A causal model of ion interference enables assessment and correction of ratio compression in multiplex proteomics

Moritz Madern[1,2,#], Wolfgang Reiter[1,2], Florian Stanek[3], Natascha Hartl[1], Karl Mechtler[3], Markus Hartl[1,2]*

[1] Max Perutz Labs, Mass Spectrometry Facility, Vienna Biocenter Campus (VBC), Dr. Bohr-Gasse 7, Vienna, Austria

[2] University of Vienna, Center for Molecular Biology, Department for Biochemistry and Cell Biology, Vienna Biocenter Campus (VBC), Dr. Bohr-Gasse 9, Vienna, Austria

[3] Research Institute of Molecular Pathology (IMP), Vienna BioCenter (VBC), Vienna, Austria

[#] Current address: Division of Immunobiology, Institute of Immunology, Center for Pathophysiology, Infectiology and Immunology, Medical University of Vienna, Vienna, Austria.

* Corresponding author information: markus.hartl@univie.ac.at

Key words: isobaric label, TMT, ion interference, ratio compression

Number of words: roughly **9300** (including M&Ms, without figure & table legends)

ABBREVIATIONS

DDA: Data dependent acquisition

DE: Differential expression

DIA: Data independent acquisition

FAIMS: High-Field Asymmetric-Waveform Ion Mobility MS

ID: Identification

IFI: Interference free index

LFQ: Label free quantification

MS: Mass spectrometry

MS2: MS/MS

MS3: MS/MS/MS

OIL: Observed interference level

PIC: Total peptide ion current

PIF: Precursor ion fraction

PPF: Precursor purity fraction

PSM: Peptide spectrum match

PTM: Post translational modification

ROC: Receiver operating characteristic

RTS: Real time search

TIC: Total ion current

TMT: Tandem mass tag

ABSTRACT

Multiplex proteomics using isobaric labeling tags has emerged as a powerful tool for the simultaneous relative quantification of peptides and proteins across multiple experimental conditions. However, the quantitative accuracy of the approach is largely compromised by a phenomenon termed ion interference, causing fold changes to appear compressed. The degree of compression is generally unclear, and the contributing factors are poorly understood. In this study, we thoroughly characterize ion interference at the MS2 level by using a defined two-proteome experimental system. We address the poor agreement between the apparent precursor purity in the isolation window and the actual level of interference in MS2-scans, a discrepancy that is only effectively resolved by considering potential co-fragmenting peptide ions hidden within the "noise". We further propose a comprehensive modeling strategy for obtaining accurate, feature-wise assessments of interference. Finally, we present a simple algorithm to calculate interference-corrected reporter ion intensities of peptides and proteins, thereby successfully removing ratio compression.

53 # INTRODUCTION

54 In mass spectrometry (MS) based proteomics, isobaric labeling systems like Tandem Mass Tags (TMT)
55 have been established as a widely used tool to efficiently quantify proteins across multiple conditions
56 within a single experiment[1]. Their high quantitative throughput is due to their ability to multiplex up to
57 18 samples simultaneously[2] using sample-specific isobaric labels that allow joint processing and
58 measurement of peptides. Upon fragmentation, sample-specific reporter ions are generated[3], disclosing
59 relative abundance of peptides and proteins. As all samples are measured in a single experiment, the
60 approach is largely free of missing values and also facilitates more complex sample preparation
61 strategies like offline-fractionation[1] or PTM enrichment[1,4].

62 Despite the elegance in design, multiplex proteomics struggles with accurate relative quantification
63 because of a phenomenon termed ion interference. Concomitant fragmentation of targeted precursor
64 peptide ions along with interfering non-precursor ions results in a mixture of different sources in
65 observed reporter ion signatures[5,6]. As the majority of proteins in a typical large-scale proteomic
66 experiment are non-differentially expressed, relative quantitative differences of precursor peptides
67 become suppressed by a uniform background of interfering reporter ion signal in all channels. This effect
68 is also referred to as "ratio compression". While simple approaches, such as increased sample
69 fractionation[7] or decreased isolation window widths[8], help to reduce the likelihood of such interference,
70 the problem still remains influential on the quantitative data, albeit to a smaller extent. The problem is
71 further exacerbated by the fact that the degree of compression varies for individual peptide
72 measurements. Interestingly, despite ratio compression, the ability to distinguish differentially from
73 non-differentially expressed proteins using isobaric labeling-based quantification is comparable or even
74 outperforms label free quantification (LFQ) approaches due to an increased precision and data
75 completeness[9]. However, compressed fold changes make it more difficult to interpret results in a
76 biologically meaningful way, especially when specific biological conditions result in the absence of
77 certain proteins. Therefore, different strategies have been developed to address and mitigate the ratio
78 compression problem.

79 Technical solutions encompass dual isolation width acquisition[8], High-Field Asymmetric-Waveform
80 Ion Mobility MS (FAIMS) [10] or MS3 measurements[11,12]. Unfortunately, these technical approaches
81 come with their own set of limitations as they require a specific technical setup for data acquisition[8,10–
82 12], and the measurement of additional scans per precursor generally decreases measurement speed and
83 therefore number of IDs[10–12]. Wühr *et al.* described an alternative quantification strategy in which the
84 complementary fragment ion cluster in the MS2 spectrum is used for the quantitative profiling[13].
85 However, most current MS instruments lack the required resolution to separate heavy N and C isotope
86 mass difference in the higher m/z region of the complementary ions which reduces the multiplexing
87 capabilities to a maximum of 8 channels[14]. Another hurdle lies in the complementary ion formation
88 itself, as currently available isobaric labeling systems are not optimized to generate complementary
89 fragment ions[15]

Finally, there are strategies that address ratio compression *in silico*. Savitski *et al*. propose interference correction based on the precursor purity in the isolation window of MS1 scans[16]. Niu *et al*. estimate the level of interference via K-y1 and R-y1 ion intensities in MS2-scans, ascribing them to either target or interfering peptides[7]. Other approaches propose to correct ratio compression by modelling the quantitative relationship between observed and theoretical fold changes provided by a synthetic peptide spike-in during sample preparation[6,17]. Similarly, modeling observed fold changes in dependence of precursor purities and signal to noise ratios for different peptides of the same protein has been proposed[18]. Altogether, these strategies take a rather phenomenological and atomistic approach in tackling the interference problem, since an assumed relationship between ion interference and a single, partly linked measured variable is leveraged to gauge the severity of ratio compression in an experiment. This simplified view of the problem can mainly be ascribed to the limited causal understanding of interference itself and all its contributing factors. As such, the proposed computational methods are not suited to fully account for ratio compression on the level of individual quantified peptides, which in turn limits their capacity in addressing ratio compression in an experiment as a whole.

In this study, we address ratio compression by building on a holistic mechanistic understanding of the underlying phenomenon itself, i.e. ion interference. To this end, we designed a controlled two-proteome experiment which allowed us to measure interference directly at the MS2 level for a substantial fraction of peptides, as well as to monitor the accompanied ratio compression effect on known theoretical fold changes. Taking advantage of this experimental design, we first explored the dependence of interference on commonly varying measurement parameters, as well as their impact on differential expression (DE) detection. We then aimed to causally describe the measured interference at MS2 level on the basis of linear regression modeling. We found that interference is only duly explained when accounting for "noise" at the MS1 level, as the apparent ion purity in the isolation window of target peptides generally underestimates the true extent of interference observed in MS2-spectra. Furthermore, by generalizing this comprehensive modeling approach, we reliably and accurately estimated the feature-wise level of interference in any MS2-quantified multiplex proteomics dataset - information that is otherwise hidden. Finally, we demonstrated how this gain in information can be used to calculate interference-corrected reporter ion intensities, and in consequence, to decompress fold changes affected by ratio compression.

119 # RESULTS

120 ## Experimental design

121 To investigate ion interference in detail, we designed an artificial two-proteome experiment comprising
122 TMTpro 16plex-labeled, tryptic peptides of budding yeast (W303-1A) and of human cells (Jurkat E6-
123 1) (**Fig 1A**). Our system allows direct measurement of interference for yeast peptides at the MS2 level
124 via the three reporter channels 128N, 129N and 130N (**Fig 1B**), which is reflected by total reporter ion
125 intensities of peptides after data normalization (**Fig 1C**). Further, the overall peptide mass ratio of human
126 to yeast peptides was set at approximately 10:1 (**Fig 1A**), ensuring comparatively little ion interference
127 from yeast peptides but a sufficient number of quantified yeast peptides to support the analysis. The
128 experiment also included one sample of excessive TMT label in absence of peptides, represented by the
129 channel 126C, to assess the contribution of chemical noise to ion interference. However, its signal was
130 negligible compared to the reporter ion signal in channels 128N, 129N and 130N (**Fig 1C**), showing that
131 ion interference and therefore ratio compression are generally not caused by quenched or hydrolyzed
132 TMT label molecules, but instead by other labeled peptide ions that are incidentally co-isolated and co-
133 fragmented during MS2-based quantification of target peptides.

134 ## Measuring interference at MS2 level

135 We first examined how ion interference is impacted by commonly varying sample- and measurement-
136 specific characteristics. In particular, we were interested in the factors which, all other things being
137 constant, alter the extent of measured interference at the MS2 level, such as sample complexity (i.e.
138 varying degrees of sample fractionation), isolation window width, quantification strategy (MS2,
139 FAIMS-MS2 and FAIMS-MS3), maximum MS2 injection time, injection amount and gradient length,
140 respectively. Ion interference was determined by measuring the fraction of reporter ion signal of yeast
141 peptides from yeast-free channels (128N, 129N, 130N) over all channels (see Observed Interference
142 Level (OIL) metric (**Equation 1**), Materials & Methods). Only yeast peptide-spectrum matches (PSM)
143 aggregated to contain unique feature combinations in sequence, modifications and charge state were
144 considered. Within a single comparison, yeast peptide features were further filtered for independent
145 quantification in all involved measurements, ensuring that results are not compromised by varying
146 depths of quantification or from the stochasticity inherent to data-dependent acquisition (DDA). Our
147 results indicate that reduced sample complexity (**Fig 2A**) and narrowed isolation windows (**Fig 2B**)
148 mitigate interference levels at the MS2 level, as previously established in literature[7,8]. We also confirm
149 a previously described decrease in interference levels for FAIMS-MS2 quantification over standard
150 MS2 quantification and almost completely interference-free TMT quantification via FAIMS-MS3 RTS
151 (real-time search)[10,11] (**Fig 2C**). Varying maximum MS2 injection times (**Fig 2D**), as well as varying
152 sample injection amounts (**Fig 2E**) yielded no notable change in interference at MS2 level. However,
153 our data suggest minimal interference for intermediate gradient lengths (**Fig 2F**). We attribute this

154 finding to the combination of changed elution profiles and the applied instrument settings, resulting in
155 fragmentation of peptides at more favorable elution times[19].

156 Next, we investigated how these changes in interference affect the detection of differential expression.
157 This is of particular interest since interference reduces the effect size of group differences which in turn
158 affects statistical hypothesis testing. We conducted pairwise differential expression testing via the
159 Limma-trend testing procedure[20] on the two sample groups 100:9 (channels 131C, 132C, 133C) and
160 100:12 (channels 132N, 133N, 134N), resulting in theoretical fold changes of 1.33 for yeast and 1.0 for
161 human peptides, respectively (**Fig 1A**). To prevent selection bias, our analysis continued to exclusively
162 focus on unique peptide features that were quantified within all measurements of a single comparison.
163 We used ROC curves to compare diagnostic abilities to correctly classify yeast peptide features as
164 differentially expressed versus human peptides as non-differentially expressed. (**Fig 2G - 2L**). Our
165 results show that ion interference is not the only factor governing reliable detection of differential
166 expression in isobaric labeling-based quantification. In addition, the reporter ion signal strength appears
167 pivotal for increasing measurement precision, resulting in reduced variance between replicate samples.
168 (**Supp Fig 1A**). In fact, we observed significant changes in within-group variances for all comparative
169 measurements (**Fig 2M - 2R**), which, when paired with observed interference levels, explain the
170 differences between ROC-curves (**Fig 2A - 2F**). For example, despite higher levels of interference in
171 the more complex samples (**Fig 2A**), increased precision due to higher levels of reporter ions
172 compensated the loss of accuracy (**Fig 2G** and **2M**). In other words, we observed that interference
173 determines a tradeoff between accuracy and precision, with both balancing each other out with respect
174 to their combined influence on differential expression detection. Importantly, in order to focus on the
175 effect of sample complexity in isolation, we did not scale up the total injection amount in the
176 measurements of reduced complexity via fractionation. Incidentally, an increase in the total sample
177 injection amount showed improved DE detection due to a gain in precision from the additional precursor
178 ion signal strength (**Fig 2K** and **2Q**). Therefore, performing fractionation which in turn enables an
179 increased total injection amount should prove highly beneficial for DE detection, while mitigating ratio
180 compression at the same time.

181 With regard to isolation window widths, our data indicate that increased precision from larger isolation
182 windows can outperform DE detection results from measurements with smaller windows (**Fig 2H,**
183 compare 0.7 Th and 0.4 Th), even though the latter is significantly less affected by interference (**Fig**
184 **2B**). Notably, in addition to a general higher transmission efficiency, windows larger than 0.7 Th also
185 manage to co-isolate +1 and -1 isotopes of triply charged precursor peptides, providing possible
186 explanations as to why a window of 0.7 Th classified markedly better than a window of 0.4 Th (**Fig**
187 **2H**). On the other hand, a window of 2.0 Th, which showed a median observed interference level above
188 0.5 (indicating that more than 50% of the reporter signal in more than half of yeast peptide measurements
189 were derived from human peptides), appeared to suffer in accuracy to an extent that the increase in
190 precision could not compensate (**Fig 2B**, **2H**, and **2N**).

191 The interpretation of ROC curves for comparisons between maximum injection times or gradient lengths
192 was complicated by the occasional difference in the number of PSMs per peptide feature. These
193 differences in PSMs arose from the associated changes in measurement speed or total measurement
194 time, respectively, which favor the redundant measurement of chromatographic peaks. Still, the
195 improved classification in measurements of intermediate gradient lengths is striking (**Fig 2L**), which
196 emphasizes the importance of reconciling the chromatographic separation of peptides with the MS
197 instrument settings that determine their targeting for MS2 quantification at specific points in their elution
198 profile for maximum accuracy and precision[19].

199 Finally, given our specific instrument settings (**Supp Table 1**), we observed FAIMS-MS2 and in
200 particular FAIMS-MS3 RTS to considerably outperform standard MS2 (**Fig 2I**). FAIMS-MS2 and
201 FAIMS-MS3 RTS could benefit from increased accuracy due to interference reduction (**Fig 2C**) without
202 losing much in precision (**Fig 2O**). Nevertheless, it is important to put these results into a broader
203 perspective. We have so far exclusively investigated shared and uniquely charged peptide features to
204 prevent bias in our analysis. However, variation in the addressed measurement parameters resulted in
205 vastly different numbers of unique quantified peptides (**Supp Fig 1B - 1G**). This shift in peptide
206 numbers naturally influences the quantitative results on protein level, where an increased number of
207 unique peptide identifications per protein is of advantage in DE detection. Altogether, maximizing
208 identifications, accuracy (i.e. minimizing interference) and DE detection at the same time is generally
209 not trivial to achieve, as there often exists a tradeoff between the three. The choice of a measurement
210 strategy is thus best guided by the available instrumentation, as well as the specific aims and limitations
211 of the individual experiment.

212 Modelling interference at MS2 level

213 In order to improve our understanding of ion interference on both a mechanistic and causal level, we
214 examined the interference-caused reporter ion signal in channels 128N, 129N and 130N in detail. We
215 speculated that ion interference would be best-explained by the specific ion composition in the precursor
216 peptide's isolation window, which should provide insight on the degree of co-isolation of interfering
217 peptide ions. This reasoning is not new. For example, the Precursor Ion Fraction (PIF) [21], implemented
218 in the quantitative proteomics software platform MaxQuant[22], and similar purity metrics estimate a
219 precursor purity by calculating the fraction of ion signal attributed to the precursor peptide (with respect
220 to the total ion signal) in the isolation window of bordering MS1 scans. Usually, these purity metrics
221 are used to filter out PSMs deemed too impure for reliable quantification. Here, we introduce a purity
222 metric termed PPF (Precursor Purity Fraction), which equally ranges from 0 to 1 by definition (**Equation**
223 **5**). Calculated PPF values correlated well with existing purity metrics (**Supp Fig 2A - 2C**), but
224 performed better in the modelling described below (see next section) due to subtle differences in
225 calculation (see methods).

However, regardless of the purity metric used, isolation window purity was a rather weak predictor for the actual level of interference observed in MS2 spectra. More specifically, interference at MS2 level was substantially underestimated, and even PSMs that appeared pure in their respective isolation window (i.e. PPF = 1) generally exhibited considerable interference at the level of reporter ions (i.e. OIL > 0) (**Fig 3A**). We attribute this discrepancy to how Orbitrap raw files are recorded during measurement. To minimize raw data file size and obtain clean spectra, signals below a certain threshold are considered "noise" and consequently removed. What remains as reduced information in Orbitrap raw file spectra are thus a) all peaks that possess intensity values above a local noise threshold; and b) noise values attributed to each centroided peak, which are proportional to the local noise threshold (personal communication with Thermo Fisher). Thus, we hypothesize that by removing these low-intensity signals, a substantial number of interfering labeled ions stays hidden at the MS1 level, especially when signal to noise ratios of target precursor ions are low. At the same time, noise values in the raw data could still sufficiently reflect the potential strength of the hidden signal. To explore these hypotheses, we employed linear regression modeling. The dependent variable in the linear model was defined as the spectrum-wise total observed interference at MS2 level, calculated as the summed reporter ion signal in channels 128N, 129N and 130N of yeast peptide measurements. In line with our previous observation, a model that does not account for noise performed poorly in predicting interference at MS2 level (**Fig 3B, upper**). We therefore adapted the model formula accordingly by including a regressor variable that expresses the potential noise contribution. Assuming that interfering ions hidden within the noise are in fact other labeled peptide ions of relatively low abundance, we calculated this variable by combining information drawn from two different sources: first, the estimated maximum amplitude of such ions, believed to be directly proportional to the noise values in the spectrum; and second, the estimated frequency of such ions, which was calculated from measurement-specific empirical peptide densities that span the m/z-retention time plane (**Supp Fig 2D**) (see methods). The updated model (**Equation 7**) managed to faithfully predict the interference-caused reporter ion signal for all yeast peptide measurements, including PSMs that appear pure in their respective isolation window (**Fig 3B, lower)**.

Estimating interference at MS2 level

Accurately estimating ion interference in MS2-quantified multiplex proteomics could prove extremely valuable for correct qualitative and quantitative interpretations of experimental results. Regression modeling has the potential to achieve this goal - however, the specific implementation above requires direct measurements of interference at MS2-level to support model training. In a typical, research-oriented multiplex proteomics experiment this will not be the case. We therefore set out to adequately adjust the above model in order to accommodate data from any MS2-quantified multiplex proteomics experiment. With that in mind, we first redefined the dependent variable of the model as the *total* measured reporter ion intensity in MS2 spectra. Importantly, most of the reporter ion signal is expected to originate from the targeted precursor peptide ion itself. It is therefore vital to have the contribution of

263 precursor peptide ions reflected as an additional regressor variable in the model, thereby allowing us to
264 attribute the remaining, unaccounted signal to interference. However, this task was complicated by
265 drastic differences in generated total reporter ion intensity (**Fig 3C**), which we ascribe to variable
266 fragmentation efficiencies for distinct peptide classes. These classes differed in properties like charge
267 state, the number of TMT labels, as well as the presence or absence of potentially charged amino acids.
268 Moreover, they also varied in frequency and composition between experiments due to what we assume
269 are predominantly instrument- and measurement-specific settings (e.g. normalized collision energies,
270 chromatography buffer properties, etc.). To account for this heterogeneity in fragmentation behavior, as
271 well as for differences between individual datasets, we came up with a decision tree algorithm to infer
272 distinct empirical peptide classes with similar fragmentation efficiencies from the experimental data
273 itself (**Fig 3D**). Empirical classes determined this way were then incorporated as separate regressor
274 variables into the linear model. Hence, differences in peptide fragmentation efficiencies could be
275 accommodated by distinct model parameters that describe the unique class-specific contribution of
276 precursor ion intensities to the total reporter ion intensity.

277 In most datasets, we observed a slight non-linear relationship between peak intensities in the MS1
278 isolation window and the corresponding total reporter ion intensities at MS2 level. This observation is
279 likely due to the circumstance that all reported intensities in Orbitrap spectra are subject to a non-trivial
280 transformation and scaling of measured intensities, which includes a spectrum-wise subtraction of global
281 noise averages, as well as a non-linear intensity scaling that aims to counteract the equally non-linear
282 Orbitrap signal response (personal communication with Thermo Fisher Scientific). While these signal
283 processing steps are generally desired, we reason that a perfectly linear dependence of MS2-intensities
284 on MS1 precursor intensities cannot always be guaranteed. Therefore, instead of explaining the total
285 reporter ion intensity at MS2 level by the signal in the MS1 isolation window, our model was built on
286 the total peptide fragment ion-intensity recorded in *the exact same* MS2-scan. This relationship proved
287 to be always linearly correlated since both measures derived from the same spectrum. In calculation, the
288 total peptide fragment ion intensity equals the total ion current (TIC) of the MS2-scan minus the total
289 reporter ion intensity, hence we abbreviated it to "total peptide ion current", or PIC in short. The PIC
290 was then further divided into interference- and non-interference-caused signal based on the isolation
291 window-derived precursor purity fraction (PPF), similar to how the signal in the isolation window would
292 be divided into precursor and non-precursor-derived ions. While this approach is arguably less intuitive
293 for explaining the mechanism behind reporter ion generation, we found it perfectly resolves non-linear
294 dependence. Thus, we arrived at a holistic linear model regression formula that comprehensively
295 describes the total reporter ion intensity, accounting for contributions from both precursor as well as
296 interfering ions, which is applicable to any MS2-quantified multiplex proteomics experiment (**Equation**
297 **8**).

298 Training this model on our experimental data revealed a trend of continually improved prediction results
299 by including more of the relevant predictor variables (**Fig 3E**). To exploit this explanatory power, we
300 then used the model parameter estimates to infer the ion purity in individual MS2 scans, expressed as

301 "estimated interference level" (EIL, **Equation 9**). This metric is essentially another PSM-wise purity
302 metric, with the difference that it also considers the contribution of noise in ion interference, which
303 promises increased accuracy for assessing ion purity. Indeed, unlike conventional purity metrics, we
304 observed EIL to accurately estimate the fraction of interference in yeast peptide scans, with further
305 increasing precision when aggregated to peptide or protein level (**Fig 4A**). The clear improvement over
306 standard purity metrics can be explained by the fact that the fitted model ascribes about half of the
307 estimated interference to "noise", which is not directly inferable from the MS1 isolation window.

308 To validate the robustness of model-fitting and subsequent EIL estimation, we applied our modelling
309 strategy to previously published datasets on MS2-quantified TKO9 and TKO11 proteomic standards[23,24].
310 These commercial TMT-based standards each comprise three multiplexed congenic yeast KO strain
311 proteomes. Hence, these datasets allow direct observation of interference at MS2 level for peptides
312 corresponding to the deleted genes. Paolo *et al*. [23] took advantage of this property and calculated the so-
313 called "interference-free index" (IFI), which translates to 1-OIL in calculation. Here, we compared OIL
314 values to model-based EIL values that have been calculated entirely on the basis of the TKO data (**Supp**
315 **Fig 3A** and **3B**). The results emphasize the robustness of model-based EIL estimation, and its advantage
316 over standard purity metrics. Moreover, we also applied the same procedure to FAIMS-MS2
317 quantification data, and found it compatible by treating each compensation voltage within a run as its
318 own separate measurement (**Supp Fig 3C**).

319

320 Interference correction

321 Equipped with accurate feature-wise estimates of interference, we addressed ratio compression. A
322 previously described computational approach calculated interference-corrected reporter ion intensities
323 by a spectrum-wise subtraction of interference-caused reporter ion signal, estimated from the ion purity
324 in the isolation window, from the measured reporter ion signal[16]. We adapted this approach using our
325 model-based EIL, along with other minor changes (see methods), to correct for ratio compression.

326 The performance of interference correction was subsequently evaluated on the basis of known
327 theoretical fold changes of proteins in our yeast-human mixture experiment. The group differences
328 comprised expected yeast protein ratios of 2.0, 1.5, and 1.33, resulting from the pairwise group
329 comparisons between groups 100:12 vs 100:6, 100:9 vs 100:6, and 100:12 vs 100:9, respectively (**Fig**
330 **1A**). Conversely, human proteins carried theoretical fold changes of 1 throughout all group comparisons,
331 mimicking proteins of unchanged expression levels. The effect of interference correction is illustrated
332 as volcano plots resulting from DE testing (**Fig 4B - 4D**), highlighting two key aspects of our approach.
333 First, the correction strategy updated with the model-based EIL metric considerably improves the
334 accuracy of calculated yeast protein fold changes - not just compared to uncorrected intensities, but also
335 compared to intensities stemming from the original correction approach that considers traditional purity
336 metrics for estimating interference (here: PPF-corrected). Notably, for this analysis, PSMs were not
337 filtered on arbitrary intensity and/or signal to noise thresholds. Hence, the model-based interference

338 correction approach provided an accuracy comparable to MS3-quantification but without the
339 accompanied loss in peptide and protein IDs. Second, as the accuracy increased, we observed an equal
340 loss in precision, particularly concerning the non-differentially expressed human proteins, reflected by
341 the increased spread of fold changes around their expected value. Interestingly, this observation matches
342 the already addressed tradeoff between accuracy and precision that a change in interference entails -
343 however, here this tradeoff is entirely brought about computationally. Similarly, ROC-analyses illustrate
344 that despite the drastic difference in measurement accuracy, the classification results showcase widely
345 unchanged differential expression detection results (**Fig 4E - 4G**). Collectively, these results support
346 that a model-based interference correction strategy presents a viable alternative to MS3-quantification
347 for addressing ratio compression in multiplex proteomics.

348

349

350 # DISCUSSION

351 Ion interference in multiplex proteomics has long been investigated as the cause of ratio compression,
352 yet without producing a thorough characterization and clear mechanistic explanation of ion interference
353 as the underlying phenomenon itself. In this study, we sought to investigate ion interference
354 comprehensively by using an artificially created experimental system specifically designed for this task.
355 We first elucidated the dependence of interference on key measurement parameters, as well as the
356 associated impact on DE detection. Our data illustrate that, as a consequence of increased signal strength,
357 interference leads to a substantial gain in precision that largely negates the effect of reduced accuracy
358 in statistical testing.

359 Our main experimental aim, however, was to clearly identify and formulate the contributing factors that
360 explain the interference-caused reporter ion signal at MS2 level. We addressed this problem by utilizing
361 linear regression modelling. Crucially, our analysis suggests that labeled peptide ions hidden in the
362 "noise" are absolutely essential in explaining interference, as a sizable fraction of measured interference
363 at MS2 level could not be explained by visible ion peaks within target peptides' isolation windows.
364 Moreover, we could confidently rule out any major contribution of hydrolyzed or quenched TMT label
365 in the sample. The observation that FAIMS-MS2 quantification exhibits reduced, yet still ample
366 interference further supports the notion that interfering ions predominantly consist of multiply charged
367 peptide ions.

368 Based on this mechanistic understanding of ion interference, we were able to generalize the linear
369 regression modelling approach and make it applicable to any MS2-quantified multiplex proteomics
370 dataset. This allowed us to faithfully separate precursor and interference-caused reporter ion
371 contributions, and consequently to obtain accurate estimates of interference at MS2 level. However, two
372 generalizations were made in the process. First, we assumed the isolation efficiency to be perfectly
373 uniform across the whole isolation window range. While the trend in instrument development certainly
374 advances in that direction, in reality the isolation efficiency describes a nonlinear function that peaks in
375 the center of the window[25] and that varies from instrument to instrument, potentially even between
376 measurements. However, since the deviation from uniformity is rather mild, we do not think that this
377 simplification should compromise the modelling approach. Alternatively, one could try to
378 experimentally determine the isolation efficiency for a given instrument[25] or infer it from the data itself,
379 to further refine the modeling. The second generalization we made stipulates that every ion peak at MS1-
380 level corresponds to a (labeled) peptide ion. This assumption neglects the possibility of singly charged,
381 non-peptide ion contamination which does not generate reporter ions upon fragmentation.
382 Computationally removing singly-charged signals from all MS1-spectra could address this issue, yet we
383 think that fractionation or FAIMS should provide sufficient removal of such ions in cases where
384 contamination poses a problem.

385 The majority of isobaric labeling experiments are so far performed on Orbitrap FT-mass spectrometers
386 due to constraints on instrument resolution that limit the multiplexing capabilities. In future, the

387 technology might well expand to other mass spectrometers, e.g. TOF (time of flight) instruments. These
388 instruments fundamentally differ in signal detection and processing[26], particularly with respect to noise,
389 and it will be interesting to see how this affects ion interference. Consequently, the presented modelling
390 strategy will have to be adapted accordingly. Moreover, ion mobility separation technology other than
391 FAIMS, as for example the trapped ion mobility cell in the timsTOF (Bruker), might efficiently remove
392 ion interference and provide more accurate quantitative data[27].

393 Finally, we illustrated the potential of interference correction on the basis of model-based interference
394 estimates. However, a limitation to be considered concerns the assumption that most proteins in the
395 experiment are similarly expressed, thereby causing interference to evenly affect all reporter channels
396 in the experiment. While this premise is generally reasonable, it might not be universally justifiable.
397 Thus, when differences between multiplexed samples are large, an alternative correction strategy might
398 instead involve filtering and/or differential weighting of individual protein's PSMs based on confidence
399 in measurement accuracy, e.g. given by ion purity metrics calculated from the isolation window. Such
400 strategies have already been discussed and implemented by others[16,28,29], and we think that accurate
401 model-based interference estimates as presented here are ideally suited to enhance their performance. It
402 should be noted though that any strict filtering of PSMs will always coincide with a loss in the number
403 of identifications.

404 In summary, our work provides in-depth insights into the nature of ion inference and provides directions
405 for minimizing interference and optimizing DE testing in multiplex proteomics data generation. We
406 have further introduced the concept of linear regression modelling as a means to assess and combat ratio
407 compression in-silico. On that basis, we believe that this work improves the general understanding of
408 this fundamental problem in multiplex proteomics, and guides researchers towards effectively coping
409 with ion interference and ratio compression in future experiments.

410

411

412

413 # MATERIALS AND METHODS

414 ## Cell culture

415 *S. cerevisiae* yeast cells W303-1A were cultured shaking (200 rpm) at 30°C until mid-log phase
416 (OD600nm ~1) in 50 ml rich medium (YPD; 1% yeast extract, 2% peptone, 2% glucose). Cells were
417 harvested by filtration (Protran 0.45 µm nitrocellulose membrane, Amersham) and frozen in liquid
418 nitrogen. Human Jurkat E6-1 cells were cultured in Roswell Park Memorial Institute (RPMI) 1640
419 media (Fibco, Grand Island, NY), supplemented with 10% fetal bovine serum (Gibco), penicillin and
420 streptomycin (Gibco). Cells were then pelleted and washed with 1x phosphate buffered saline (PBS)
421 prior to freezing in liquid nitrogen.

422

423 ## Proteomic sample preparation

424 Frozen yeast cell pellets were resuspended in lysis buffer (8 M urea, 50 mM Tris buffer pH8, 150 mM
425 NaCl, 1 mM PMSF, 5 mM sodium butyrate, benzonase and protease inhibitor cocktail) and subsequently
426 lysed by bead-beating (FastPrep®-24 device, MP Biomedicals). Frozen Jurkat cells were dissolved in
427 lysis buffer and lysed via sonication. After centrifugation at 16,000 x g at 4°C for 10 min, yeast and
428 human crude extract's supernatants were transferred into fresh tubes, and proteins in solution were
429 precipitated by the addition of three times the volume of cold acetone and subsequently incubated
430 overnight at -20°C. Following centrifugation at 15,000 x g and the subsequent removal of acetone,
431 protein pellets were dissolved in 8 M Urea, 50 mM ammonium bicarbonate (ABC) buffer and treated
432 with 10 mM DTT (Dithiothreitol) for 45 min to reduce protein disulfide bonds. Subsequent alkylation
433 of reduced cysteines was performed by adding iodoacetamide (IAA) to a final concentration of 20 mM,
434 followed by incubation for 30 min in the dark. The remaining IAA was quenched and samples were
435 diluted to 4 M urea for digestion with Lys-C 1:50 (w/w). After incubation for 2 h at room temperature,
436 samples were further diluted to a final concentration of 1 M urea, and digestion with trypsin 1:50 (w/w)
437 was carried out overnight at 37°C. Finally, tryptic peptides were desalted using SepPak C18 cartridges
438 (Waters), and eluates were vacuum centrifuged for 30 min prior to lyophilization.

439

440 ## Two-Proteome controlled experiment

441 Lyophilized tryptic yeast and human peptides were reconstituted in 0.1% trifluoroacetic acid (TFA) to
442 reach a peptide concentration of 3 µg/µl each. 4 distinct artificial two-proteome mixtures were created
443 as follows: A constant amount of human peptide solution (300 µl) was initially prepared for each
444 mixture. Then, a comparatively small, variable amount of yeast peptide solution was added to each
445 solution, namely 0, 18, 27 and 36 µl, respectively; resulting in theoretical fold changes of 1.33, 1.5 and
446 2.0 between yeast peptides. The yeast-human mixtures were subsequently lyophilized and again
447 reconstituted in 120 µl 100 mM TEAB (Sigma) each. Next, each mixture was split into 3 technical
448 replicates, resulting in 12 peptide samples in total ready for isobaric labeling. TMT labeling was

449 conducted according to a labeling scheme that minimizes isotopic interference across groups (**Fig 1A**).
450 Additionally, a thirteenth sample without peptides was labeled in parallel to interrogate the effect of
451 hydrolyzed or quenched label on ion interference (**Fig 1A**).

452

453 ## TMT-labeling

454 Dried TMTpro 16plex 0.5 mg label reagents (Thermo Fisher) were taken up in 30 µl anhydrous
455 acetonitrile (ACN) and combined with 30 µl 100 mM TEAB peptide solution, resulting in a ratio of
456 TMT label to peptide of approximately 2:1 (w/w); and a final peptide concentration of about 4.2 µg/µl
457 per 60 µl labeling reaction. After incubation for 1 h, a small aliquot of each reaction was combined in
458 100 µl of 0.1% TFA. This pooled sample was subsequently measured on the mass spectrometer to
459 ascertain a labeling efficiency of >99% in each channel by using an in-house R script for the analysis.
460 The labeling reactions were then quenched by addition of 7 µl of 5% hydroxylamine and incubation for
461 25 min at 400 rpm at RT. Finally, all 13 labeled samples were combined 1:1, desalted via SepPak C18
462 cartridges, vacuum centrifuged and lyophilized prior to high pH fractionation.

463

464 ## High pH fractionation

465 Dried labeled peptides were reconstituted in 100 µl 10 mM ammonium formate pH 6.8. After sonication,
466 the solution was injected into UltiMate 3000 Dual LC pHPLC System equipped with a C18 column
467 (xBridge Peptide BEH C18, 25 cm x 4.6 mm, 3.5 µm, Waters). 1 ml fractions were collected during
468 separation on a 5-50% ACN 1 ml/min gradient in 10 mM ammonium formate buffer pH 6.8 at a flow
469 rate of 1 ml/min. The resulting 1 ml fractions were vacuum centrifuged to evaporate the ACN, and then
470 pooled according to a cyclical pooling scheme to create 6 fractionation pools of reduced complexity.
471 The pools were desalted via SepPak C18 cartridges and stored at -80°C until measurement on the LC-
472 MS system.

473

474 ## Measurement on the LC-MS system

475 The main results of this study (**Fig 1, Fig 3-4**) draw on raw file data generated by measuring yeast-
476 human mixture samples of reduced sample complexity, quantified at the MS2 level. In brief, for each of
477 the 6 pools after fractionation, a total of 200 ng peptide material was measured on a Q Exactive HF-X
478 Orbitrap mass spectrometer (Thermo Fisher), coupled to the LC-system with a nano-spray ion-source
479 using coated emitter tips (PepSep, MSWil). Peptides were separated along a segmented 2 h 2-40% ACN
480 gradient on the Ultimate 3000 RSLC nano-flow chromatography system (Thermo-Fisher) using a pre-
481 column for sample loading (Acclaim PepMap C18, 2 cm × 0.1 mm, 5 µm, Thermo-Fisher), and a C18
482 analytical column (Acclaim PepMap C18, 50 cm × 0.75 mm, 2 µm, Thermo-Fisher) for separation, at a
483 flowrate of 230 nl/min. On the mass spectrometer, the following instrument settings were applied. For

484 MS1 scans: resolution 120 k, maxAGC 1e6, maximum IT 60 ms; for MS2 scans: resolution 45 k,
485 maxAGC 2e5, maximum IT 120 ms, respectively. Precursor ions were targeted for fragmentation in a
486 top 20 method using a dynamic exclusion time of 20 s. Target ions were isolated at an isolation window
487 width of 0.7 Th and fragmented at NCE of 35. Precursor ion charge states of 1 or >7 were excluded. All
488 spectra were recorded in profile mode.

489 For the comparative measurements of varying measurement-specific parameters (**Fig 2**), samples were
490 measured on the Q-Exactive HF-X (Thermo Fisher) and Orbitrap Eclipse Tribid mass spectrometers
491 (Thermo Fisher) as described in detail in **Supp Table 1**, unless specified otherwise due to the
492 measurement parameters being tested.

493

494 Database search and raw file data extraction

495 Raw files were searched with MaxQuant[22] (version 1.6.14) against a concatenated yeast and human
496 database (Uniprot release 2020.01, canonical sequences) and a common contaminants database, while
497 allowing up to two missed cleavages on tryptic peptides. Search parameters considered oxidation of
498 methionine as well as N-terminal acetylation as variable modifications, and included
499 carbamidomethylation of cysteine residues as a fixed modification. TMTpro label modifications on N-
500 termini and lysine residues were searched as either variable or fixed. The results were filtered to reach
501 an expected FDR of 1% at the PSM-level. Separately, spectral features like m/z, intensity and noise
502 values of centroided peaks were extracted directly from Thermo raw files via an in-house command line
503 application written in C# that uses the Thermo's RawFileReader library. The combined results were
504 imported into the R statistical environment and further analyzed there.

505 In addition to MaxQuant, part of the data were searched using FragPipe[30] v16.0 as well as the Proteome
506 Discoverer™ v2.5.0.400 software platform (Thermo Fisher) with MSAmanda[31] v2.0.0.16129. Details
507 are described in the supplemental material.

508

509 Extraction of reporter ion intensities and isotopic impurity correction

510 Reporter ion intensities were retrieved from the raw data by taking the maximum observed peak intensity
511 value within a $\pm$ 0.002 Th window at the expected reporter ion m/z. For this task, we leaned on the code
512 of the R proteomics software MSnbase[32] and utilized the reporter ion classes pre-defined therein.
513 Further, we used linear algebra to correct for TMT isotopic impurities as implemented in MSnbase[32].
514 The impurity matrix was calculated manually from the product sheet provided by the manufacturer (lot
515 number VC294906).

516

Estimation of Run-Specific Peptide Density

Retention time and m/z values of all identified, uniquely charged peptide features were used to construct a run-specific two-dimensional kernel density estimate, evaluated on a 200 x 200 grid. The calculation employed the kde2d function from the R package MASS[33]. Resulting density values were rescaled to sum up to 1. Further, the square root of all density values was taken to account for extreme differences in the m/z and retention time plane, and density values of 0 were substituted with the minimum calculated positive density. Finally, the 200 x 200 density grid was interpolated for each individual PSM at the precursor-specific m/z and retention time coordinates (**Supp Fig 2D**).

Calculation of relevant variables

OIL The "observed interference level" (OIL) denotes the PSM-specific fraction of the interference-caused reporter ion signal with respect to the total reporter ion signal measured at MS2 level. It is calculated as the ratio between the average normalized reporter ion intensity of channels where no quantitative signal is expected (but nonetheless observed due to interference), and the average normalized reporter ion intensity of all other channels. To ensure values between 0 and 1, calculated OIL values above 1 are capped at 1. For yeast PSMs in the yeast-human mixture experiment (**Fig 1A**), the calculation equaled

$$OIL = min\left\{\frac{\mathrm{Avg}\left(\mathrm{I}_{128\mathrm{N}}, \mathrm{I}_{129\mathrm{N}}, \mathrm{I}_{130\mathrm{N}}\right)}{\mathrm{Avg}\left(\mathrm{I}_{128\mathrm{C}}, \mathrm{I}_{129\mathrm{C}}, \mathrm{I}_{130\mathrm{C}}, \mathrm{I}_{131\mathrm{C}}, \mathrm{I}_{132\mathrm{C}}, \mathrm{I}_{133\mathrm{C}} \mathrm{I}_{132\mathrm{N}}, \mathrm{I}_{133\mathrm{N}}, \mathrm{I}_{134\mathrm{N}}\right)}, 1\right\}, \quad (1)$$

where $I_c,\ c \in \{126C, 127N, \dots, 134N\}$ is the normalized reporter ion intensity of channel c.

PPF The "precursor purity fraction" (PPF) denotes the PSM-specific fraction of signal stemming from precursor peptide ions (including isotopes) with respect to the total measured signal in the isolation window of bordering MS1 scans. It is very similar to existing ion purity metrics like PIF[21] (**Sup Fig 2A-C**). First, all centroided peak intensities $I_{ip},\ i \in \{-1, +1\},\ p \in \{1, 2, \dots, k_i\}$ found in the isolation windows of the two bordering MS1 scans indexed $i,\ i \in \{-1, +1\}$ are extracted, where k_i denotes the total number of distinct peaks in the isolation window of MS1-scan i. It is then checked if the precursor ion peak can be found among the peaks in the preceding MS1 scan ($i = -1$) within a margin of error of $\pm$ 0.0025 Th. If the check is unsuccessful, the information of the second last MS1 scan is used instead, provided that the precursor ion peak is present there (this was typically the case). We found this situation to occur more frequently for measurements on Exploris and Eclipse instruments than on HF-X instruments. If the precursor ion peak is still not found, its intensity is imputed at the declared precursor m/z by taking the minimum intensity value over all observed peaks in the preceding MS1 scan ($i = -1$). Concerning following MS1 scan ($i = +1$), if the respective isolation window is entirely empty, the scan is skipped and not used in the calculations further downstream. Else, if the isolation window is non-empty but the precursor ion peak is not found at the declared precursor m/z ($\pm$ 0.0025 Th), the precursor peak intensity is imputed as just described. However, in the rare event that the precursor peak cannot be

found in either bordering MS1 scans, the following MS1 scan is also skipped $i = +1$). Next, the peak intensities $I_{ip},\ i \in \{-1, +1\},\ p \in \{1, 2, \dots, k_i\}$ in the isolation windows of bordering MS1 scans are inversely weighted by their absolute distance in retention time t_i to the MS2-scan that produced the PSM. This results in weighted intensities I^*_{ip}:

$$d_i = |\, t_i - t_{MS2} \,|, \quad i \in \{-1, +1\} \tag{2}$$

$$w_i = 1 - \frac{d_i}{\sum_i \mathrm{d_i}}, \quad i \in \{-1, +1\} \tag{3}$$

$$I^*_{ip} = w_i \cdot I_{ip}\ , \quad i \in \{-1, +1\},\ p \in \{1, 2, \dots, k_i\} \tag{4}$$

Further, heavy (+1 Da) and light (-1 Da) isotope peaks of the precursor peptide are determined using a mass error tolerance of $\pm$ 0.00125 Th. This results in the subset $I^*_{precursor}$ containing all weighted peak intensities I^*_{ip} that can be ascribed to the precursor peptide, including heavy and light isotope ions, detected in the isolation window of bordering MS1 scans. The precursor purity fraction PPF is then defined as

$$PPF = \frac{\sum_i \sum_p I^*_{\mathrm{ip}} \cdot \mathbb{1}\left(I^*_{\mathrm{ip}}\right)}{\sum_i \sum_p I^*_{ip}}, \qquad \mathbb{1}\left(I^*_{\mathrm{ip}}\right) = \begin{cases} 1 & \text{if } I^*_{ip} \in \mathrm{I}^*_{\mathrm{precursor}} \\ 0 & \text{if } I^*_{ip} \notin \mathrm{I}^*_{\mathrm{precursor}} \end{cases},\ i \in \{-1, +1\},\ p \in \{1, 2, \dots, k_i\} \tag{5}$$

TIW The "total intensity in the isolation window" (TIW) is defined as

$$TIW = \sum_i \sum_p I^*_{ip}\,, \quad i \in \{-1, +1\},\ p \in \{1, 2, \dots, k_i\} \tag{6}$$

PIC The "total peptide ion current" (PIC) is calculated as the total ion current (TIC) of an MS2 scan minus the total reporter ion intensity of the same scan. This metric therefore equals the sum of all peak intensities stemming from peptide fragment ions in MS2 scans.

Multiple linear regression modelling

We employed linear regression modelling to explain the interference-caused reporter ion signal in our yeast-human mixture experiment. The model formula is

$$Y \sim 0 + nonprecursor{:}rawfileCharge + noiseEstimate, \tag{7}$$

where:

Y is the dependent variable of the model and denotes the total reporter ion intensity measured in channels 128N, 129N and 130N of yeast PSMs.

$nonprecursor = TIW \cdot (1 - PPF)$. This predictor variable denotes the total intensity of non-precursor interfering ions in the isolation window.

$rawfileCharge$ is a categorical variable denoting the charge state of the precursor ion as reported by the mass spectrometer and documented in the raw file. Charge states above 3 were set to 3 to ensure sufficient numbers of observations in each category.

583 $noiseEstimate$ is a PSM-wise estimate of the signal of non-observable interfering ions in the isolation
584 window. It is calculated as the product of the average reported noise value of peaks in the isolation
585 window of the preceding MS1 scan, and the interpolated run-specific peptide density at the precursor-
586 specific m/z and retention time coordinate.

587 Further, "0" denotes an intercept fixed at 0, and the colon denotes interaction effects between predictor
588 variables. The model parameters coefficients were estimated via robust linear regression modelling with
589 bisquare weighting using the rlm function from the R package MASS[33]. This robust estimation
590 procedure accounts for outliers as well as extreme heteroscedasticity. The model was trained separately
591 on data of each individual measurement prior to normalization.

592 Moreover, we propose an extended linear model that comprehensively explains the total reporter ion
593 intensity in MS2-scans in any MS2-quantified multiplex proteomics experiment. The model formula is

594 $$Y \sim 0 + precursor{:}pepChar + nonprecursor{:}rawfileCharge + noiseEstimate, \quad (8)$$

595 where:

596 Y is the dependent variable of the model and denotes the total reporter ion intensity in MS2-scans.

597 $precursor = PIC \cdot PPF.$

598 $pepChar\ is$ a categorial variable that specifies empirical classes of specific combinations of observed
599 precursor peptide characteristics (**Fig 3D-E**). It is calculated using a greedy classification algorithm that
600 continuously creates binary splits in peptide characteristic predictor variables in order to better explain
601 observed differences in the fragmentation efficiency of precursor peptides.

602 $nonprecursor = PIC \cdot (1 - PPF).$

603 $rawfileCharge$ and $noiseEstimate$ are defined as above.

604 The model was trained separately on the data of each individual measurement run prior to data
605 normalization between samples. As before, parameter estimates were calculated via robust linear
606 regression modelling with bisquare weighting.

607

608 Calculation of the estimated interference level (EIL)

609 Training a linear regression model (Equation 8) on MS2-quantified multiplex proteomics data results in
610 estimated model parameters $\hat{\beta}$ that can be used to predict the PSM-wise level of interference in MS2-
611 scans. We define the "estimated interference level" or EIL as

612 $$\text{EIL} = min\left\{1 - \frac{x_{precursor}^{T}\hat{\beta}_{precursor}}{x^{T}\hat{\beta}}\ ,\ 1\right\}, \quad (9)$$

613 where:

$\hat{\beta} = (\hat{\beta}_1, \hat{\beta}_2, \ldots, \hat{\beta}_k)^T$ is the k x 1 vector of estimated parameter coefficients,

x is a k x 1 vector containing the PSM-specific realizations of predictor variables,

$x^T\hat{\beta}$ is therefore the fitted value, i.e. the PSM-specific model-estimated total reporter ion intensity;

$\hat{\beta}_{precursor} = (\hat{\beta}_1, \hat{\beta}_2, \ldots, \hat{\beta}_p)^T$ is a p x 1, p < k subset of $\hat{\beta}$ containing only the coefficients that describe the contribution of the precursor peptide to Y,

$x_{precursor}$ is the corresponding p x 1 , p < k subset of x, such that

$x_{precursor}^T \hat{\beta}_{precursor}$ is the PSM-specific model-estimated total reporter ion intensity generated by precursor peptide ions, including heavy and light isotopes.

Interference correction

Similar to Savitski *et al.*[16], we propose a PSM-wise interference correction algorithm that generates interference-corrected reporter ion intensities from the estimated interference at MS2 level. The major differences to the original approach lie in using a different kind of impurity metric for the calculation - the estimated interference level (EIL) - instead of an impurity metric that is derived from the isolation window purity like 1-PPF; as well as substituting any resulting corrected reporter ion intensity values with the minimum respective MS2 peak intensity value, in case they should fall below that threshold. In more detail, the calculations for our yeast-human mixture experiment were

$$\bar{I} = \frac{1}{|C|}\sum_c I_c, \quad c \in C = \{128C, 129N, \ldots, 134N\}, \tag{10}$$

where I_c denotes the untransformed, normalized reporter ion intensity of reporter channel c, and $\bar{I}$ thus equals the average normalized reporter ion intensity. Importantly, missing values as well as empty channels were excluded in this calculation. Then, corrected normalized reporter ion intensities $I_c^{corrected}$ were calculated as

$$I_c^{corrected} = I_c - \bar{I} \cdot min\{\text{EIL}, \ 0.8\}. \tag{11}$$

In this formula, EIL values are capped at an arbitrary cutoff of 0.8 to mitigate overcorrection. Finally, corrected normalized reporter intensities $I_c^{corrected}$ that reach below the minimum observed MS2 peak intensity value were substituted with that value to mitigate artefacts from overcorrection. Missing reporter intensity values, albeit rare, were substituted equally.

Data Transformation, normalization and statistical testing

Reporter ion intensities were $\log_2$-transformed and subsequently normalized using a cyclic LOESS normalization strategy contained in the R package Limma[34]. However, prior to calculating OIL values

645 (Equation 1) and interference-corrected reporter ion intensities (Equation 11), $\log_2$-transformed
646 normalized reporter ion intensities were retransformed from log-space by applying the corresponding
647 inverse function on the transformed data (i.e. exponentiation with base 2). Statistical testing for
648 differential expression between groups was conducted via the Limma-trend testing procedure[20] on $\log_2$-
649 transformed normalized reporter ion intensities.

650 # FIGURE LEGENDS

651 **Figure 1 - Overview of artificial two-proteome experiment designed to investigate ion interference.**
652 **(A)** Schematic depiction of sample composition and TMTpro 16plex labeling scheme. Samples were
653 defined mixtures of *S. cerevisiae* (violet) and *H. sapiens* Jurkat (grey) peptides from whole-proteome
654 digests. Relative mass abundance ratios (w/w) of human to yeast are displayed on top of each sample
655 group. **(B)** Illustration of how human-caused ion interference was measured at the MS2 level via
656 channels 128N, 129N and 130N for a targeted *S. cerevisiae* peptide. (**C)** Calculated total channel
657 intensities of quantified yeast (left) and human (right) peptides after between-sample normalization.

658 **Figure 2 - Dependence of measurement parameters on ion interference and differential expression**
659 **detection. (A-F)** Boxplots showing distributions of observed interference levels (OIL) for yeast peptide
660 features independently quantified in all conditions within a comparison. P-values denote statistical
661 significance for overall differences between conditions (Friedman test). (A) Varying sample
662 complexities. Numbers 1, 1/3 and 1/6 correspond to factors of sample complexity reduction via
663 fractionation – e.g. 1 implies just one measurement run of the unfractionated sample, while 1/6 implies
664 six measurement runs in total for the six fractionated parts of reduced sample complexity. (B) Varying
665 gradient lengths, in hours (h). (C) Varying injection amounts, in nanogram (ng). (D) Varying isolation
666 window widths, in Thomson (Th). (E) Varying quantification strategies. MS3 measurements were also
667 real time-searched (RTS). (F) Different maximum injection times, in milliseconds (ms). **(G-L)** ROC-
668 curves calculated from independently quantified peptide features. Measurements are the same as shown
669 in (A-F). The true positive rate was calculated as the fraction of shared yeast peptide features correctly
670 classified as differentially expressed at a certain significance level. The false positive rate was calculated
671 as the fraction of shared human peptide features incorrectly classified as differentially expressed at a
672 certain significance level. **(M-R)** Median within-group standard deviations of independently quantified
673 peptide features. Measurements are the same as shown in (A-F).

674 **Figure 3 - Modeling ion interference at the MS2 level. (A)** Relationship between observed
675 interference levels (OIL) and MS1 ion impurities (1-PPF) of quantified yeast peptides. Each data point
676 corresponds to a single yeast PSM. **(B)** Model predictions (i.e. fitted values) of interference-caused
677 reporter ion signal for quantified yeast peptides. Each data point corresponds to a single yeast PSM. The
678 dashed line represents the identity function (y = x). The underlying model formulas used for calculation
679 of fitted values are depicted as text on top of each plot (see methods section for details). A pseudo-count
680 of 1 was added to each value to ensure finite values after log-transformation. **(C)** Dependence of total
681 reporter ion signal on MS1 precursor intensity. Each data point corresponds to a single PSM. Different
682 colors code for distinct empirical peptide classes as determined by a classification algorithm. Since axes
683 are log2-transformed, differences in slopes appear as differences in intercepts. **(D)** Classification of
684 peptides into distinct classes using a decision tree algorithm. Each internal node reflects a binary split
685 that explains the most variance upon separation of peptides into two distinct subsets based on a single
686 physicochemical property (e.g. contains histidine vs contains no histidine). These properties comprise

687 the precursor charge state (z); the number of TMT-labels (Labels); the inclusion amino acids histidine
688 (H), arginine (R) and lysine (K); and the existence of additional charges not explained by the typically
689 charged functional groups in the peptide's chemical structure (Extra). Terminal nodes reflect distinct
690 empirical peptide classes. The algorithm ensures a sufficient number of observations (>100) in each
691 terminal node by preventing splits that would reduce the number below a certain threshold. **(E)** Model
692 predictions (i.e. fitted values) of total reporter ion signal for quantified peptides. Each data point
693 corresponds to a single PSM. The dashed line represents the identity function (y = x). The underlying
694 model formulas used for calculating fitted values are depicted as text on top of each plot (see methods
695 section for details). An increasing number of predictor variables were sequentially included into the
696 model from left to right.

697 **Figure 4 – Estimation and correction of ion interference. (A)** Relationship between observed
698 interference levels (OIL) and MS1 ion impurities (1-PPF) (top row), or estimated interference level
699 (EIL) (bottom row). The four columns reflect four different levels of yeast feature aggregation (PSMs,
700 peptides, proteins, and proteins with at least 2 unique peptides), with increased aggregation from left to
701 right. Aggregation of metrics EIL and PPF was performed by weighting individual values with the
702 average reporter ion intensity of associated PSMs. Each data point corresponds to a single yeast feature
703 at the respective level of aggregation. **(B-D)** Volcano plots illustrating protein level DE testing results
704 and fold change accuracy and precision for three distinct interference correction strategies, here shown
705 as distinct rows**.** From top to bottom: Uncorrected intensities; PPF-corrected intensities (reporter ion
706 intensities were corrected by subtraction of estimated interference signal calculated on the basis of
707 precursor purity fractions PPF); and EIL-corrected intensities (reporter ion intensities were corrected by
708 subtraction of estimated interference signal calculated on the basis of model-based estimated
709 interference levels EIL). Expected theoretical fold changes between yeast proteins constitute 1.33 (B),
710 1.5 (C) and 2.0 (D), as reflected by vertical red lines in the individual plots. Each data point corresponds
711 to a single protein, filtered for at least 2 unique and unambiguous peptides. Statistical testing was
712 performed via the Limma-trend testing procedure**.** Aggregation to protein intensities was performed by
713 summation of normalized PSM reporter ion intensities. **(E-G)** ROC-analysis of the pairwise group
714 comparisons in (B-D). The true positive rate was calculated as the fraction of yeast proteins correctly
715 classified as differentially expressed at a certain significance level. The false positive rate was calculated
716 as the fraction of human proteins incorrectly classified as differentially expressed at a certain
717 significance level.

718 # SUPPLEMENTAL FIGURE LEGENDS

719 **Supplemental Figure 1 – Supplement to Figure 2. (A)** Observed mean-variance trend of log2-
720 transformed reporter ion intensities from MS2-quantified yeast and human peptides. The solid blue
721 represents a loess-fit with a span parameter of 0.05. Each data point corresponds to a single PSM. **(B-**
722 **G)** Relative numbers of quantified peptides in comparative measurements with changes in single
723 measurement parameters. The highest number in each comparison was set to 1. Measurements are the
724 same as shown in Fig 2.

725 **Supplemental Figure 2 – Illustrations for calculated variables for modeling. (A-B)** Scatterplots
726 comparing precursor purity fraction (PPF) to other related MS1 ion purity metrics. Each data point
727 corresponds to a unique PSM. r denotes the calculated Pearson correlation coefficient. Measurements
728 are the same as shown in Figure 2**. (**A) Comparison with PIF metric, as implemented in MaxQuant. (B)
729 Comparison with Purity metric, as implemented in FragPipe. (C) Comparison with Isolation Interference
730 metric, as implemented in the Proteome Discoverer software MSAmanda**. (D)** Peptide density estimates
731 calculated for all PSMs in a single measurement. Higher density values represent higher numbers of
732 unique peptide features quantified in the vicinity of respective m/z and retention time coordinates.

733 **Supplemental Figure 3 – Evaluation of interference estimation in other data sets. (A-C)**
734 Relationship between observed interference levels (OIL) and MS1 ion impurities (1-PPF) (top row), or
735 estimated interference level (EIL) (bottom row). Each data point corresponds to a single PSM for which
736 the interference-caused reporter ion signal could be measured in the respective experiment. **(**A) Data
737 from published triplicate measurements (0.5 h gradients) of the TKO9 standard. Observed interference
738 levels (OIL) were inferred for proteins of the three KO genes met6, pfk2 and ura2. **(B)** Data from
739 published triplicate measurements (0.5 h gradient) of the TKO11 standard. Observed interference levels
740 (OIL) were inferred for proteins of the three KO genes met6, his4 and ura2. **(C)** Data from measurements
741 (2 h gradient) of the yeast-human mixture experiment using FAIMS-MS2 quantification. Measurements
742 are the same as shown in Figure 2C (FAIMS-MS2).

743

744 # SUPPLEMENTAL TABLE LEGENDS

745 **Supplemental Table 1 – Supplement to Figure 2.** Summary of sample and instrument-specific
746 statistics for comparative measurements of different quantification strategies.

747

ACKNOWLEDGEMENTS

We thank Teresa Preglej and Wilfried Ellmeier for providing the Jurkat cells, Sabrina Jenull and Karl Kuchler for providing yeast cells used in setup experiments, and Gerhard Dürnberger for the support with the Proteome Discoverer software platform. We are grateful to David Hollenstein and Johannes Griss for their critical feedback on the manuscript, and Andreas Bögehold and Thermo Fisher Global Support for the correspondence on orbitrap measurement details. We thank the Vienna Biocenter Core Facilities (VBCF) for providing the LC-MS instrument pool. MM, NH, WR and MH were supported by the Austrian Science Fund (FWF) Special Research Program F70.

AUTHORS' CONTRIBUTION

MM and MH conceptualized the study and designed experiments. MM, WR and NH performed experiments. All R code and statistical analysis was written and performed by MM. FS and KM provided computational tools and resources. MM, WR, and MH wrote the paper. All authors edited the text. All authors read and approved the final manuscript.

763 # REFERENCES

764 1. Carr, S. A. *et al.* Reproducible workflow for multiplexed deep-scale proteome and
765 phosphoproteome analysis of tumor tissues by liquid chromatography–mass spectrometry. *Nat.*
766 *Protoc.* **13**, 1632–1661 (2018).

767 2. Li, J. *et al.* TMTpro-18plex: The Expanded and Complete Set of TMTpro Reagents for Sample
768 Multiplexing. *J. Proteome Res.* **20**, 2964–2972 (2021).

769 3. Thompson, A. *et al.* Tandem mass tags: A novel quantification strategy for comparative analysis
770 of complex protein mixtures by MS/MS. *Anal. Chem.* **75**, 1895–1904 (2003).

771 4. Paulo, J. A. & Schweppe, D. K. Advances in quantitative high-throughput phosphoproteomics
772 with sample multiplexing. *Proteomics* **21**, e2000140 (2021).

773 5. Ow, S. Y. *et al.* iTRAQ underestimation in simple and complex mixtures: 'the good, the bad and
774 the ugly'. *J. Proteome Res.* **8**, 5347–55 (2009).

775 6. Karp, N. A. *et al.* Addressing accuracy and precision issues in iTRAQ quantitation. *Mol. Cell.*
776 *Proteomics* **9**, 1885–97 (2010).

777 7. Niu, M. *et al.* Extensive Peptide Fractionation and y1 Ion-Based Interference Detection Method
778 for Enabling Accurate Quantification by Isobaric Labeling and Mass Spectrometry. *Anal. Chem.*
779 **89**, 2956–2963 (2017).

780 8. Roumeliotis, T. I., Weisser, H. & Choudhary, J. S. Evaluation of a Dual Isolation Width
781 Acquisition Method for Isobaric Labeling Ratio Decompression. *J. Proteome Res.* **18**, 1433–
782 1440 (2019).

783 9. Hogrebe, A. *et al.* Benchmarking common quantification strategies for large-scale
784 phosphoproteomics. *Nat. Commun.* **9**, 1045 (2018).

785 10. Schweppe, D. K. *et al.* Characterization and optimization of multiplexed quantitative analyses
786 using high-field asymmetric-waveform ion mobility mass spectrometry. *Anal. Chem.* **91**, 4010–
787 4016 (2019).

788 11. Ting, L., Rad, R., Gygi, S. P. & Haas, W. MS3 eliminates ratio distortion in isobaric multiplexed
789 quantitative proteomics. *Nat. Methods* **8**, 937–940 (2011).

790 12. O'Connell, J. D., Paulo, J. A., O'Brien, J. J. & Gygi, S. P. Proteome-Wide Evaluation of Two
791 Common Protein Quantification Methods. *J. Proteome Res.* **17**, 1934–1942 (2018).

792 13. Wühr, M. *et al.* Accurate multiplexed proteomics at the MS2 level using the complement reporter
793 ion cluster. *Anal. Chem.* **84**, 9214–21 (2012).

794 14. Johnson, A., Stadlmeier, M. & Wühr, M. TMTpro Complementary Ion Quantification Increases

795 Plexing and Sensitivity for Accurate Multiplexed Proteomics at the MS2 Level. *J. Proteome Res.*
796 **20**, 3043–3052 (2021).

797 15. Pappireddi, N., Martin, L. & Wühr, M. A Review on Quantitative Multiplexed Proteomics.
798 *Chembiochem* **20**, 1210–1224 (2019).

799 16. Savitski, M. M. *et al.* Measuring and managing ratio compression for accurate iTRAQ/TMT
800 quantification. *J. Proteome Res.* **12**, 3586–98 (2013).

801 17. Ahrné, E. *et al.* Evaluation and Improvement of Quantification Accuracy in Isobaric Mass Tag-
802 Based Protein Quantification Experiments. *J. Proteome Res.* **15**, 2537–47 (2016).

803 18. O'Brien, J. J. *et al.* Compositional Proteomics: Effects of Spatial Constraints on Protein
804 Quantification Utilizing Isobaric Tags. *J. Proteome Res.* **17**, 590–599 (2018).

805 19. Savitski, M. M. *et al.* Delayed fragmentation and optimized isolation width settings for
806 improvement of protein identification and accuracy of isobaric mass tag quantification on
807 orbitrap-type mass spectrometers. *Anal. Chem.* **83**, 8959–8967 (2011).

808 20. Phipson, B., Lee, S., Majewski, I. J., Alexander, W. S. & Smyth, G. K. Robust hyperparameter
809 estimation protects against hypervariable genes and improves power to detect differential
810 expression. *Ann. Appl. Stat.* **10**, 946–963 (2016).

811 21. Michalski, A., Cox, J. & Mann, M. More than 100,000 detectable peptide species elute in single
812 shotgun proteomics runs but the majority is inaccessible to data-dependent LC-MS/MS. *J.*
813 *Proteome Res.* **10**, 1785–1793 (2011).

814 22. Tyanova, S., Temu, T. & Cox, J. The MaxQuant computational platform for mass spectrometry-
815 based shotgun proteomics. *Nat. Protoc.* **11**, 2301–2319 (2016).

816 23. Paulo, J. A., O'Connell, J. D. & Gygi, S. P. A Triple Knockout (TKO) Proteomics Standard for
817 Diagnosing Ion Interference in Isobaric Labeling Experiments. *J. Am. Soc. Mass Spectrom.* **27**,
818 1620–5 (2016).

819 24. Gygi, J. P. *et al.* A Triple Knockout Isobaric-Labeling Quality Control Platform with an
820 Integrated Online Database Search. *J. Am. Soc. Mass Spectrom.* **31**, 1344–1349 (2020).

821 25. Scheltema, R. A. *et al.* The Q exactive HF, a benchtop mass spectrometer with a pre-filter, high-
822 performance quadrupole and an ultra-high-field orbitrap analyzer. *Mol. Cell. Proteomics* **13**,
823 3698–3708 (2014).

824 26. Kaufmann, A. & Teale, P. Capabilities and Limitations of High-Resolution Mass Spectrometry
825 (HRMS): Time-of-flight and Orbitrap™. in *Chemical Analysis of Non-antimicrobial Veterinary*
826 *Drug Residues in Food,First Edition* (ed. Jack F. Kay, J. D. M. and J. W.) 93–139 (John Wiley
827 & Sons, Inc., 2016). doi:10.1002/9781118696781.ch3.

27. Ogata, K. & Ishihama, Y. Extending the Separation Space with Trapped Ion Mobility Spectrometry Improves the Accuracy of Isobaric Tag-Based Quantitation in Proteomic LC/MS/MS. *Anal. Chem.* **92**, 8037–8040 (2020).

28. Fischer, M. & Renard, B. Y. IPQF: A new peptide-to-protein summarization method using peptide spectra characteristics to improve protein quantification. *Bioinformatics* **32**, 1040–1047 (2016).

29. Iwasaki, M., Tabata, T., Kawahara, Y., Ishihama, Y. & Nakagawa, M. Removal of Interference MS/MS Spectra for Accurate Quantification in Isobaric Tag-Based Proteomics. *J. Proteome Res.* **18**, 2535–2544 (2019).

30. Teo, G. C., Polasky, D. A., Yu, F. & Nesvizhskii, A. I. Fast Deisotoping Algorithm and Its Implementation in the MSFragger Search Engine. *J. Proteome Res.* **20**, 498–505 (2021).

31. Dorfer, V. *et al.* MS Amanda, a universal identification algorithm optimized for high accuracy tandem mass spectra. *J. Proteome Res.* **13**, 3679–3684 (2014).

32. Gatto, L. & Lilley, K. S. Msnbase-an R/Bioconductor package for isobaric tagged mass spectrometry data visualization, processing and quantitation. *Bioinformatics* **28**, 288–289 (2012).

33. Venables, W. N. & Ripley, B. D. *Modern Applied Statistics with S.* (Springer-Verlag, 2004). doi:10.1007/b97626.

34. Ritchie, M. E. *et al.* limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* **43**, e47 (2015).

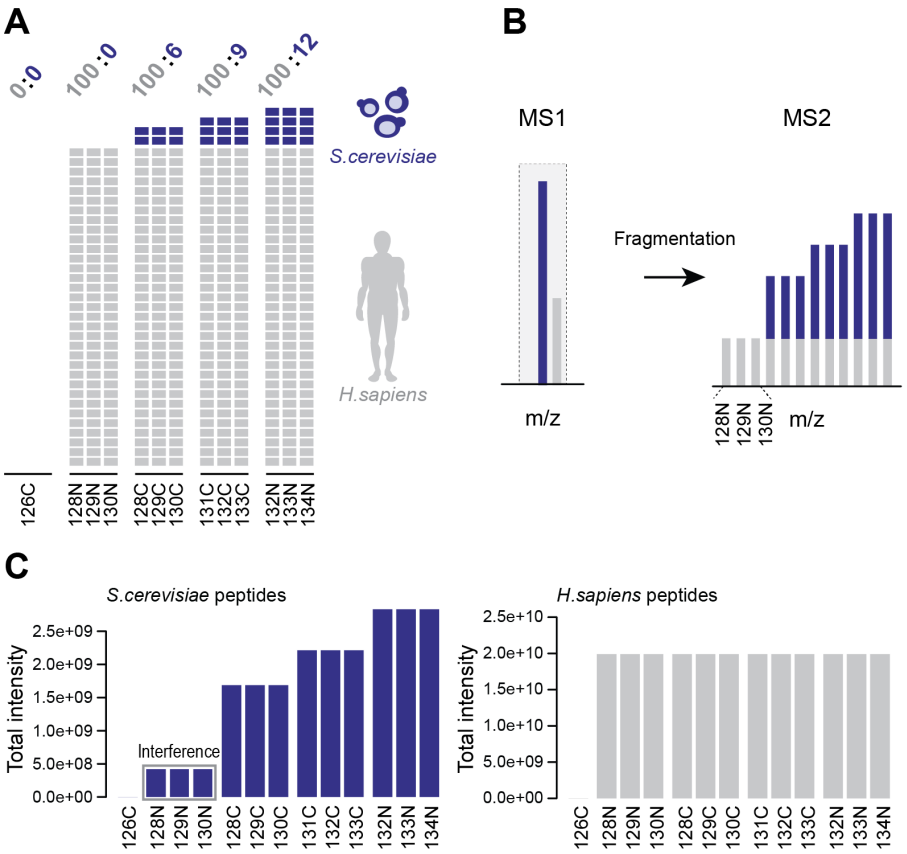

A
0:0
100:0
100:6
100:9
100:12
S.cerevisiae
H.sapiens
126C
128N 129N 130N
128C 129C 130C
131C 132C 133C
132N 133N 134N
B
MS1
MS2
Fragmentation
m/z
128N 129N 130N
m/z
C
S.cerevisiae peptides
H.sapiens peptides
Total intensity
Interference
2.5e+09
2.0e+09
1.5e+09
1.0e+09
5.0e+08
0.0e+00
2.5e+10
2.0e+10
1.5e+10
1.0e+10
5.0e+09
0.0e+00
126C 128N 129N 130N 128C 129C 130C 131C 132C 133C 132N 133N 134N

Figure 1

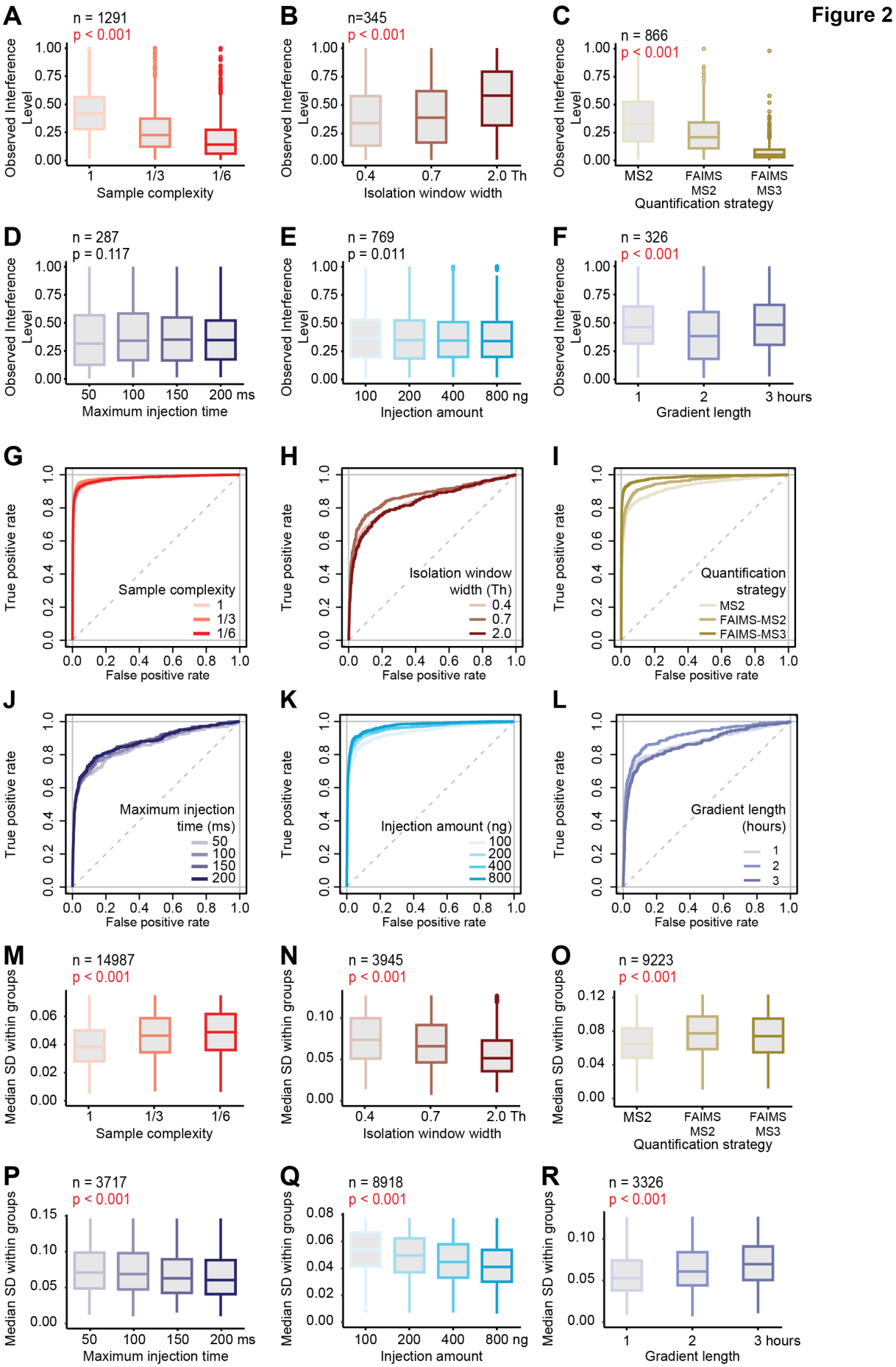
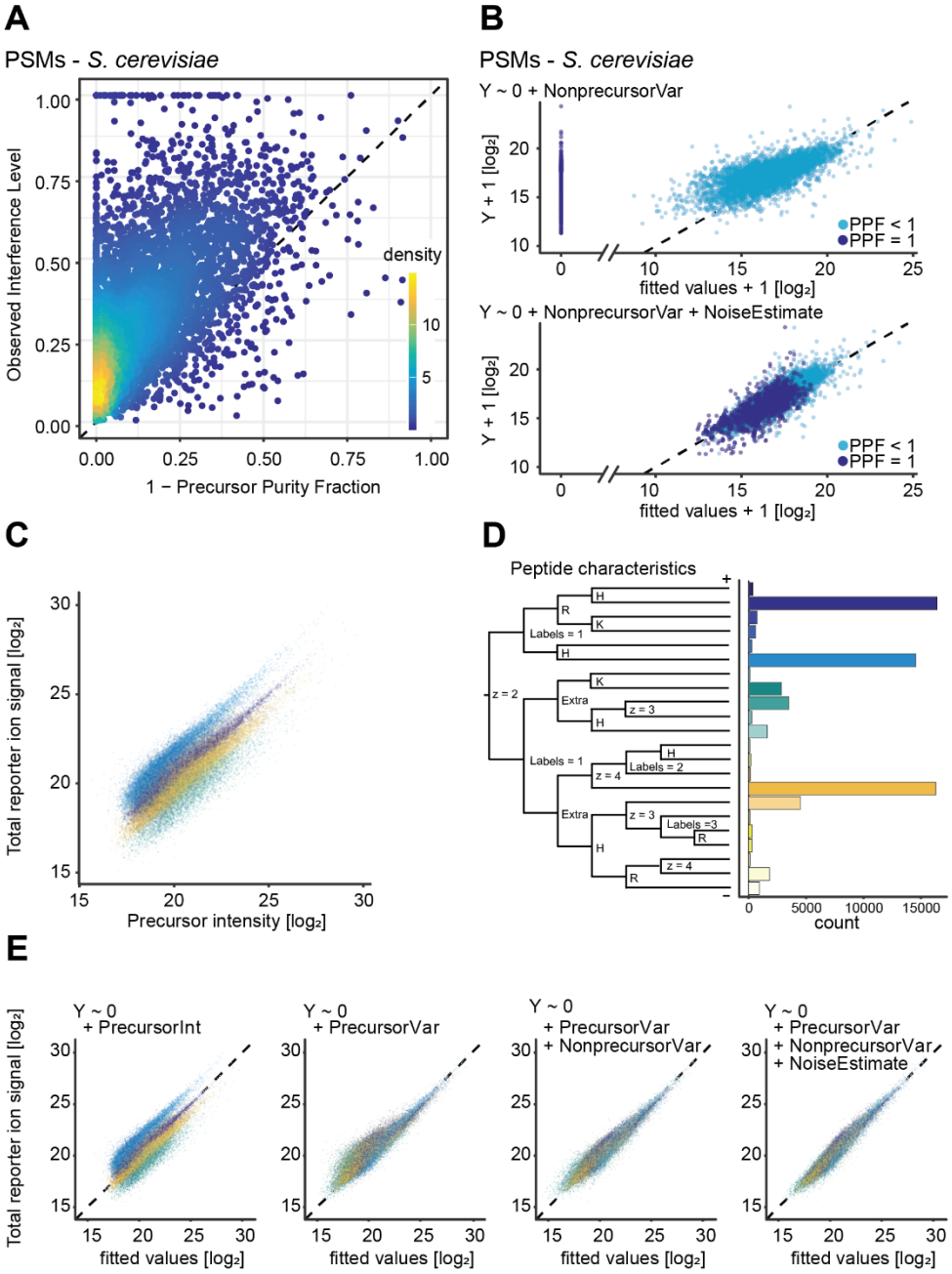


Figure 2
A
n = 1291
p < 0.001
Observed Interference Level
Sample complexity
1
1/3
1/6
B
n=345
p < 0.001
Isolation window width
0.4
0.7
2.0 Th
C
n = 866
p < 0.001
Quantification strategy
MS2
FAIMS MS2
FAIMS MS3
D
n = 287
p = 0.117
Maximum injection time
50
100
150
200 ms
E
n = 769
p = 0.011
Injection amount
100
200
400
800 ng
F
n = 326
p < 0.001
Gradient length
1
2
3 hours
G
Sample complexity
1
1/3
1/6
H
Isolation window width (Th)
0.4
0.7
2.0
I
Quantification strategy
MS2
FAIMS-MS2
FAIMS-MS3
J
Maximum injection time (ms)
50
100
150
200
K
Injection amount (ng)
100
200
400
800
L
Gradient length (hours)
1
2
3
True positive rate
False positive rate
M
n = 14987
p < 0.001
N
n = 3945
p < 0.001
O
n = 9223
p < 0.001
P
n = 3717
p < 0.001
Q
n = 8918
p < 0.001
R
n = 3326
p < 0.001
Median SD within groups

Figure 3



853

854

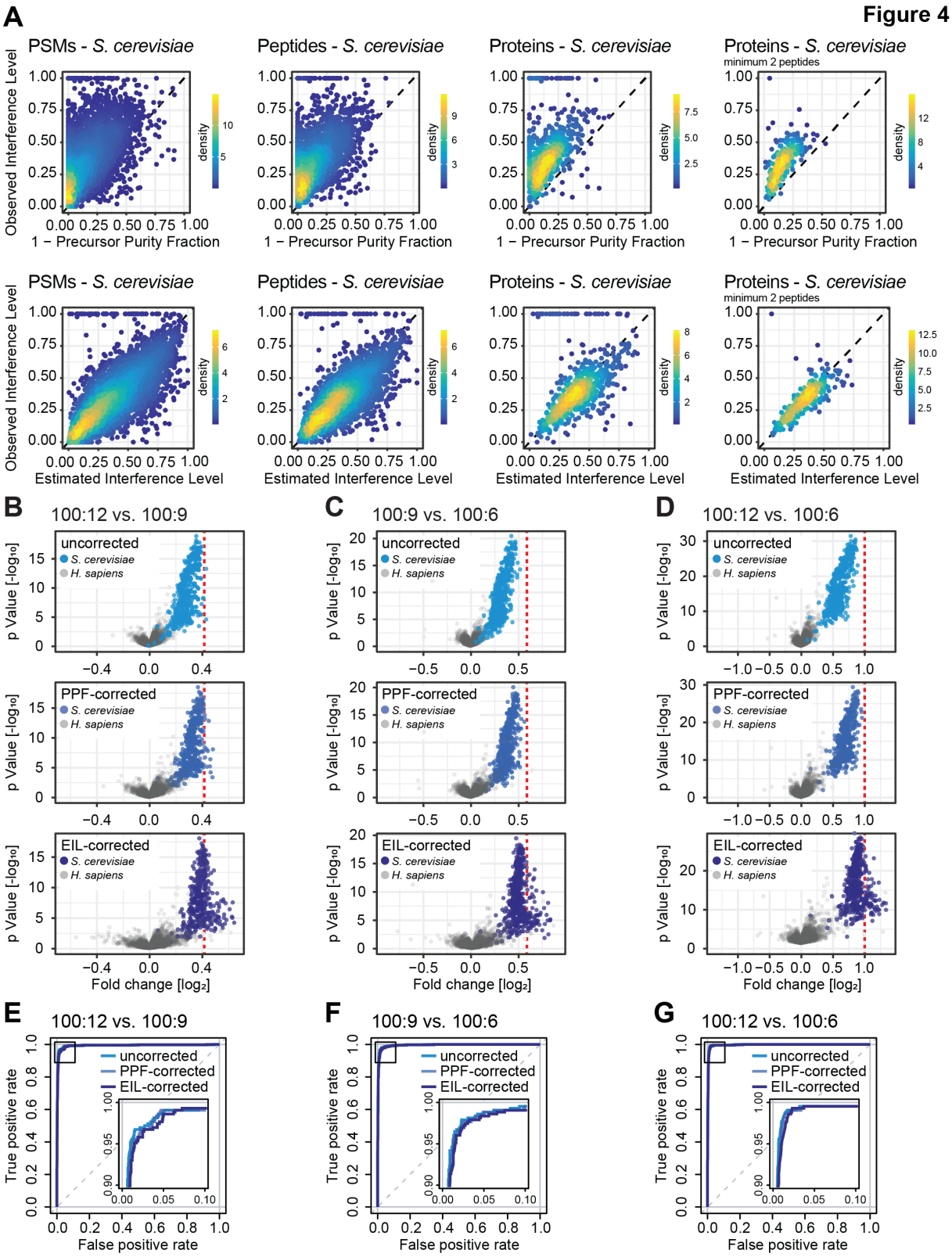
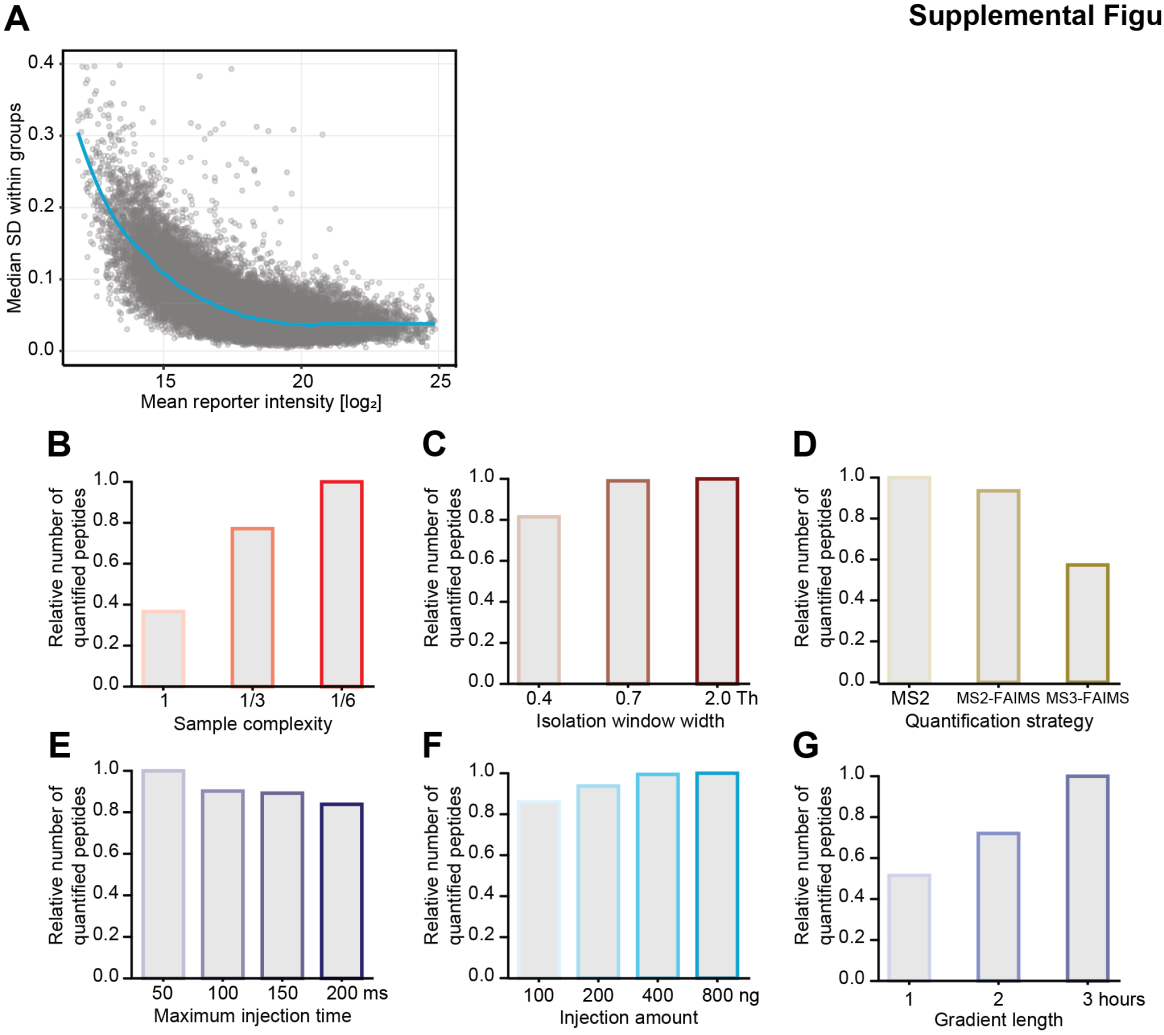


Figure 4
A
PSMs - S. cerevisiae
Peptides - S. cerevisiae
Proteins - S. cerevisiae
Proteins - S. cerevisiae
minimum 2 peptides
Observed Interference Level
1 − Precursor Purity Fraction
Estimated Interference Level
density
B
C
D
100:12 vs. 100:9
100:9 vs. 100:6
100:12 vs. 100:6
uncorrected
PPF-corrected
EIL-corrected
S. cerevisiae
H. sapiens
p Value [-log10]
Fold change [log2]
E
F
G
True positive rate
False positive rate

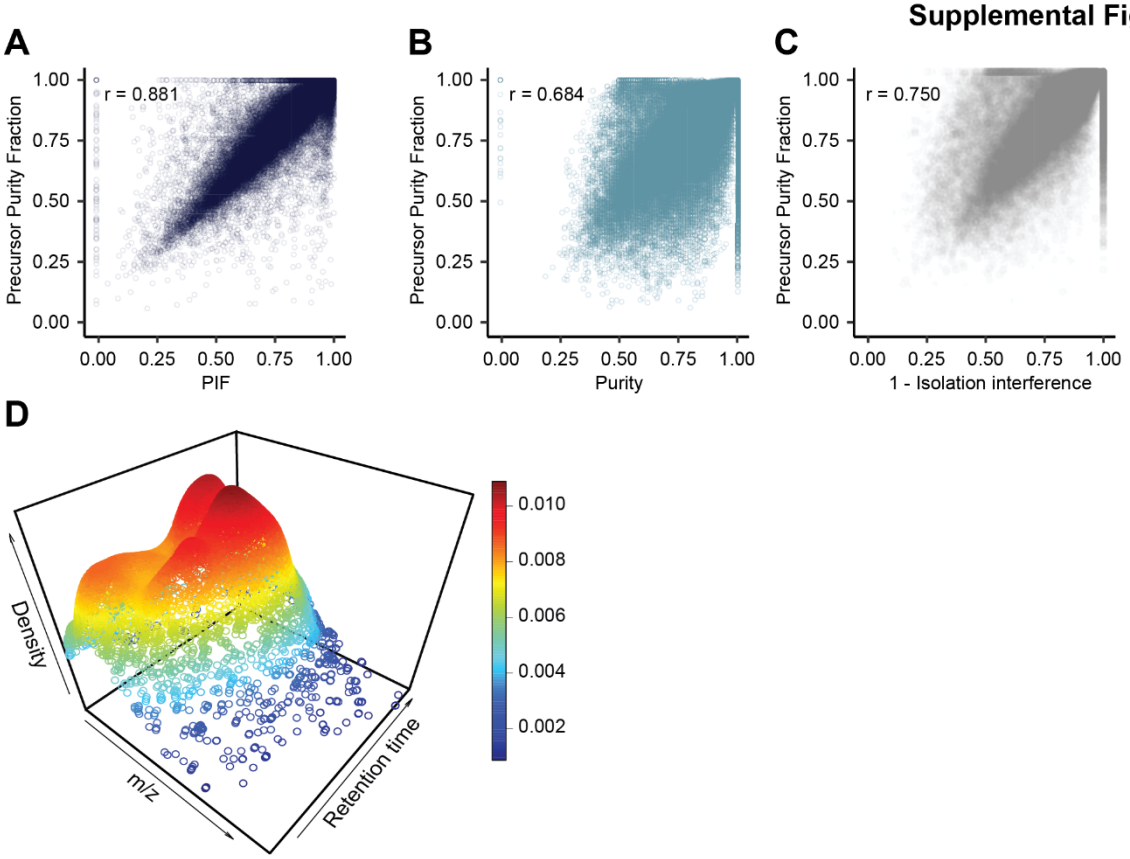
855

856

Supplemental Figure 1



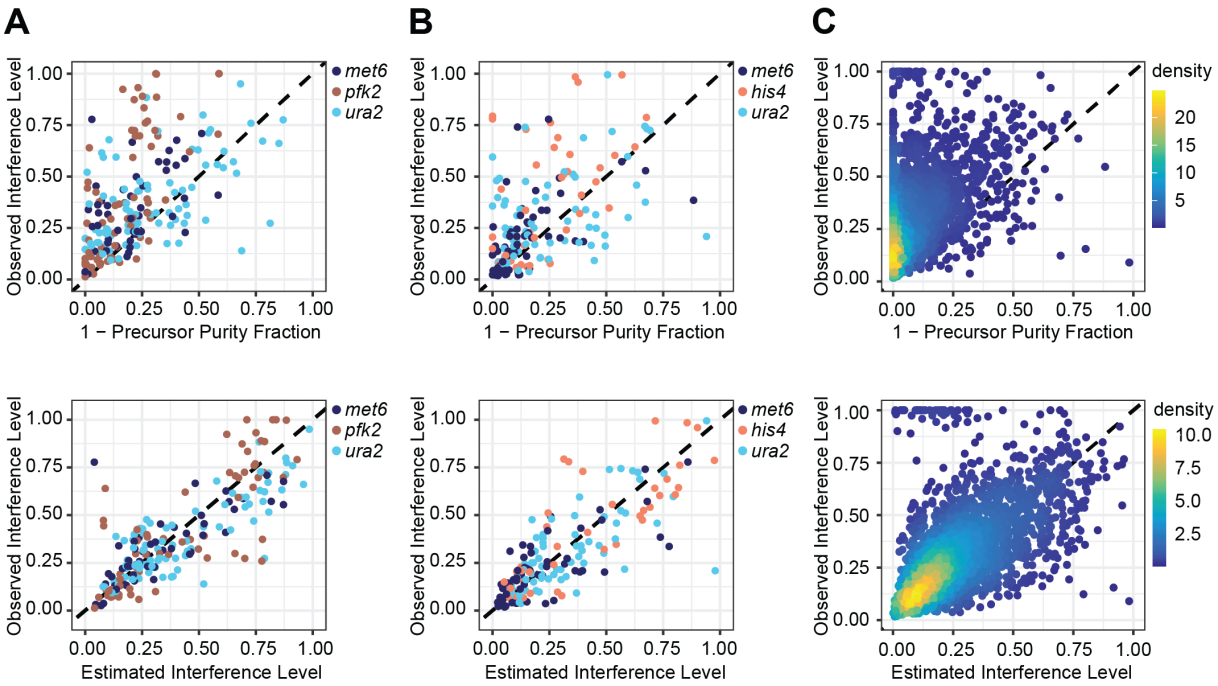
Supplemental Figure 2



859

860

Supplemental Figure 3



861

862

863 **Supplemental Table 1**

	General measurement settings									
	CVs	**MS instrument**	**Number of samples**	**Gradient length**	**Cycle time**	**MS1 res**	**MS1 AGC target**	**MS1 maximum IT**	**Isolation window**	**Intensity threshold**
MS2	-	Eclipse	2	2h	3s	120k	400000	50ms	0.7Th	25000
FAIMS-MS2	-40, -55, -70	Eclipse	2	2h	1.2s	60k	400000	50ms	0.7Th	25000
FAIMS-MS3 RTS	-40, -55, -70	Eclipse	2	2h	1.2s	120k	400000	50ms	0.7Th	10000

	MS2 settings					MS3 settings					
	MS2 detector type	**MS2 AGC target**	**MS2 collision energy**	**MS2 Orbitrap res**	**MS2 maximum IT**	**MS3 detector type**	**MS3 AGC target**	**MS3 collision energy**	**MS3 Orbitrap res**	**MS3 maximum IT**	**MS3 Number of notches**
MS2	Orbitrap	150000	HCD 34%	50k	100ms	-	-	-	-	-	-
FAIMS-MS2	Orbitrap	150000	HCD 34%	50k	100ms	-	-	-	-	-	-
FAIMS-MS3 RTS	IonTrap	10000	CID 30%	-	50ms	Orbitrap	150000	HCD 50%	50k	150ms	10

864