed in Table S2.

The mitochondrial tag combinatorial library consisted of 24 clusters (strains DSN1−24), hereafter referred to as the "designed set". The designed set varied mitochondrial tags within four subsets of expression strength. Within each subset, the mitochondrial tags are varied in six combinations (*nif E*, *nif N*, *nif B* tagged with SU9 or INDH combined with *nif H* and *nif M* tagged with SU9, SOD2, or INDH, all others tagged with SU9). Based on previous work with the *A. vinelandii nif* cluster,[12] all genes were targeted to be expressed at a constant ratio. Relative to *nif N*, the expression of *nif H* was targeted to be a factor of 10 higher, *nif D* and *nif K* 6 times higher, *nif U* and *nif S* 2 times higher, and all others equal to *nif N*. Expression was set "very high", "high", "medium", and "low" for each subset. The expression strength of each selected promoter and terminator combination with GFP is depicted in Figure 1a.

The expression combinatorial library consisted of 72 clusters (strains DOE1−72), hereafter referred to as the "factorial set." The factorial set fixed mitochondrial tags (*nif E*, *nif N*, *nif B* tagged with INDH, *nif H* and *nif M* tagged with SOD2, all others tagged with SU9) and assigned each gene one of two expression levels according to a factorial design output by the software package JMP. High and low expressing combinations of promoters and terminators were chosen. The expression strength of these combinations, tested with GFP, is depicted in Figure 1b.
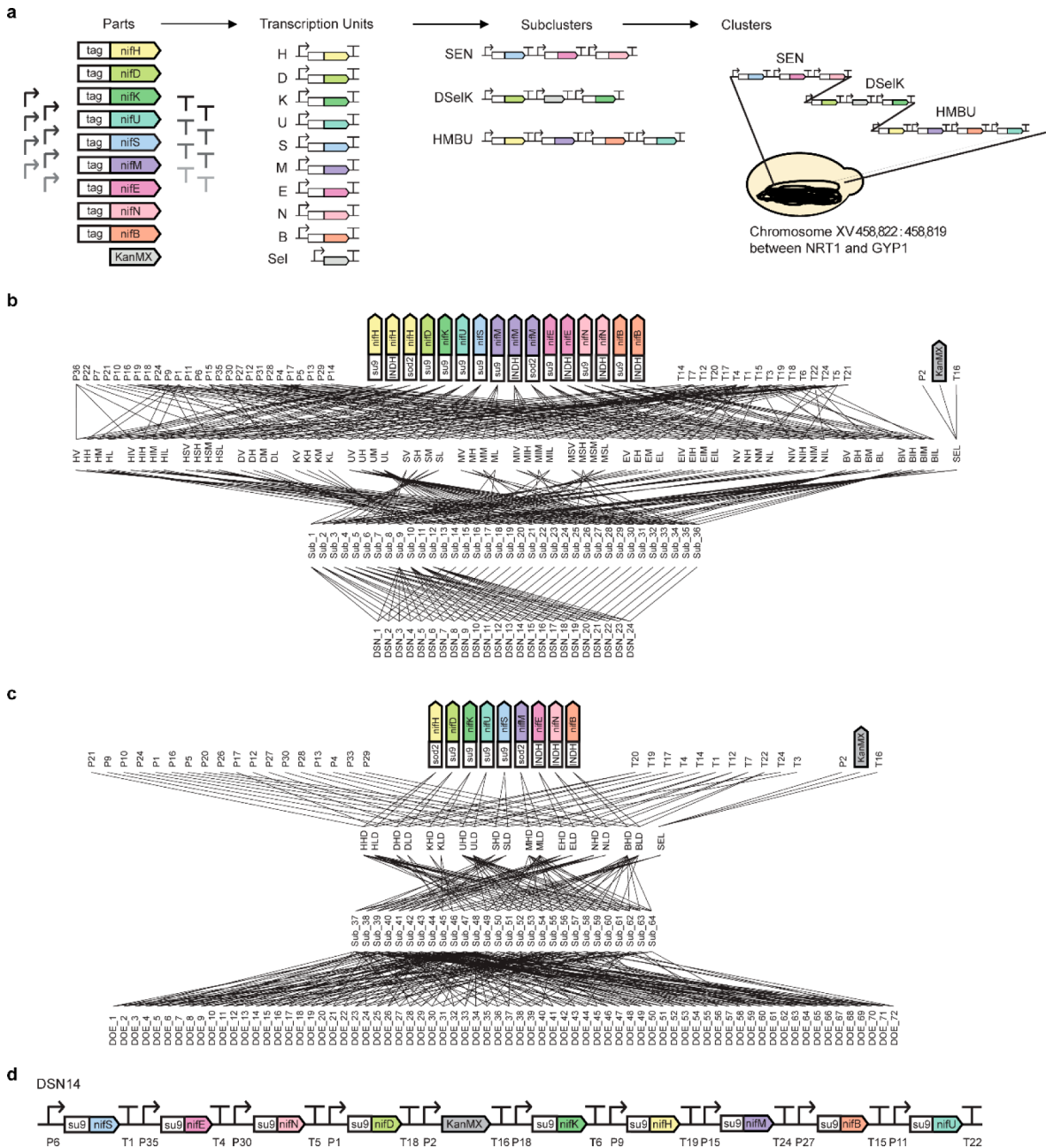
Library Assembly. A hierarchical assembly strategy based on TypeIIS cloning was used to construct all clusters.[20,37] A simplified version of this strategy is depicted in Figure 2a, with a more detailed version described in the Methods section and Supporting Information. Clusters were built up from a set of tagged genes, an antibiotic resistance gene for selection, and transcriptional parts. The parts were combined into transcription units via TypeIIS assembly. The transcription unit destination vectors defined gene order and allowed assembly into three subclusters, SEN, DSelK, and HMBU. These subclusters contained homologous sequences to one another and to the *S. cerevisiae* genome, and could therefore be assembled into full clusters by homologous recombination during transformation. All assembly steps for the designed set are depicted by the assembly tree in Figure 2b, while the assembly tree for the factorial set is shown in Figure 2c. An example of final *nif* gene organization is shown from DSN14 (Figure 2d). Each assembly step is also tabulated in Table S3.

Following gene assembly, a subset of clusters was selected for genomic DNA isolation and verification of cluster assembly by polymerase chain reaction (PCR). A further subset was confirmed by Sanger sequencing of the PCR products (Table S4). All clusters tested were correct, with the exception of two clusters within the designed set that were determined to have a newly discovered bacterial transposon inserted into *nif M* during the cloning process. This transposon was observed to only affect those transcription units which contained *nif M* expressed with P4 and T7, therefore the two clusters containing this transcription unit (DSN1 and DSN4) were omitted from further analysis yielding a final number of 94 yeast clones.

Yeast Nif Protein Expression. To confirm protein expression, total protein extracts were prepared from aerobically grown yeast cultures and analyzed by immunoblot using antibodies specifically targeting each Nif protein. Levels of Nif proteins in each strain were given a score of 0 if the protein was undetected, 1 if detected, and 2 if detected and the band was more pronounced, qualitatively indicating more protein being

Figure 2. Hierarchal assembly. (a) Assembly strategy for transcription units, subclusters and full clusters inserted by homologous recombination in the genome of *S. cerevisiae*. The standard KanMX selection cassette, abbreviated as "Sel," was used to select for G418 resistance as a result of successful integration. (b and c) Assembly tree of the designed (b) and factorial (c) set, where each line represents an assembly step (see Methods section and Supporting Information for further details). Note the variants of the tags on each gene in the designed set (b), while the tags are unchanged for each gene in the factorial set (c). (d) Scheme of *nif* gene organization in DSN14.

present (Table S4). Using expression and protein migration data from these 94 strains, efficiency of mitochondria targeting signals and promoters for each Nif protein could be assessed (Figure 3).

The three mitochondrial targeting signals tested here successfully expressed Nif proteins. One striking exception was INDH, which when fused to *nifH* did not produce detectable NifH protein levels in any of the eight clones tested, independent of the promoter being used. In contrast, with the use of SU9 or SOD2 tags NifH accumulated in most constructs (Figure 3a−c and Table S5). Although sufficient separation of NifE and NifN was not possible by SDS-PAGE, as NifE and NifN in each clone were expressed with the same mitochondria

targeting signal, the NifEN immunoblot result could be included in the mitochondria signal analysis. However, as different promoters were used for *nifE* and *nifN*, they were excluded from the promoter analysis (Supporting Information Figure S1 and Table S6).

Evaluation of promoter sequences that successfully produced detectable Nif proteins lead to the following general conclusions (Figure 3d−j and Table S6). Promoter 1 (from actin gene) rendered good expression levels, as demonstrated by results obtained using NifD, NifH, and NifS antibodies. In other cases, expression of proteins using the same promoter sequence gave strikingly different results depending on the gene. One example is promoter 24 (from the glyceraldehyde-3-
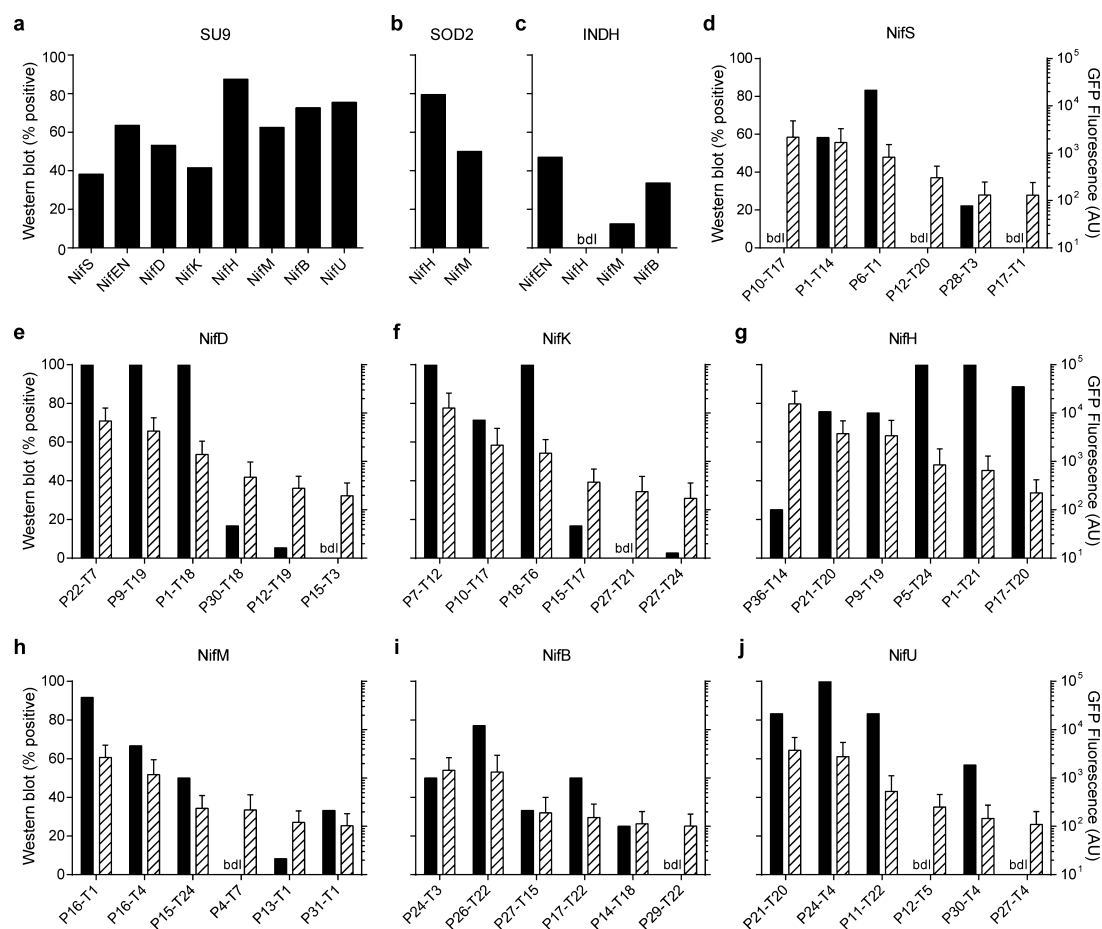
Figure 3. Analysis of promoters/terminators and mitochondria targeting signals. Efficiency of each mitochondria targeting signal (a−c) and promoter (d−j) in generating detectable Nif proteins. As no NifH protein was detected when tagged with INDH (a−c), INDH-NifH was excluded in the promoter analysis (d−j). Solid bars indicate % positive clones as detected by Western blotting using antibodies targeting Nif proteins (a−j, left axis). Striped bars indicate observed fluorescence when GFP was expressed by corresponding transcription unit, with the geometric mean and standard deviation of biological replicates plotted in arbitrary units for each promoter−terminator combination (d−j, right axis, see also Figure 1). Data in (d−j) is ordered according to GFP fluorescence levels. bdl, below detection limit.

phosphate dehydrogenase gene) that resulted in NifU protein in all 35 strains used, but only in half of the six clones where it was used for NifB expression. Another example is promoter 17 (from telomerase inhibitor gene) that generated detectable expression of NifH in nearly 90% (31/35) of the strains, but only NifB protein in half (3/6) of the strains and no NifS protein (0/4). These four strains lacking NifS expression had SU9 as the mitochondria tag, suggesting that promoter 17 might fail when combined with this N-terminal sequence. However, in the case of NifB, the three positive clones expressing detectable protein all contained SU9, indicating that the combination of promoter and leading sequence (P17+SU9) per se was likely not the reason for nondetectable NifS. In addition, SU9 sequence fused to *nifS* showed accumulation of NifS protein in five out of six strains when controlled by promoter 6 (from elongation factor 2 gene), further demonstrating that promoter 17 + SU9 was not the sole reason for the failure of these strains to accumulate NifS.

Regression Analysis. To analyze the effect of gene tag and expression level on detection of the protein in the immunoblot, least-squares regression was performed using the statsmodels Python library. Bands corresponding to Nif proteins in individual immunoblots were scored by no detection (0), detection (1), and strong detection (2). Clusters were then

scored by summing the immunoblot scores for each protein. Thus, a cluster could have a maximum score of 18 (9 genes, each with a maximum score of 2) and a minimum score of 0. Regression analysis of cluster score against all factors within the cluster was performed. The models used for regression, the associated R^2 values, and coefficients for selected fit equations are listed in Table S7. The python script used to perform all regression analysis is provided in the Supporting Information.

In general, transcription unit strength and mitochondrial tags were not predictive of detection by immunoblotting at the cluster level ($R^2 = 0.59$ for the designed set and $R^2 = 0.50$ for the factorial set). This implies that other factors, such as misassembly, mutation, and antibody sensitivity determine how many genes in a cluster were detected. Since protein level rather than transcription was assessed, we cannot eliminate the possibility of instability of the mRNA transcripts or the nascent proteins, as there appears to be one or more truncation products of Nif proteins (discussed below). The fit equation for the designed set (Table S7) shows that the SU9 tag, the expression strength, and the interaction terms between the SU9 tag and the expression strength all have the largest positive coefficients, indicating that those factors are positively correlated with detection by immunoblotting.
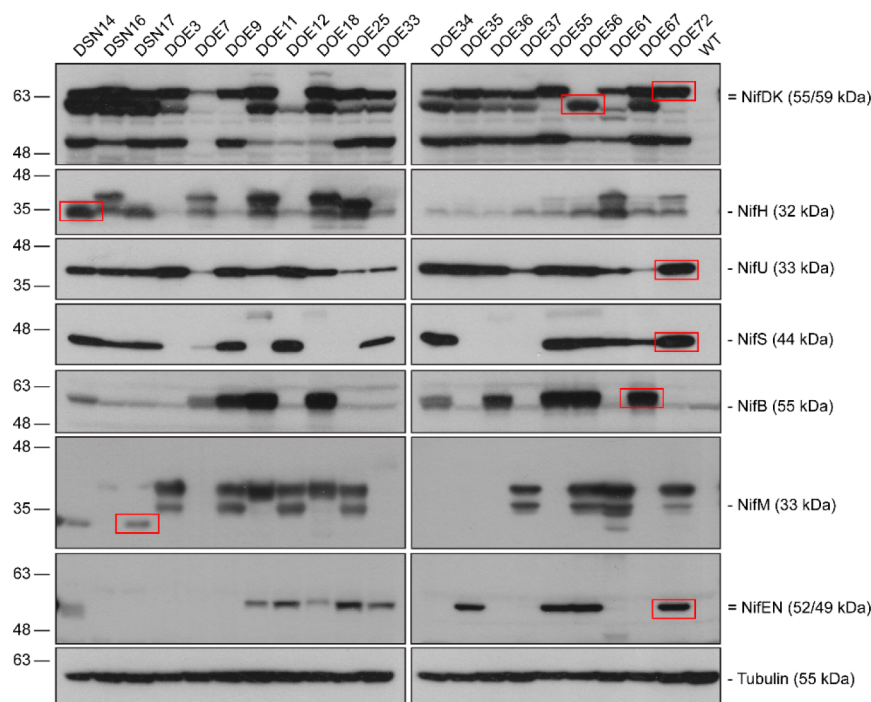
Figure 4. Nif protein expression in 20 selected yeast clones. Immunoblot analysis of 20 transformed yeast clones, showing expression and migration of NifDK, NifH, NifU, NifS, NifB, NifM, and NifEN proteins, as well as tubulin (loading control). Theoretical masses of corresponding *A. vinelandii* Nif proteins are indicated. Wild-type (WT) yeast clone is included as control of Nif antibodies specificity. Red squares highlight migration at the location of corresponding *A. vinelandii* proteins.

Ordinary least-squares regressions were also performed at the individual gene, not cluster level. Variability in the quality of fit was observed. NifS detection was uncorrelated to its expression strength ($R^2 \sim 0.16$ in the factorial set and ~ 0.08 in the designed set), while NifM, NifB, and NifU also showed poor fits (R^2 values between 0.3 and 0.6 in both sets). However, NifD detection better correlated with its expression strength ($R^2 = 0.77$ in the factorial set and 0.71 in the designed set). NifH detection was uncorrelated with expression in the factorial set ($R^2 = 0.14$), but better correlated in the designed set, where the mitochondrial tag was varied ($R^2 = 0.65$). The INDH tag was used for NifH in the designed set, which did not produce detectable protein. For NifH in the designed set, NifH detection was negatively correlated with the INDH tag (coefficient = −0.69), and positively correlated with both SOD2 and SU9 tags (coefficients of 0.81 and 0.56, respectively). The cluster with detectable expression of all nine Nif proteins by Western blotting was the designed set cluster DSN14 (Figure 4 and Table S4). This cluster also showed strong immunoblot detection of seven of the nine genes, giving a cluster score of 16 of a possible 18. This cluster contains all genes tagged with SU9 and with expression set at the high level (Figure 2d). This is in agreement with the conclusions of the regression analysis.

Nif Protein Processing. On the basis of the initial immunoblot-based screening, the 20 most promising strains were selected for further analysis. These strains were grown again and independently analyzed to verify protein expression levels (Figure 4). To ensure that Nif proteins were properly translated and processed, their migration on SDS-PAGE gels were compared to the migration of corresponding protein from *A. vinelandii* total protein extracts, or purified *A. vinelandii* Nif proteins.[12] Difference in migration could indicate faulty translation start, improper processing of mitochondria targeting

sequences or Nif protein instability/protein degradation. In this regard, a prominent and faster-migrating polypeptide around 50 kDa was recognized by the NifDK antibodies in many yeast strains, in addition to the expected NifDK double band. This could indicate that NifD and/or NifK are prone to degradation when expressed in a eukaryotic host. As NifD and NifK forms a tetramer consisting of two NifD and NifK subunits, insertion of this aberrant variant into NifDK could prevent NifDK functionality. Another possibility is that this variant appears when NifDK is not properly assembled, and that the putative degradation product is indicative of failing NifDK maturation.

To understand the origin of the faster migrating polypeptide, purified *A. vinelandii* NifDK protein was separated by SDS-PAGE and immunoblotted using antibodies raised specifically against NifD or NifK (Figure S2a). The result showed slower NifD migration, although the theoretical size of NifD is several kDa below that of NifK (55.3 and 59.4 kDa, respectively). In this regard, the presence of the 50 kDa band in the yeast protein extracts coincided with the upper and slower migrating NifD (or NifD together with NifK), compared to clones that only showed NifK expression, confirming an NifD origin of the 50 kDa band (Figure S2b). This was further supported using NifD and NifK specific immunoblots of yeast protein extracts (Figure S2c), in which the NifD antibodies detected a band around 50 kDa. Interestingly, expression of mitochondrial targeted N-terminally His-tagged NifD (SU9-His-NifD) with NifK (SU9-NifK) from plasmids using galactose induced promoters showed that full-length NifD was recognized by the His-tag antibody, while the 50 kDa form was not, strongly suggesting that the 50 kDa variant is originating from N-terminal processing of NifD (Figure S2d). A similar faster migrating NifD isoform is observed in *A. vinelandii* cells upon deletion of the *nifK* gene (unpublished data), indicating that

indeed this NifD form is produced upon incomplete NifDK assembly.

Targeting and Processing of Mitochondrial Signals. As most Nif proteins form homodimers or complexes with other Nif proteins, correct N- and C-termini can be critical for Nif functionality. Our analysis showed that SU9 mostly rendered polypeptides of correct size, while, for example, SOD2 often resulted in two distinct bands, where the slower migrating form likely corresponded to non (or partly)-processed SOD2 fusion proteins (e.g., compare migration of NifH in DSN14 and DSN17 (SU9-NifH) with DSN16, DOE7, and DOE11 (SOD2-NifH) in Figure 4). In other cases, SOD2 generated two polypeptides that migrated slower than the expected *A. vinelandii* protein, and as seen using SU9 signal (e.g., compare DSN17 (SU9-NifM) with DOE3 (SOD2-NifM) in Figure 4). While INDH failed to generate detectable protein for most NifH and NifM clones Figure 3c, it was efficient in generating NifB of apparently correct size (Figure 4).

Targeting of Nif proteins using the three leader sequences was confirmed by isolating enriched mitochondrial fractions from selected strains (Figure S3). As expected, mitochondria targeting was achieved using all three mitochondria targeting signals (SU9, SOD2, and INDH). Importantly, the lower NifD isoform was enriched in mitochondria isolations, indicating its appearance was not due to faulty translation initiation, but that processing/degradation appears to take place in the organelle itself. In line with results from the *A. vinelandii nifK* mutant mentioned above, this finding could indicate that excessive NifD processing is initiated by inefficient or absent NifDK tetramer formation.

Yeast NifDK Formation. Maturation of NifDK tetramer depends on the action of the *nifUSHMDK* gene products. To test NifDK complex formation, protein extract prepared under anoxic conditions were obtained from yeast strain DSN14. DSN14 was selected due to good expression and accurate mitochondrial targeting and processing of NifU, NifS, NifH, NifM, NifD, and NifK (Figure 5 and Figure S4). Two individual strategies were used to test NifDK assembly. First, proteins were separated on anoxic native gels, where eventual NifDK complexes would migrate intact. Immunoblot analysis showed that polypeptides cross-reacting with NifDK, as well as the NifD and NifK specific antibodies, comigrated (Figure 6a). As NifDK is known to migrate differently depending on the clusters present[38−40] (e.g., FeMo-co present at holo-NifDK increases its migration compared to the P-cluster containing, but FeMo-co-less, apo-NifDK), the purified *A. vinelandii* holo- and apo-NifDK proteins were included as comparisons. Yeast NifD and NifK migrated closely, but not identical to any of the two *A. vinelandii* NifDK forms, indicating that some conformational or biochemical discrepancies exist between the proteins expressed in the two hosts and suggesting that some of the critical components of NifDK could be missing in the yeast expressed protein.

As a second and independent approach to verify NifD and NifK direct interaction, an inducible yeast expression vector was created where His-tagged NifD and nontagged NifK, both with SU9 leader sequences, were expressed using the galactose inducible galactokinase 1 and 10 (GAL1 and GAL10) promoters. Anaerobic Co^{2+} affinity chromatography of protein extracts from DSN14 cells, or wild-type yeast cells complemented with galactose inducible expression of NifUSHM,[17] transformed with the His-NifD/NifK expression vector and grown with galactose as the carbon source, showed that NifK
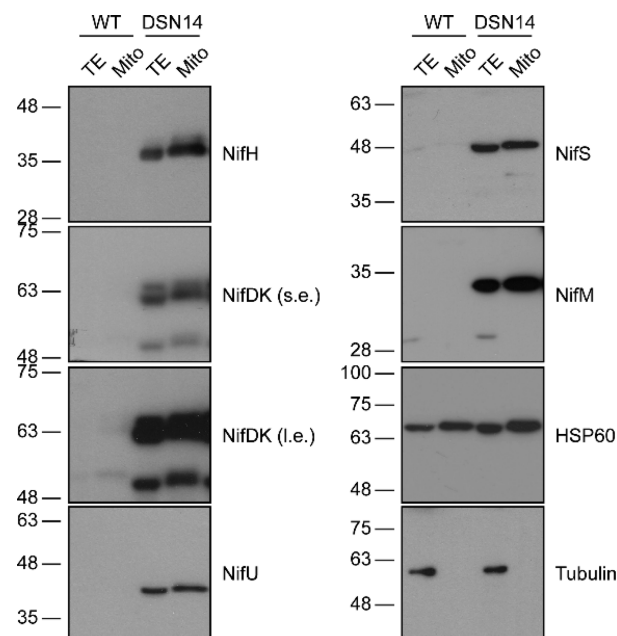

Figure 5. Mitochondria targeting of Nif proteins in DSN14. Immunoblot analysis of total extracts (TE) and mitochondria isolations (Mito) from wild-type yeast (WT) and DSN14, where Nif proteins are expressed and targeted to mitochondria using SU9 leader sequences. Analysis using antibodies recognizing cytoplasmic (tubulin) and mitochondria (HSP60) control proteins is included. s.e. and l.e., short and long exposure.

copurified with His-NifD and further suggested that NifDK complex is indeed formed (Figure 6b and Figure S5). In addition, the faster migrating N-terminally processed isoform of NifD (lacking His-tag) was copurified with full-length His-NifD, suggesting that tetramers of NifDK, and not only eventual NifDK dimers, are formed.

NafY Binding to Yeast NifDK. NafY protein is shown to bind both the iron−molybdenum cofactor (FeMo-co) and to FeMo-co-less apo-NifDK.[41] Interaction between NafY and apo-NifDK can be assessed using anoxic gel electrophoresis, where NifDK-bound NafY shows retarded migration with respect to free NafY.[41] NafY incubated with pure *A. vinelandii* apo-NifDK comigrated with apo-NifDK as expected (Figure 6c), whereas binding to holo-NifDK was significantly weaker. The majority of NafY, when incubated with holo-NifDK, showed faster migration compared to NafY alone. This is likely to result from binding of NafY to holo-NifDK-derived FeMo-co,[41−43] as FeMo-co-bound NafY has been shown to migrate faster on anoxic native gels.[44] While most of NafY incubated with anaerobic DSN14 cell-free extract migrated as in the control sample, a small population of NafY polypeptides showed retarded migration at the site of yeast NifDK. This result suggests some affinity of NafY for yeast NifDK, supporting that NifDK in DSN14 is partly matured. However, no difference in yeast NifDK migration was observed upon incubation with purified FeMo-co, and no NifDK activity (assessed by acetylene reduction[8,45]) was measured following FeMo-co insertion (data not shown), indicating that the yeast expressed protein was not readily activated by FeMo-co. Therefore, to what extent the formed NifDK tetramer is matured in *S. cerevisiae* and what factor(s) are missing remains to be investigated.

Summary. We generated and analyzed 94 strains of *S. cerevisiae* with nine *A. vinelandii nif* genes (*nifHDKUSMBEN*)
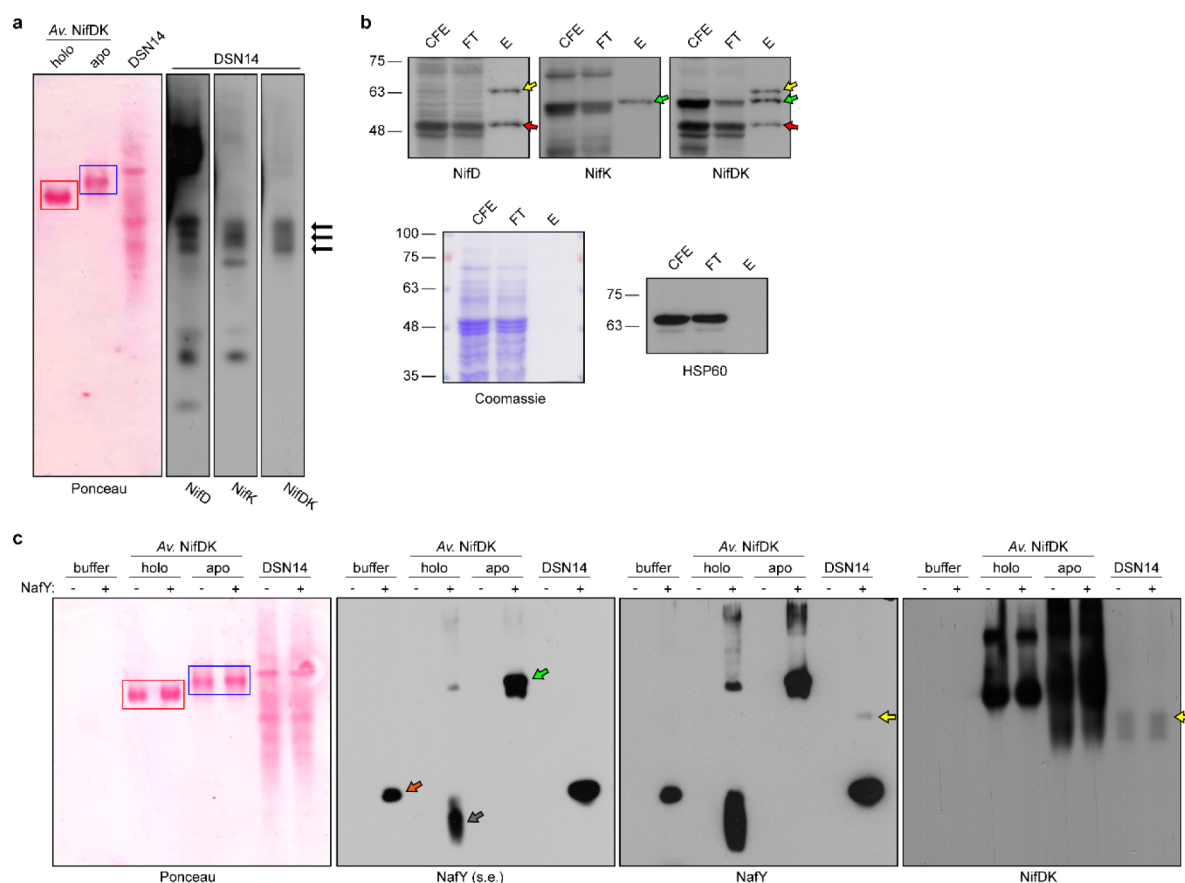
Figure 6. NifDK tetramer formation. (a) Ponceau and immunoblot analysis of yeast DSN14 protein extract, as well as purified *A. vinelandii* holo- and apo-NifDK proteins (red and blue squares, respectively), separated on anoxic native gels. Polypeptides cross-reacting with NifD, NifK, and NifDK antibodies are highlighted with black arrows. (b) Co-purification of NifK with His-tagged NifD. NifK (green arrows) comigrates with His-tagged NifD (yellow arrows), indication complex formation of NifK and His-NifD polypeptides. In addition, the faster migrating species, presumably N-terminally processed form of NifD (red arrows) is copurifying with His-tagged NifD, suggesting that NifDK tetramer is formed. (c) Ponceau and immunoblot analysis of yeast DSN14 protein extract, as well as purified *A. vinelandii* holo- and apo-NifDK proteins (red and blue squares, respectively), separated on anoxic native gels in the absence or presence of NafY. Co-migration of NafY protein with apo-NifDK is indicated by green arrow, as compared to unbound NafY (brown arrow). Faster migrating, presumably FeMo-co bound NafY, is indicated by gray arrow. A population of NafY protein comigrating with yeast NifDK is indicated by yellow arrow. s.e., short exposure.

integrated into the yeast genome. Gene products were targeted to the yeast mitochondria using three different mitochondria leading sequences, and their expression was regulated using 29 distinct promoter and 18 terminator sequences aiming for a suitable stoichiometric ratio of Nif protein production. Successful expression, targeting and mitochondrial signal trimming was confirmed in aerobically grown yeast cultures by immunoblotting with Nif targeting antibodies. This work elucidated SU9 as a promising mitochondrial tag to use in future *nif* cluster assemblies, and that efficient Nif protein expression in eukaryotic cell is possible.

Accurate NifDK protein assembly is of absolute necessity for future attempts to generate eukaryotic organisms (e.g., plants) able to fix atmospheric dinitrogen, and to test the possibilities of eukaryotic hosts to perform this activity is required. We therefore determined efficiency of Nif protein production and maturation in yeast protein extracts, with a special emphasis on NifDK. Apo-NifDK is defined as a P-cluster containing, but FeMo-co-less, NifDK protein consisting of two NifD and NifK subunits, respectively. Formation of apo-NifDK depends on the activity of NifSUHM in addition to NifDK, where NifS and NifU are involved in synthesis and delivery of [Fe−S] clusters to NifDK (and NifH), and NifH is necessary for the maturation

of the P-cluster in apo-NifDK.[40] NifM is important for maturation of active NifH,[46−48] although its mechanism is not completely clarified.

Our results show that NifS, NifU, NifH, NifM, NifE, NifN, and NifB proteins could be efficiently expressed and accumulated at mitochondria. We could also detect formation of NifDK protein, indicating that the two subunits are capable of forming a complex upon import into yeast mitochondria. When coexpressed with NifHUSM, yeast NifDK showed some affinity toward NafY, indicative of the presence of P-clusters (or P-cluster precursor). However, formed apo-NifDK was not readily activated by FeMo-co *in vitro*, suggesting that coexpression of additional *nif* genes might be necessary for efficient NifDK formation to take place in the eukaryotic cell. Proper *nif* gene expression dynamics could also be an important factor for NifDK formation. *A. vinelandii* cells entering nitrogen-fixing conditions show a peak of *nif U* and *nif S* expression preceding *nif HDK* expression, probably to ensure the supply of Fe−S clusters to newly formed nitrogenase components.[12] It is possible that engineering nitrogenase in nondiazotrophs would require the incorporation of promoters adding temporal control to *nif* gene expression.[49] Further work will be needed to elucidate the biochemical properties of this

protein and to understand the level of NifDK maturation achieved.

This study highlights the strength of yeast as platform to rapidly design, test, and evaluate expression and functionality of heterologous Nif proteins targeted to mitochondria, compared to using plants. This platform will be important in further attempts to express even more complex combinations of Nif proteins, which will likely be needed to generate a functional nitrogenase in a eukaryotic organism.

■ METHODS

Strains, Media, and Molecular Biology for Generation of Yeast Strains. S. cerevisiae CEN.PK113−7D (MATa URA3 TRP1 LEU2 HIS3 MAL2−8c SUC2) was the host strain for all constructs and grown at 30 °C in YPD media with 200 μg/mL G418 added when appropriate. Yeast transformations were carried out according to the lithium acetate method.[50,51] Chemically competent E. coli DH5α (New England Biolabs) was used as a cloning strain and grown at 37 °C in LB media with appropriate antibiotics (100 μg/mL carbenicillin, 25 μg/mL chloramphenicol, or 25 μg/mL kanamycin) and inducer (100 μL of 40 mg/mL X-gal) was spread and dried on plates for blue/white screening when appropriate.

All Sanger sequencing reactions were performed by Quintara Biosciences. Plasmid isolations were performed with Qiagen Qiaprep kits where purity and yield were desired or Zymo Zyppy kits when speed was desired. Genomic DNA was isolated using the Promega Wizard Genomic DNA Preparation Kit. Gel electrophoresis was carried out using an Agilent bioanalyzer according to manufacturer instructions.

BsaI was purchased from New England Biolabs. BbsI was purchased from Thermo Fisher Scientific. High concentration T4 DNA ligase was purchased from Promega. Assembly reactions were performed as previously described.[20]

All PCR primers were ordered from IDT (Table S1). All PCRs used Q5 2X Master Mix from New England Biolabs. PCRs were performed on BioRad 1000 series thermocyclers. Synthetic double stranded DNA, when required, was ordered as a gBlock from IDT.

Cluster Genes and Localization Tags. Yeast-optimized versions of A. vinelandii nif H, nif D, nif K, nif U, nif S, nif E, nif N, and nif B (GenScript) were selected to form a minimal nitrogen fixation cluster in S. cerevisiae. An N-terminal His tag was placed in NifD. These genes were directed to the mitochondrial matrix by using three different N-terminal peptide tags: SOD2,[34] SU9,[35] and INDH.[36] While the peptide sequence was maintained, the DNA sequence of the mitochondrial tags was varied to minimize risk of recombination in yeast. Furthermore, the mitochondrial tag for only a subset of the genes was varied. The tags for nif H and nif M were SOD2, SU9, and INDH. The tags for nif E, nif N, and nif B were SU9 and INDH. All other genes used SU9 only. Gene and tag sequences are listed in Table S1.

Yeast Gene Expression Parts. Promoters and terminators were selected from a set of characterized yeast gene expression parts (Eric M. Young, Johannes A. Roubos, D. Benjamin Gordon, and Christopher A. Voigt. *Composability and design of Saccharomyces cerevisiae gene expression* 2017, in preparation). Promoter and terminator sequences are listed in Table S1. Yeast promoter and terminator sequences were cleaned of any BsaI or BbsI sites by point mutation, and then placed into level 0 TypeIIS assembly vectors for versatile cloning into any transcription unit.

To characterize expression strength, promoters and terminators were cloned into respective characterization vectors. The promoter characterization vector, pEMY01AB-PRO, contains homologous sequences to the yeast genome upstream of the promoter cloning site, and a fragment of a yeast codon-optimized GFP downstream. The terminator characterization vector, pEMY01C-TER, contains an overlapping fragment of GFP upstream of the terminator cloning site, and homology to the *Ashbya gossypii* TEF1 promoter used to drive expression of the G418 resistance gene. Using these characterization vectors, any promoter could be combined with any terminator due to the homology occurring in the shared GFP protein. A combination can be integrated into the yeast genome and selected by G418 resistance.

Using homologous recombination and automated liquid handling (LabCyte Echo), thousands of unique transcription units could be created from these characterization vectors. To measure GFP expression strength, yeast strains with unique integrated promoter−terminator pairs were cultured for 24 h in an Infors Multitron shaker at 30 °C with 500 μL of SC+G418 media in a deep-well microtiter plate. Yeast cultures were then diluted 1:10 in sterile water and GFP expression was quantified by flow cytometry using a MACSQuant VYB cytometer. Flow cytometry data was analyzed by taking the geometric mean and standard deviation of 20 000 events for each promoter−terminator combination (Figure 1).

Library Assembly. Hierarchical assembly allows for the combination of any gene with any transcriptional part, yet requires part standardization to eliminate restriction sites used in the assembly process. Yeast promoters and terminators were previously standardized and are maintained in Level 0 vectors with carbenicillin resistance and BsaI sites for Level 1 transcription unit assembly (Figure S6a). Promoters have the scars GTGC and AATG. Terminators have the scars TAAA and CCTC. Upon BsaI digestion, these scars form sticky ends that are then ligated together with the desired open reading frame. However, the yeast-optimized nif genes were not in Level 0 vectors, many did not have the necessary mitochondrial tags, and most needed to be modified to eliminate BsaI and BbsI sites. First, a nif gene was amplified from the parent vector with PCR using primers that added flanking BbsI sites. If a base was to be edited, an upstream fragment and a downstream fragment were amplified, with the modified bases forming a TypeIIS scar sequence and flanked by BbsI sites. Mitochondrial tags were ordered as gBlocks from IDT, also with flanking BbsI sites. This allowed, in one assembly reaction, removal of all TypeIIS sites, addition of mitochondrial tags, and placement in a Level 0 vector for further assembly (Figure S6b).

Each Level 0 vector was picked using blue/white screening and later confirmed by Sanger sequencing of the inserted gene. All nif genes assembled appropriately into Level 0 vectors, with the exception of nif E, which was toxic to *Escherichia coli* in the Level 0 vector, even when a low-copy p15a replicon was used. nif E was therefore ordered as a gBlock from IDT with flanking BsaI sites to be used in Level 1 assembly. Level 0 vectors containing nif M inhibited the growth of E. coli, but were stable.

Level 1 transcription units were assembled by digesting Level 0 vectors with BsaI and ligating them into a chloramphenicol-resistant Level 1 destination vector (Figure S6c). Level 1 destination vectors are designed to determine gene order within the cluster, and are designed to have yeast terminator-like elements as a spacer between transcription units (see spacer design). All desired Level 1 transcription units could be

constructed, although *nif E* was constructed using a p15a replicon rather than ColE1. All Level 1 transcription units are described in Table S3. Each Level 1 assembly was picked using red/white screening and later confirmed by Sanger sequencing.

Level 2 subclusters were assembled by digesting Level 1 vectors with BbsI and ligating into a kanamycin-resistant Level 2 destination vector (Figure S6d). The left subcluster contains *nif S*, *nif E*, and *nif N*. The center subcluster contains *nif D*, the KanMX yeast selection marker, and *nif K*. The right subcluster contains *nif H*, *nif M*, *nif B*, and *nif U*. All Level 2 subclusters are described in Table S3. These subclusters were constructed with the aid of the Labcyte Echo acoustic liquid handling device. Each Level 2 subcluster was picked using blue/white screening and later confirmed by BsaI digestion and gel electrophoresis. A subset was later confirmed by Sanger sequencing. Two subclusters containing *nif M* failed at this stage because the observed fragment size was larger than expected on the gel. This was later determined to be due to transposon insertion.

Subclusters were cut with BsaI to generate linear fragments for yeast transformation. By adding homology fragments with identity to the yeast chromosome XV, subclusters recombined and integrated into this chromosome. This is depicted in Figure S6e. Subclusters were combined with the aid of the Labcyte Echo acoustic liquid handling device. Genome integrations were selected by growth in the presence of G418. A subset of clusters were selected for genomic DNA isolation and verification of insertion by PCR. A further subset was confirmed by Sanger sequencing of the PCR products. All Level 1 vectors are diagrammed in Figure S7 and all yeast-integrated clusters are described in Table S3.

Terminator-like Spacer Design. As an additional safeguard against read-through, terminator-like elements were placed in between each transcription unit in this work. By combining known efficiency, spacing, and polyadenylation elements in forward and reverse orientation, a double terminator-like spacer sequence was generated using a python script (Figure S8). Varying elements and adding random bases between elements diversified the spacer sequence to limit recombination. Spacer sequences are listed in Table S1.

Generation of Plasmids for Galactose-Induced Expression. *E. coli* DH5α was used for storage and amplification of yeast expression vectors, and grown at 37 °C in Luria–Bertani medium supplemented with appropriate antibiotics. Yeast optimized coding sequences for SU9-His-NifD (containing an N-terminal 8xHis-tag) and SU9-NifK were cloned into pESC-LEU (Agilent Technologies), generating pRHB887 (GAL1p::mlsSU9-his8-nifD and GAL10p::mlsSU9-nifK). Plasmid pRHB887 was transformed into *S. cerevisiae* W303–1a (MATa leu2–3, 112 trp1–1 can1–100 ura3–1 ade2–1 his3–11,15) strain GF11[17] (GAL1p::mlsSOD2-his7-nifH and GAL10p::mlsSOD2-flag-nifM in pESC-His, and GAL1p::mls-Su9-nifU and GAL10p::mlsSu9-nifS in pESC-URA), generating SB01Y. To make the vector compatible with transformation into prototrophic *S. cerevisiae* CEN.PK113–7D clone DSN14, the hygromycin marker hphMX4 was amplified from pMJ696 (identical to pAG32[52]) by PCR to replace the LEU2 auxotrophic marker of pRHB887, generating pN2SB16. pN2SB16 was transformed into DSN14, generating SB05Y.

Growth of Yeast Clones and Preparation of Nif Total and Mitochondria Enriched Protein Extracts. *S. cerevisiae* CEN.PK113–7D (and derivative strains constructed herein) and *S. cerevisiae* CEN.PK2–1C (MATa ura3–52 trp1–289 leu2–3,112 his3Δ1 MAL2–8c SUC2) (used to prepare wild-type protein extracts) were grown in flasks at 28 °C and 200 rpm in synthetic drop-out (SD) medium (1.9 g/L yeast nitrogen base, 5 g/L ammonium sulfate, 20 g/L glucose, and Kaiser drop-out mixture[53] (SC -His -Leu -Trp -Ura, FORMEDIUM) supplemented with 40 mg/L tryptophan, 40 mg/L histidine, 20 mg/L uracil, 60 mg/L leucine, and 20 mg/L adenine). Total yeast protein extracts for screening of Nif protein expression levels were performed using mild alkali treatment.[54] Similar loading on SDS-PAGE gels was obtained by preparing samples according to optical density, and was confirmed using Ponceau staining of nitrocellulose membranes and immunoblotting with tubulin antibodies. Mitochondria-enriched isolations were prepared following Diekert et al.,[55] and purity verified using tubulin (cytoplasmic) and HSP60 (mitochondria) targeting antibodies.

Cultures for NifDK purification were grown following a previously reported procedure,[17] in a 4-L fermenter (BIO-STAT) containing rich medium (1% yeast extract, 0.75% bactopeptone, 0.5% bactotryptone, 0.5% ammonium sulfate, 2.5% galactose), supplemented with 25 mg/L ammonium iron(III) citrate, 1.25 mM magnesium sulfate, 1.5 mM calcium chloride, and trace element solution.[17] pH was maintained around 5 using 0.8 M ammonium hydroxide. Plasmid for the inducible expression of NifK and His-tagged NifD in transformed DSN14 (SB05Y) was maintained by supplementing the inoculum growth media with 300 μg/L hygromycin. Plasmids for the inducible expression of Nif proteins in SB01Y was maintained by growing the inoculum in SD medium using their auxotrophic requirements.[17]

Antibodies and Immunoblotting Quantification. Antibodies used for immunoblotting and quantification in this study were as follows: polyclonal antibodies detecting NifDK, NifEN, NifH, NifM, NifU, NifS, NifB, and NafY were raised against purified preparations of the corresponding *A. vinelandii* proteins. Quantification of NifD and NifK polypeptides using NifDK antibodies was accomplished due the apparent difference in migration. NifEN was quantified collectively, as the NifEN polypeptides could not be sufficiently separated by SDS-PAGE. For immunoblotting scoring, both mature Nif protein with processed mitochondria leader sequence and non-processed protein was included.

NifD and NifK specific antibodies were a kind gift from Dr. Yuichi Fujita, Nagoya University. Monoclonal antibodies against tubulin (3H3087, sc-69971, Santa Cruz), His-tag (H-3, sc-8036, Santa Cruz), and HSP60 (LK2, NBP2–34671H, Novus Biologicals) are commercially available.

Preparation of Anaerobic Yeast Cell-Free Extracts. *S. cerevisiae* cells were resuspended in anaerobic lysis buffer (100 mM Tris-HCl pH 8.0, 400 mM NaCl, 20% glycerol) supplemented with 2 mM dithionite (DTH), 1 mM phenyl-methylsulfonyl fluoride (PMSF), 2.5 μM pepstatin A, 1 μg/mL leupeptin, 2 μg/mL aprotinin, 2.5 μM E-64, and 1.5 μM phosphoramidon disodium salt. The cells were lysed in an Emulsiflex-C5 homogenizer (Avestin Inc.) at 25 000 lb per square inch. Cell-free extracts were obtained after removing debris by centrifugation at 50 000g for 1 h at 4 °C under anaerobic conditions, followed by 0.2 μM filtration (Nalgene Rapid-Flow, Thermo Scientific).

Anoxic Native Gels and NafY Binding. Separation of NifDK by discontinuous anaerobic native PAGE was performed inside a glovebox (MBraun). All buffers were previously made anaerobic by sparging with N_2. Anaerobic yeast cell-free extracts were mixed with 2× loading buffer (50 mM Tris-HCl, pH 7.9,

30% glycerol) and separated using a water-cooled SE 600 RUBY vertical unit (GE Healthcare) in 4.5 L running buffer (25 mM Tris-HCl, pH 8.3, 192 mM glycine) supplemented with 2 mM DTH. For NafY binding, 6 μg of NafY was added to 400 μg of anoxic yeast protein extract (100 mM Tris-HCl, pH 8.0, 20% glycerol, supplemented with 2 mM DTH, 1 mM PMSF, 1 μg/mL leupeptin, and 5 μg/mL DNase I), or 30 μg pure *A. vinelandii* NifDK, at a final volume of 75 μL and incubated at room temperature for 60 min. For FeMo-co insertion, 5 μL of pure FeMo-co (at 294 μM, based on iron concentration) was added to 185 μL of yeast extract in a final volume of 400 μL and incubated at 30 °C for 60 min. Before electrophoresis, 2× loading buffer was added to all samples. Gels of 18 cm (7% resolving gel, pH 8.6 and 4% stacking gel, pH 7.0) were made anaerobic at constant current, 12.5 mA per gel for 2 h. Protein samples (20 μL) were separated at constant current, 25 mA per gel, for 6 h at 4 °C. Following gel electrophoresis, proteins were transferred to nitrocellulose membranes and analyzed by immunoblotting following standard methods.

NifDK Activity. NifDK activity of protein purified from *S. cerevisiae* cells was analyzed by acetylene reduction assay following incubation of pure FeMo-co,[42] and after addition of excess of purified *A. vinelandii* NifH protein and ATP-regenerating mixture (1.23 mM ATP, 18 mM phosphocreatine, 2.2 mM MgCl$_2$, 3 mM DTH and 40 mg of creatine phosphokinase).[45] Positive control reactions were carried out with pure preparations of *A. vinelandii* NifH and NifDK (apo and holo forms) proteins.[8]

NifDK Purification. His-tagged NifD was purified by Co^{2+} affinity chromatography under anaerobic conditions (<0.1 ppm of O$_2$) using an AKTA Prime FPLC system (GE Healthcare) inside a glovebox (MBraun). All buffers were previously made anaerobic by sparging with N$_2$. Anaerobic cell-free extract from 50 to 100 g of cell paste was loaded at 2 mL/min onto a column filled with 5 mL of IMAC resin (GE Healthcare) equilibrated with buffer A (25 mM Tris-HCl pH 7.4, 400 mM NaCl, 20% glycerol) and washed with three successive washes of Buffer B supplemented with 0, 10, and 30 mM imidazole, respectively. Bound protein was eluted with buffer A containing 250 mM imidazole. The eluted fractions were concentrated using a 100 kDa cutoff pore (Amicon Ultra-15, Millipore) and then desalted in PD10 columns equilibrated with buffer A. Purified NifDK was frozen and stored in liquid N$_2$.

ASSOCIATED CONTENT

Ⓢ Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acssynbio.6b00371.

Figure S1, analysis of promoter efficiency in generating detectable NifEN protein; Figure S2, migration of NifD and NifK polypeptides; Figure S3, mitochondria targeting of Nif proteins; Figure S4, expression of Nif proteins in DSN14; Figure S5, NifDK tetramer formation in DSN14; Figure S6, detailed TypeIIS hierarchy for cluster assembly; Figure S7, level 1 destination vector design; Figure S8, spacer design (PDF)

Sequences (XLSX)

Part assignment (XLSX)

Assembly (XLSX)

Quality control and results (XLSX)

Mitochondrial signal efficiency (XLSX)

Promoter efficiency (XLSX)

Regression analysis (XLSX)

AUTHOR INFORMATION

Corresponding Authors
*E-mail: lm.rubio@upm.es.
*E-mail: cavoigt@gmail.com.

ORCID ⊙
Luis M. Rubio: 0000-0003-1596-2475

Author Contributions
S.B., E.Y., E.A.S., and G.L.T. carried out the experimental work and contributed to experimental design and data analysis. M.V. carried out experimental work. L.M.R, and C.A.V. contributed to experimental design and data analysis. S.B., E.Y., E.A.S., C.A.V., and L.M.R. wrote the paper. C.A.V. and L.M.R initiated and directed this research.

Notes
The authors declare no competing financial interest.

REFERENCES

(1) Dos Santos, P. C., Fang, Z., Mason, S. W., Setubal, J. C., and Dixon, R. (2012) Distribution of nitrogen fixation and nitrogenase-like sequences amongst microbial genomes. *BMC Genomics* 13, 162.

(2) Hoffman, B. M., Lukoyanov, D., Yang, Z.-Y. Y., Dean, D. R., and Seefeldt, L. C. (2014) Mechanism of nitrogen fixation by nitrogenase: The next stage. *Chem. Rev.* 114, 4041−4062.

(3) Spatzal, T., Aksoyoglu, M., Zhang, L., Andrade, S. L. A., Schleicher, E., Weber, S., Rees, D. C., and Einsle, O. (2011) Evidence for interstitial carbon in nitrogenase FeMo cofactor. *Science* 334, 940.

(4) Einsle, O., Tezcan, F. A., Andrade, S. L. A., Schmid, B., Yoshida, M., Howard, J. B., and Rees, D. C. (2002) Nitrogenase MoFe-protein at 1.16 A resolution: a central ligand in the FeMo-cofactor. *Science* 297, 1696−1700.

(5) Rubio, L. M., and Ludden, P. W. (2008) Biosynthesis of the iron-molybdenum cofactor of nitrogenase. *Annu. Rev. Microbiol.* 62, 93−111.

(6) Stokstad, E. (2016) The nitrogen fix. *Science* 353, 1225−1227.

(7) Curatti, L., and Rubio, L. M. (2014) Challenges to develop nitrogen-fixing cereals by direct nif-gene transfer. *Plant Sci.* 225, 130−137.

(8) Curatti, L., Hernandez, J. A., Igarashi, R. Y., Soboh, B., Zhao, D., and Rubio, L. M. (2007) In vitro synthesis of the iron-molybdenum cofactor of nitrogenase from iron, sulfur, molybdenum, and homocitrate using purified proteins. *Proc. Natl. Acad. Sci. U. S. A.* 104, 17626−17631.

(9) Dixon, R. A., and Postgate, J. R. (1972) Genetic transfer of nitrogen fixation from Klebsiella pneumoniae to Escherichia coli. *Nature* 237, 102−103.

(10) Wang, X., Yang, J.-G., Chen, L., Wang, J.-L., Cheng, Q., Dixon, R., and Wang, Y.-P. (2013) Using Synthetic Biology to Distinguish and Overcome Regulatory and Functional Barriers Related to Nitrogen Fixation. *PLoS One* 8, e68677.

(11) Wang, L., Zhang, L., Liu, Z., Zhao, D., Liu, X., Zhang, B., Xie, J., Hong, Y., Li, P., Chen, S., Dixon, R., and Li, J. (2013) A Minimal

Nitrogen Fixation Gene Cluster from Paenibacillus sp. WLY78 Enables Expression of Active Nitrogenase in Escherichia coli. *PLoS Genet. 9*, 1−11.

(12) Poza-Carrión, C., Jiménez-Vicente, E., Navarro-Rodríguez, M., Echavarri-Erasun, C., and Rubio, L. M. (2014) Kinetics of nif gene expression in a nitrogen-fixing bacterium. *J. Bacteriol. 196*, 595−603.

(13) Scharff, L. B., and Bock, R. (2014) Synthetic biology in plastids. *Plant J. 78*, 783−798.

(14) Ivleva, N. B., Groat, J., Staub, J. M., and Stephens, M. (2016) Expression of active subunit of nitrogenase via integration into plant organelle genome. *PLoS One 11*, e0160951.

(15) Lill, R., and Muhlenhoff, U. (2008) Maturation of iron-sulfur proteins in eukaryotes: mechanisms, connected processes, and diseases. *Annu. Rev. Biochem. 77*, 669−700.

(16) Niazi, A. K., Mileshina, D., Cosset, A., Val, R., Weber-Lotfi, F., and Dietrich, A. (2013) Targeting nucleic acids into mitochondria: progress and prospects. *Mitochondrion 13*, 548−558.

(17) Lopez-Torrejon, G., Jimenez-Vicente, E., Buesa, J. M., Hernandez, J. A., Verma, H. K., and Rubio, L. M. (2016) Expression of a functional oxygen-labile nitrogenase component in the mitochondrial matrix of aerobically grown yeast. *Nat. Commun. 7*, 11426.

(18) Lill, R. (2009) Function and biogenesis of iron-sulphur proteins. *Nature 460*, 831−838.

(19) Temme, K., Zhao, D., and Voigt, C. A. (2012) Refactoring the nitrogen fixation gene cluster from Klebsiella oxytoca. *Proc. Natl. Acad. Sci. U. S. A. 109*, 7085−7090.

(20) Smanski, M. J., Bhatia, S., Zhao, D., Park, Y., B A Woodruff, L., Giannoukos, G., Ciulla, D., Busby, M., Calderon, J., Nicol, R., Gordon, D. B., Densmore, D., and Voigt, C. A. (2014) Functional optimization of gene clusters by combinatorial design and assembly. *Nat. Biotechnol. 32*, 1241−1249.

(21) Smanski, M. J., Zhou, H., Claesen, J., Shen, B., Fischbach, M. A., and Voigt, C. A. (2016) Synthetic biology to access and expand nature's chemical diversity. *Nat. Rev. Microbiol. 14*, 135−149.

(22) Chan, L. Y., Kosuri, S., and Endy, D. (2005) Refactoring bacteriophage T7. *Mol. Syst. Biol. 1*, 2005.0018.

(23) Lee, M. E., DeLoache, W. C., Cervantes, B., and Dueber, J. E. (2015) A Highly Characterized Yeast Toolkit for Modular, Multipart Assembly. *ACS Synth. Biol. 4*, 975−986.

(24) Nevoigt, E., Kohnke, J., Fischer, C. R., Alper, H., Stahl, U., and Stephanopoulos, G. (2006) Engineering of promoter replacement cassettes for fine-tuning of gene expression in Saccharomyces cerevisiae. *Appl. Environ. Microbiol. 72*, 5266−5273.

(25) Blount, B. A., Weenink, T., Vasylechko, S., and Ellis, T. (2012) Rational Diversification of a Promoter Providing Fine-Tuned Expression and Orthogonal Regulation for Synthetic Biology. *PLoS One 7*, e33279.

(26) Sun, J., Shao, Z., Zhao, H., Nair, N., Wen, F., Xu, J.-H., and Zhao, H. (2012) Cloning and characterization of a panel of constitutive promoters for applications in pathway engineering in Saccharomyces cerevisiae. *Biotechnol. Bioeng. 109*, 2082−2092.

(27) Curran, K. A., Karim, A. S., Gupta, A., and Alper, H. S. (2013) Use of expression-enhancing terminators in Saccharomyces cerevisiae to increase mRNA half-life and improve gene expression control for metabolic engineering applications. *Metab. Eng. 19*, 88−97.

(28) Yamanishi, M., Ito, Y., Kintaka, R., Imamura, C., Katahira, S., Ikeuchi, A., Moriya, H., and Matsuyama, T. (2013) A genome-wide activity assessment of terminator regions in Saccharomyces cerevisiae provides a terminatome" toolbox. *ACS Synth. Biol. 2*, 337−347.

(29) Salis, H. M., Mirsky, E. A., and Voigt, C. A. (2009) Automated design of synthetic ribosome binding sites to control protein expression. *Nat. Biotechnol. 27*, 946−950.

(30) Galanie, S., Thodey, K., Trenchard, I. J., Filsinger Interrante, M., and Smolke, C. D. (2015) Complete biosynthesis of opioids in yeast. *Science 349*, 1095−1100.

(31) Beekwilder, J., van Rossum, H. M., Koopman, F., Sonntag, F., Buchhaupt, M., Schrader, J., Hall, R. D., Bosch, D., Pronk, J. T., van Maris, A. J. A., and Daran, J.-M. (2014) Polycistronic expression of a

beta-carotene biosynthetic pathway in Saccharomyces cerevisiae coupled to beta-ionone production. *J. Biotechnol. 192*, 383−392.

(32) Roehner, N., Young, E. M., Voigt, C. A., Gordon, D. B., and Densmore, D. (2016) Double Dutch: A Tool for Designing Combinatorial Libraries of Biological Systems. *ACS Synth. Biol. 5*, 507−517.

(33) Oldroyd, G. E. D., and Dixon, R. (2014) Biotechnological solutions to the nitrogen problem. *Curr. Opin. Biotechnol. 26*, 19−24.

(34) de la Loza, M. C. D., and Wellinger, R. E. (2008) A novel approach for organelle-specific DNA damage targeting reveals different susceptibility of mitochondrial DNA to the anticancer drugs camptothecin and topotecan. *Nucleic Acids Res. 37*, e26.

(35) Westermann, B., and Neupert, W. (2000) Mitochondria-targeted green fluorescent proteins: Convenient tools for the study of organelle biogenesis in Saccharomyces cerevisiae. *Yeast 16*, 1421−1427.

(36) Wydro, M. M., Sharma, P., Foster, J. M., Bych, K., Meyer, E. H., and Balk, J. (2013) The evolutionarily conserved iron-sulfur protein INDH is required for complex I assembly and mitochondrial translation in Arabidopsis. *Plant Cell 25*, 4014−27.

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

(38) Homer, M. J., Dean, D. R., and Roberts, G. P. (1995) Characterization of the γ protein and its involvement in the metallocluster assembly and maturation of dinitrogenase from Azotobacter vinelandii. *J. Biol. Chem. 270*, 24745−24752.

(39) Tal, S., Chun, T. W., Gavini, N., and Burgess, B. K. (1991) The delta-nifB (or delta-nifE) FeMo cofactor-deficient MoFe protein is different from the delta-nifH protein. *J. Biol. Chem. 266*, 10654−10657.

(40) Rubio, L. M., and Ludden, P. W. (2005) Maturation of nitrogenase: a biochemical puzzle. *J. Bacteriol. 187*, 405−414.

(41) Dyer, D. H., Rubio, L. M., Thoden, J. B., Holden, H. M., Ludden, P. W., and Rayment, I. (2003) The three-dimensional structure of the core domain of Naf Y from Azotobacter vinelandii determined at 1.8-A resolution. *J. Biol. Chem. 278*, 32150−32156.

(42) Hernandez, J. A., Phillips, A. H., Erbil, W. K., Zhao, D., Demuez, M., Zeymer, C., Pelton, J. G., Wemmer, D. E., and Rubio, L. M. (2011) A sterile alpha-motif domain in NafY targets apo-NifDK for iron-molybdenum cofactor delivery via a tethered domain. *J. Biol. Chem. 286*, 6321−8.

(43) Rangaraj, P., Ryle, M. J., Lanzilotta, W. N., Ludden, P. W., and Shah, V. K. (1999) In Vitro Biosynthesis of Iron-Molybdenum Cofactor and Maturation of the nif-encoded Apodinitrogenase. *J. Biol. Chem. 274*, 19778−19784.

(44) Rangaraj, P., Shah, V. K., and Ludden, P. W. (1997) ApoNifH functions in iron-molybdenum cofactor synthesis and apodinitroge-nase maturation. *Proc. Natl. Acad. Sci. U. S. A. 94*, 11250−11255.

(45) Shah, V. K., and Brill, W. J. (1973) Nitrogenase. IV. Simple method of purification to homogeneity of nitrogenase components from Azotobacter vinelandii. *Biochim. Biophys. Acta, Bioenerg. 305*, 445−454.

(46) Petrova, N., Gigova, L., and Venkov, P. (2000) NifH and NifM proteins interact as demonstrated by the yeast two-hybrid system. *Biochem. Biophys. Res. Commun. 270*, 863−867.

(47) Howard, K. S., McLean, P. A., Hansen, F. B., Lemley, P. V., Koblan, K. S., and Orme-Johnson, W. H. (1986) Klebsiella pneumoniae nifM gene product is required for stabilization and activation of nitrogenase iron protein in Escherichia coli. *J. Biol. Chem. 261*, 772−778.

(48) Gavini, N., Tungtur, S., and Pulakat, L. (2006) Peptidyl-prolyl cis/trans isomerase-independent functional NifH mutant of Azoto-bacter vinelandii. *J. Bacteriol. 188*, 6020−6025.

(49) Reider Apel, A., d'Espaux, L., Wehrs, M., Sachs, D., Li, R. A., Tong, G. J., Garber, M., Nnadi, O., Zhuang, W., Hillson, N. J., Keasling, J. D., and Mukhopadhyay, A. (2017) A Cas9-based toolkit to program gene expression in Saccharomyces cerevisiae. *Nucleic Acids Res. 45*, 496−508.

(50) Gietz, R. D., and Schiestl, R. H. (2007) High-efficiency yeast transformation using the LiAc/SS carrier DNA/PEG method. *Nat. Protoc. 2*, 31−34.

(51) Hegemann, J. H., and Heick, S. B. (2011) Delete and repeat: a comprehensive toolkit for sequential gene knockout in the budding yeast Saccharomyces cerevisiae. *Methods Mol. Biol. 765*, 189−206.

(52) Goldstein, A. L., and McCusker, J. H. (1999) Three new dominant drug resistance cassettes for gene disruption in Saccharomyces cerevisiae. *Yeast 15*, 1541−1553.

(53) Kaiser, C., Michaelis, S., and Mitchell, A. (1994) *Methods in Yeast Genetics: A Cold Spring Harbor Laboratory Course Manual*, Cold Spring Harbor Laboratory Press, New York.

(54) Kushnirov, V. V. (2000) Rapid and reliable protein extraction from yeast. *Yeast 16*, 857−860.

(55) Diekert, K., de Kroon, A. I., Kispal, G., and Lill, R. (2001) Isolation and subfractionation of mitochondria from the yeast Saccharomyces cerevisiae. *Methods Cell Biol. 65*, 37−51.