SUPPORTING INFORMATION

A Triple KnockOut Isobaric-labelling Quality Control Platform with an Integrated On-line Database Search

Jeremy P. Gygi[†], Ramin Rad [†], Jose Navarrete-Perea [†], Simon Younesi[†], Steven P. Gygi[†], Joao A. Paulo[†,#]

[†]Department of Cell Biology, Harvard Medical School, Boston, MA 02115, United States

#Corresponding author

Supplemental Methods: Extended sample preparation and mass spectrometry data analysis.

Figure S1: Composition and assembly of TKOpro standards.

Figure S2: TVT2.0 output customization.

Figure S3: Example output generated by the TVT2.0 of TKO16pro data acquired from three acquisition strategies.

Figure S4: Comparison of TKOpro16 and TKO11 standards.

Supplemental Methods

Materials. Tandem mass tag isobaric reagents (TMT and TMTpro) were from ThermoFisher Scientific (Waltham, MA). Water and organic solvents were from J.T. Baker (Center Valley, PA).

Yeast strains and growth conditions. *Saccharomyces cerevisiae* strains from the haploid MATalpha collection (BY4742 MATα his3Δ1 leu2Δ0 lys2Δ0 ura3Δ0) were purchased from openbiosystems. Cultures were grown in standard yeast-peptone-dextrose (YPD) media to an optical density (OD) of 0.8/mL and then harvested, as described previously [1] and highlighted below. We emphasize that the knockout strains selected (Δmet6, Δpfk2, Δura2) fulfilled three criteria. Specifically, the deleted protein: 1) has no redundant tryptic peptides in the *S. cerevisiae* proteome, 2) is non-essential for yeast growth, and 3) is abundant such that tryptic peptides are observed with short data acquisition methods.

Cell lysis and protein digestion. Yeast cultures were harvested by centrifugation, washed twice with ice cold deionized water, and resuspended in lysis buffer: 50 mM EPPS pH 8.5, 8 M urea, and protease (complete mini, EDTA-free) inhibitors (Roche, Basel, Switzerland). Cells were lysed using the MiniBeadbeater (Biospec, Bartlesville, OK) in microcentrifuge tubes at maximum speed for three cycles of 60 sec each, with 1 min pauses on ice between cycles to avoid overheating of the lysates. After centrifugation, cleared lysates were transferred to new tubes. We determined the protein concentration in the lysate using the bicinchoninic acid (BCA) protein assay (Thermo Fisher Scientific, Waltham, MA).

Proteins were subjected to disulfide reduction with 5 mM tris (2-carboxyethyl)phosphine (TCEP), (room temperature, 30 min) and alkylation with 10 mM iodoacetamide (room temperature, 30 min in the dark). Excess iodoacetamide was quenched with 10 mM dithiotreitol

(room temperature, 15 min in the dark). Methanol-chloroform precipitation was performed prior to protease digestion. In brief, four parts neat methanol was added to each sample and vortexed, one-part chloroform was added to the sample and vortexed, and three parts water was added to the sample and vortexed. The sample was centrifuged at 20,000 RPM for 2 min at room temperature and, after removing both the aqueous and organic phases, subsequently washed once with 100% methanol. Samples are not dried, rather, all but ~10 µL of methanol was aspirated to allow for better solubility of the precipitated proteins.

Samples were resuspended in 200 mM EPPS, pH 8.5 and digested at room temperature for 16 h with LysC protease at a 100:1 protein-to-protease ratio. Trypsin was then added at a 100:1 protein-to-protease ratio and the reaction was incubated 6 h at 37°C.

Tandem mass tag labeling. For the TKO11 standard, TMT reagents (0.8 mg) were dissolved in anhydrous acetonitrile (40 µL) of which 10 µL was added to the peptides (100 µg) along with 30 µL of acetonitrile to achieve a final acetonitrile concentration of approximately 30% (v/v). Specifically, for the TKO11, peptides from the Δ*met6* strain replicates were conjugated to tags 126, 127n and 127c; the Δ*pfk2* or Δ*his4* strain replicates with tags 128n, 128c and 129n; the Δ*ura2* strain replicates with tags 129c, 130n and 130c; and the two wild type replicates were labeled with 131 and 131c. Following incubation at room temperature for 1 h, the reaction was quenched with hydroxylamine to a final concentration of 0.3% (v/v). The TMT-labeled samples were pooled at a 1:1 ratio across all channels. The sample was vacuum centrifuged to near dryness and subjected to C18 solid-phase extraction (SPE) (Sep-Pak, Waters).

For the TKO9pro standard, TMT reagents (5 mg) were dissolved in anhydrous acetonitrile (400 µL) of which 12 µL was added to the peptides (100 µg) along with 30 µL of acetonitrile to

achieve a final acetonitrile concentration of approximately 30% (v/v). Specifically, for the TKO9pro, peptides from the Δmet6 strain replicates were conjugated to tags 126, 127 and 128; the Δpfk2 or Δhis4 strain replicates with tags 129, 130 and 131; and the Δura2 strain replicates with tags 132, 133 and 134. Following incubation at room temperature for 1 h, the reaction was quenched with hydroxylamine to a final concentration of 0.3% (v/v). The TMTpro-labeled samples were pooled at a 1:1 ratio across all channels. The sample was vacuum centrifuged to near dryness and subjected to C18 solid-phase extraction (SPE) (Sep-Pak, Waters).

For the TKO16 pro standard, TMT reagents (5 mg) were dissolved in anhydrous acetonitrile (40 μL) of which 12 μL was added to the peptides (400 μg) along with 30 μL of acetonitrile to achieve a final acetonitrile concentration of approximately 30% (v/v). Specifically, for the TKO9pro, peptides from the Δmet6 strain replicates were conjugated to tags 126, 127n, 127c, and 128n; the Δpfk2 or Δhis4 strain replicates with tags 128c, 129n, 129c, and 130n; the Δura2 strain replicates with tags 130c, 131n, 131c, and 132n; and the wild type replicates were labeled with 132c, 133n, 133c, and 134. Following incubation at room temperature for 1 h, the reaction was quenched with hydroxylamine to a final concentration of 0.3% (v/v). The TMTpro-labeled samples were pooled at a 1:1 ratio across all channels. The sample was vacuum centrifuged to near dryness and subjected to C18 solid-phase extraction (SPE) (Sep-Pak, Waters).

Liquid chromatography and tandem mass spectrometry. The samples were reconstituted in 5% acetonitrile and 5% formic acid for LC-MS/MS processing. Peptides were separated on a 35 cm long, 100 μm inner diameter microcapillary column packed with Accucore (2.6 μm, 150Å) resin (ThermoFisher Scientific). For each analysis, we loaded 0.5 μg of the TKO standard onto the C18 capillary column using a Proxeon NanoLC-1200 UHPLC. Peptides were separated in-

line with the mass spectrometer using gradients of 6 to 26% acetonitrile in 0.125% formic acid at a flow rate of ~500 nL/min. Standard TKO analyses are typically 45 min, as noted in the main text.

Data analysis. Mass spectra were processed using a Sequest-based (Comet) software pipeline [2, 3]. Database searching included all entries from the yeast SGD (*Saccharomyces* Genome Database) (March 20, 2015). This database was concatenated with one composed of all protein sequences in the reversed order. Searches were performed using a 50 ppm precursor ion tolerance. The product ion tolerance was set to 0.9 Da for SPS-MS3 and RTS-MS3 data, but to 0.03Da for hrMS2 data. These wide mass tolerance windows were chosen to maximize sensitivity in conjunction with Comet searches and linear discriminant analysis [4, 5]. TMT tags on lysine residues and peptide N termini (+229.163 Da for TMT or +304.207 for TMTpro) and carbamidomethylation of cysteine residues (+57.021 Da) were set as static modifications, while oxidation of methionine residues (+15.995 Da) was set as a variable modification.

Peptide-spectrum matches (PSMs) were adjusted to a 1% false discovery rate (FDR) [6, 7]. PSM filtering was performed using a linear discriminant analysis, as described previously [5], while considering the following parameters: XCorr, ΔCn, missed cleavages, peptide length, charge state, and precursor mass accuracy. For TMT-based reporter ion quantitation, we extracted the signal-to-noise (S:N) ratio for each TMT channel and found the closest matching centroid to the expected mass of the TMT reporter ion. PSMs were identified, quantified, and collapsed to a 1% peptide false discovery rate (FDR) and then collapsed further to a final protein-level FDR of 1%. Moreover, protein assembly was guided by principles of parsimony to produce the smallest set of proteins necessary to account for all observed peptides.

Peptide intensities were quantified by summing reporter ion counts across all matching PSMs so as to give greater weight to more intense ions [8, 9]. PSMs with poor quality, spectra with TMT reporter summed signal-to-noise measurements that were less than 200, or with no MS3 spectra (for MS3-based methods) were excluded from quantitation. Isolation specificity of ≥ 0.8 (i.e., peptide purity >80%) was required [9].

Supplemental references:

1. Paulo, J. A.; O'Connell, J. D.; Everley, R. A.; O'Brien, J.; Gygi, M. A.; Gygi, S. P., Quantitative mass spectrometry-based multiplexing compares the abundance of 5000 S. cerevisiae proteins across 10 carbon sources. *J Proteomics* **2016,** *148*, 85-93.
2. Eng, J. K.; Hoopmann, M. R.; Jahan, T. A.; Egertson, J. D.; Noble, W. S.; MacCoss, M. J., A deeper look into Comet--implementation and features. *J Am Soc Mass Spectrom* **2015,** *26* (11), 1865-74.
3. Eng, J. K.; Jahan, T. A.; Hoopmann, M. R., Comet: an open-source MS/MS sequence database search tool. *Proteomics* **2013,** *13* (1), 22-4.
4. Beausoleil, S. A.; Villen, J.; Gerber, S. A.; Rush, J.; Gygi, S. P., A probability-based approach for high-throughput protein phosphorylation analysis and site localization. *Nature biotechnology* **2006,** *24* (10), 1285-92.
5. Huttlin, E. L.; Jedrychowski, M. P.; Elias, J. E.; Goswami, T.; Rad, R.; Beausoleil, S. A.; Villen, J.; Haas, W.; Sowa, M. E.; Gygi, S. P., A tissue-specific atlas of mouse protein phosphorylation and expression. *Cell* **2010,** *143* (7), 1174-89.
6. Elias, J. E.; Gygi, S. P., Target-decoy search strategy for mass spectrometry-based proteomics. *Methods Mol Biol* **2010,** *604*, 55-71.
7. Elias, J. E.; Gygi, S. P., Target-decoy search strategy for increased confidence in large-scale protein identifications by mass spectrometry. *Nat Methods* **2007,** *4* (3), 207-14.
8. McAlister, G. C.; Nusinow, D. P.; Jedrychowski, M. P.; Wuhr, M.; Huttlin, E. L.; Erickson, B. K.; Rad, R.; Haas, W.; Gygi, S. P., MultiNotch MS3 enables accurate, sensitive, and multiplexed detection of differential expression across cancer cell line proteomes. *Anal Chem* **2014,** *86* (14), 7150-8.
9. McAlister, G. C.; Huttlin, E. L.; Haas, W.; Ting, L.; Jedrychowski, M. P.; Rogers, J. C.; Kuhn, K.; Pike, I.; Grothe, R. A.; Blethrow, J. D.; Gygi, S. P., Increasing the multiplexing capacity of TMTs using reporter ion isotopologues with isobaric masses. *Anal Chem* **2012,** *84* (17), 7469-78.

Figure S1:

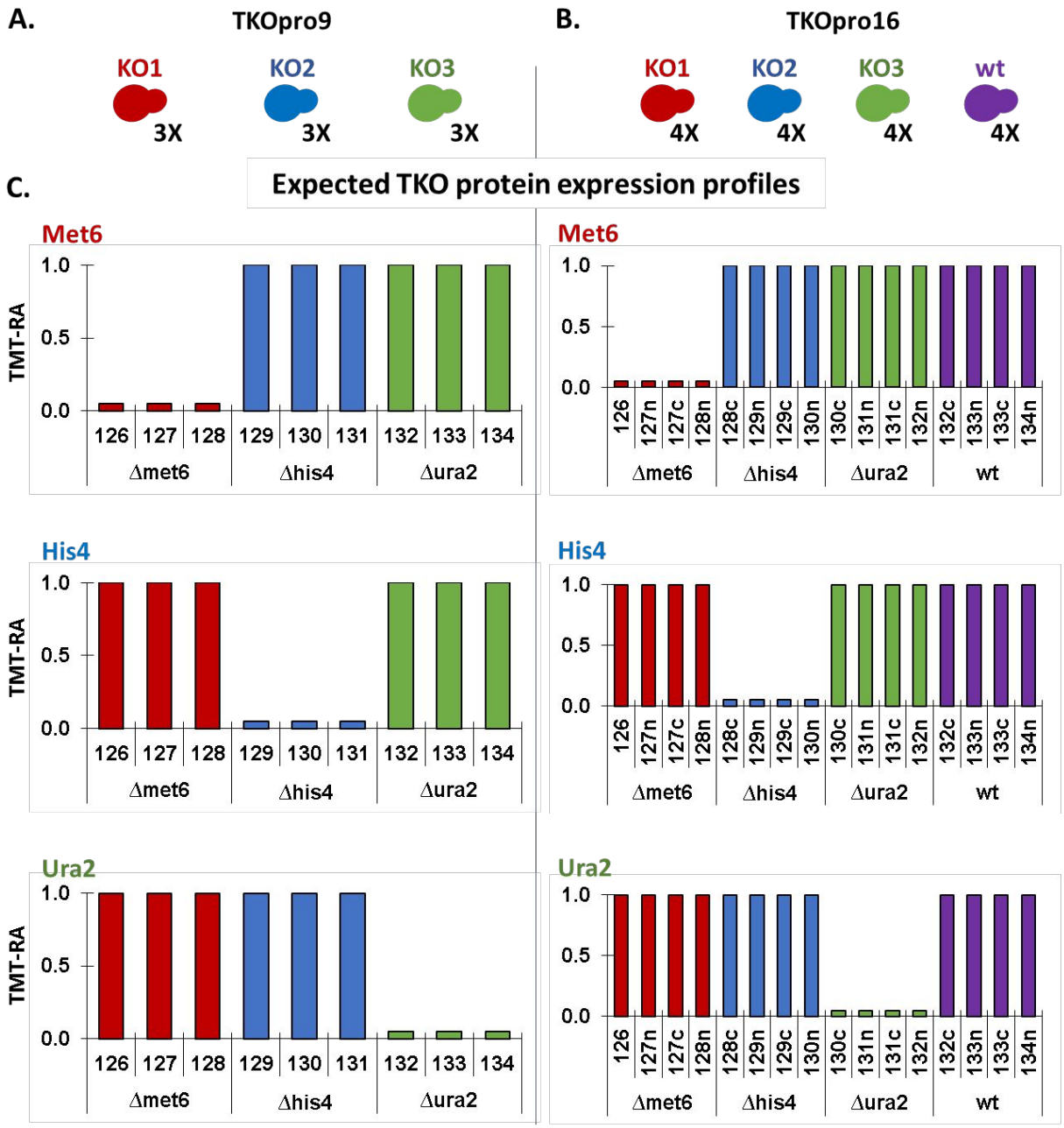

Figure S1: Composition and assembly of TKOpro standards. The TKO standard is composed of three yeast deletion strains. Variants of the TKO standard include the previously described TKO6, TKO9, and TKO11 in addition to the **A)** TKOpro9 and **B)** TKOpro16 described herein. Although the number of isobaric tags used are different among the variants, all share a similar layout. **C)** Depicted are the expected TKO protein expression profiles for each of the knockout (TKO) proteins. TMT-RA, tandem mass tag relative abundance.

Figure S2:

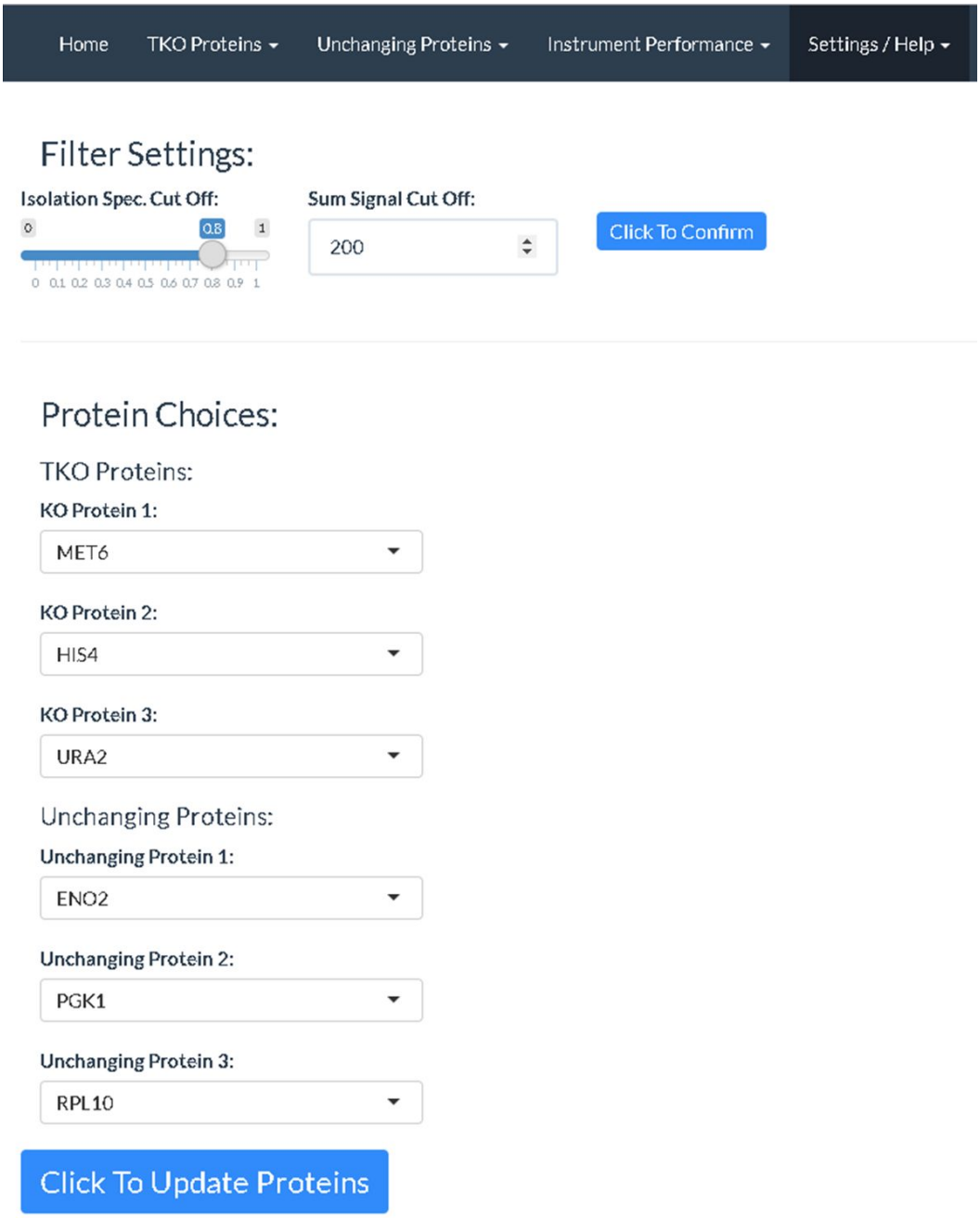

Figure S2: TVT2.0 output customization. Isolation specificity (i.e., isolation purity) may be adjusted, as well as signal-to-noise thresholds (top). Proteins selected as TKO proteins (middle) or unchanging proteins (bottom) can also be altered, allowing for fully customizable TKO standards in both 3-3-3 (TKO9, TKOpro9, TKO11) and 4-4-4 (TKO16) arrangements.

Figure S3:


A. No. of proteins and peptides
RTS-MS3
hrMS2
SPS-MS3
830(Current) 831 832 833 834 835
Total Proteins 1686 1688 1588 1559 1255 1289
Total Peptides 7132 7277 7268 7151 5688 5784
B. IFI comparison
Gene Symbol: RTS-MS3 hrMS2 SPS-MS3
830(Current) 831 832 833 834 835
MET6 0.944 +/- 0.059 | 0.957 +/- 0.05 | 0.768 +/- 0.197 | 0.757 +/- 0.292 | 0.896 +/- 0.114 | 0.919 +/- 0.111
HIS4 0.996 +/- NA | 0.885 +/- 0.08 | 0.93 +/- 0.098 | 0.97 +/- 0.042 | 0.991 +/- NA | 0.917 +/- 0.089
URA2 0.978 +/- 0.026 | 0.912 +/- 0.124 | 0.832 +/- 0.128 | 0.777 +/- 0.182 | 0.762 +/- 0.284 | 0.883 +/- 0.152
Total TKO 0.962 +/- 0.048 | 0.923 +/- 0.103 | 0.832 +/- 0.137 | 0.791 +/- 0.199 | 0.819 +/- 0.239 | 0.901 +/- 0.125
C. No. of TKO peptides
RTS-MS3 hrMS2 SPS-MS3
Gene Symbol: 830(Current) 831 832 833 834 835
MET6 9 6 3 4 8 7
HIS4 1 2 2 2 1 6
URA2 8 12 11 15 14 12
Total TKO 18 20 16 21 23 25
D. XCorr
E. PPM mass error
F. Σsignal:noise
G. Isolation purity
H. Precursor intensity
I. MS2 ion time (ms)
J. MS3 ion time (ms)
K. Peak width (min)
RTS-MS3 hrMS2 SPS-MS3

Figure S3: Example output generated by the TVT2.0 of TKO16pro data acquired from three acquisition strategies. The data and experimental set-up are outlined in **Figure 3**. The first replicate runs (830/831, blue-green/orange) are RTS-MS3, the second replicate runs (832/833, blue/fuchsia) are hrMS2, and the third replicate (834/835, lime green/yellow) are SPS-MS3. The following tables and graphs were retrieved directly from the TVT2.0 performance comparison output: **A)** number of total proteins and peptides for each run, **B)** comparison of the IFI (interference-free index), **C)** number of TKO peptides quantified, **D)** XCorr, cross correlation score, **E)** PPM mass error (mass accuracy), **F)** TMT summed signal-to-noise for reporter ions, **G)** isolation purity (isolation specificity), **H)** precursor max intensity, **I)** MS2 and **J)** MS3 ion injection times, and **K)** chromatographic peak width.

Figure S4:

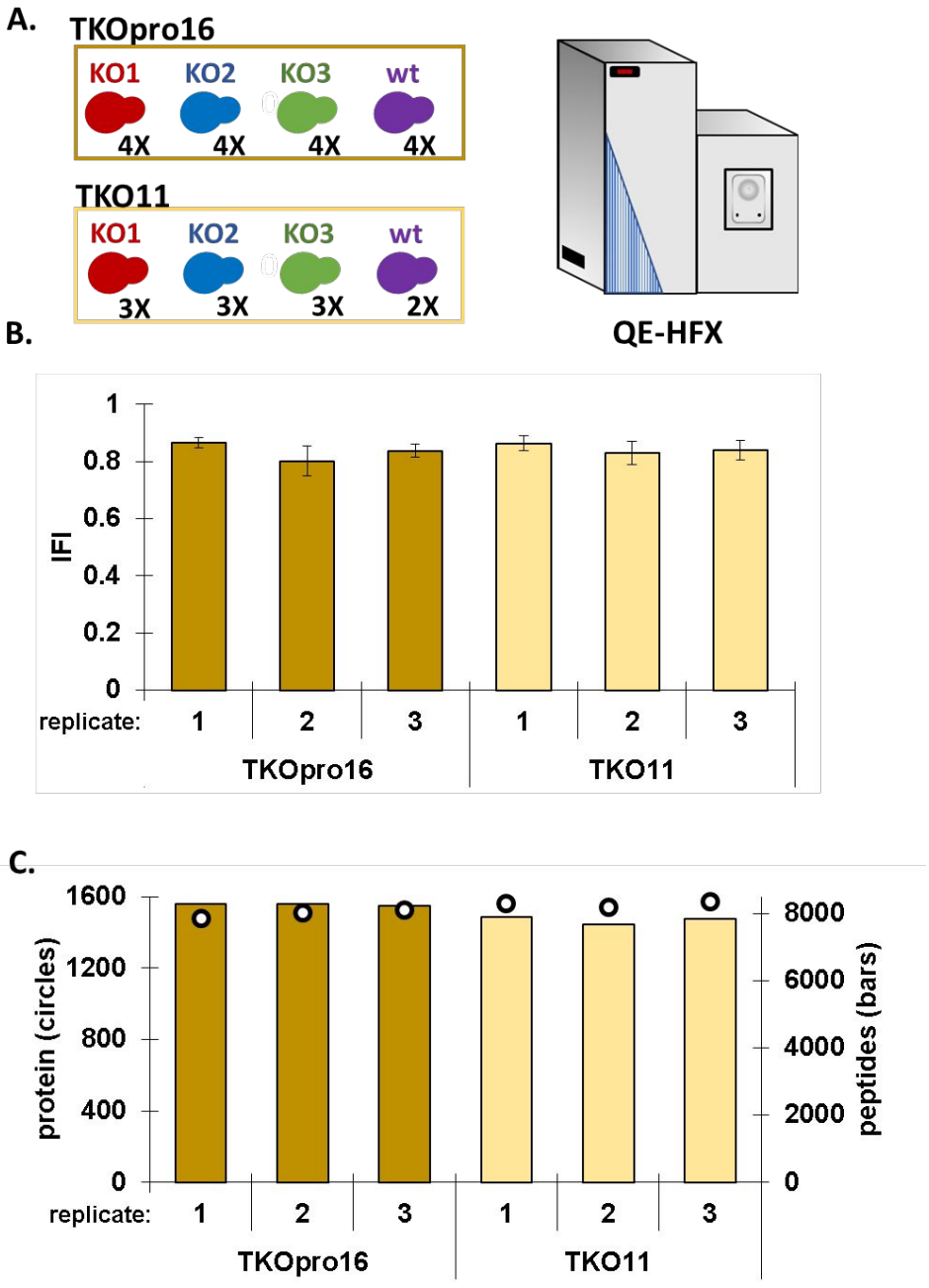

Figure S4: Comparison of TKOpro16 and TKO11 standards. A) Composition of the TKOpro16 (top left) and the TKO11 (bottom left) standards which were analyzed on a Q-Exactive HF-X mass spectrometer (right). **B)** Bar plot of the interference-free index (IFI) for triplicate analysis using the standard hrMS2-TMT method template (available in Tune2.9). **C)** Combination plot indicating the number of proteins (circle) and peptides (bars) for the data acquired above.