Detecting more peptides from bottom-up mass spectrometry data via peptide-level target-decoy competition

Andy Lin[1], Temana Short[2], William Stafford Noble[3,4], and Uri Keich[*2]

[1]Chemical and Biological Signatures, Pacific Northwest National Laboratory
[2]Department of Statistics, University of Sydney
[3]Department of Genome Sciences, University of Washington
[4]Paul G. Allen School of Computer Science and Engineering, University of Washington

Abstract

A critical statistical task in the analysis of shotgun proteomics data involves controlling the false discovery rate (FDR) among the reported set of discoveries. This task is most commonly solved at the peptide-spectrum match (PSM) level by using target-decoy competition (TDC), in which a set of observed spectra are searched against a database containing a mixture of real (target) and decoy peptides. The PSM-level procedure can be adapted to the peptide level by selecting the top-scoring PSM per peptide prior to FDR estimation. Here, we investigate both PSM-level and peptide-level FDR control methods and come to two conclusions. First, although the TDC procedure is provably correct under certain assumptions, we observe that one of these assumptions—that incorrect PSMs are independent of one another—is frequently violated. Hence, we empirically demonstrate that TDC-based PSM-level FDR estimates can be liberally biased. We propose that researchers avoid PSM-level results and instead focus on peptide-level analysis. Second, we investigate three ways to carry out peptide level TDC and show that the most common method ("PSM-only") offers the lowest statistical power in practice. The most powerful method, peptide-level FDR with PSM competition ("PSM-and-peptide"), carries out competition first at the PSM level and then again at the peptide level. In our experiments, this approach yields an average increase of 17% more discovered peptides at a 1% FDR threshold relative to the PSM-only method.

Key Words: database search, FDR control, false discovery rate

1 Introduction

The goal of most proteomics tandem mass spectrometry experiments is to detect and quantify the proteins present in a complex biological sample. Discoveries from such an experiment are most useful when they are associated with a statistical confidence estimate. Such estimates are typically provided by thresholding with respect to a specified *false discovery rate* (FDR), defined as the expected rate of false discoveries among the set of all accepted discoveries.

In practice, this type of FDR control can be carried out at one of three different levels, corresponding to three types of discoveries. First is the peptide-spectrum match (PSM) level, where each discovery is an observed spectrum linked to the peptide that is inferred to be responsible for

[*]Corresponding author

generating the spectrum. Second is the peptide level, in which multiple PSMs for the same peptide sequence are considered jointly. Third is the protein level, where evidence is accrued across all peptides associated with a given protein. The decision of whether to control the FDR at the level of PSMs, peptides, or proteins depends upon the question at hand, and it is not unusual for a single study to employ more than one type of FDR control for different purposes.

In mass spectrometry proteomics, the most widely used methods for FDR control are based on a straightforward procedure known as *target-decoy competition* (TDC) [6]. The procedure estimates and consequently controls FDR at the PSM level by comparing scores from a search of each observed spectrum against a database of real (*target*) peptides with scores for the same spectrum against a database of reversed or shuffled *decoy* peptides. This method is called target-decoy *competition* because each spectrum only receives a single score: the target and decoy peptide matches to the spectrum compete against one another, and the highest scoring match is assigned to it.

The TDC procedure relies on the underlying assumption that an incorrect match (i.e., a false discovery) is equally like to have come from a target or a decoy peptide. Hence, the number of false discoveries above some score threshold can be estimated by the number of decoys that win the competition and whose scores are above the same threshold. Thus, dividing this estimated number of false discoveries by the number of target peptide wins that score above the same threshold yields an estimate of the FDR. He et al. showed that we can control the FDR at level α by choosing the smallest score threshold for which the estimated FDR is still $\leq \alpha$, provided we add 1 to the number of decoy wins above the threshold before dividing by the number of target wins [11].

Due in part to its simplicity to understand and implement, TDC is by far the most widely used method for PSM-level FDR control in proteomics mass spectrometry. Recently, the same competition-based approach to control the FDR gained significant popularity in the statistics and machine learning communities after Barber and Candès introduced their knockoff+ procedure [1].

Motivated by the PSM-level TDC, decoy-based approaches to controlling the FDR at the peptide and protein levels have also been described. In both cases the idea is to aggregate the scores of all PSMs involving a peptide (taking the maximum of the scores) or a protein (taking the sum of the peptide scores). We thus assign a score to each target and decoy peptide (or protein, depending on the level of our analysis). With these scores in hand we can continue analogously to the PSM-level TDC, although as we will see below there is more than one way to implement how the scores are aggregated and how the competition is implemented.

Our goal in this paper is to convey two distinct messages related to decoy-based FDR control in mass spectrometry proteomics. The first is that controlling the FDR at the level of PSMs is problematic, due to unavoidable potential dependencies between the incorrect PSMs. As we mentioned above, TDC provably controls the FDR, but only under certain assumptions. Unfortunately, as others have pointed out previously, in practice these assumptions might not be met [11]. Here we provide further empirical evidence showing that PSM-level FDR estimates can be liberally biased, meaning that if you try to control the FDR at, say, 1%, then the TDC procedure is likely to control the FDR at a less stringent threshold. In our experiments, the magnitude of the effect becomes larger as the FDR threshold increases.

Our second message is that the most common way to control FDR at the peptide level is suboptimal, in the sense that it yields fewer detected peptides than a relatively simple variant of that procedure. Though it is rare in the literature to precisely specify how peptide-level FDR is performed, the few cases where we found a precise description appear to carry out PSM-level competition only [11, 14, 18]. Specifically, the explicit pairing between each target peptide and its shuffled decoy is ignored: the peptides are ordered according to their scores (taking the maximum scoring PSM involving each peptide), and the FDR is estimated and controlled using the number of decoy peptides above the threshold as noted above for the PSM-level TDC.

We investigated three variant techniques for carrying out peptide-level FDR control using TDC. These variants were motivated in part by the "picked protein" approach for protein-level FDR control [28]. The idea, whose benefits were demonstrated in [25] and [27], is to perform a "head-to-head" competition between each target protein and its paired decoy, keeping only the higher scoring of the two. Thus, we compared the above PSM-only competition approach to peptide-level analysis with two additional protocols, both of which involve a direct peptide-level competition between each target peptide and its paired decoy (so only the higher scoring of the two is kept). The first protocol, peptide-only, involves only the peptide-level competition, i.e., the PSMs are generated by separately searching the target and decoy databases. The second protocol, PSM-and-peptide, involves both PSM- and peptide-level competition, so the PSMs are generated by searching a concatenated database as in the commonly used approach. After generating the PSMs and performing the peptide-level competition both procedures proceed as in the PSM-level TDC.

We find that PSM-and-peptide, which uses competition first at the PSM level and then at the peptide level, yields the greatest number of discovered peptides, whereas the commonly used PSM-only competition yields significantly lower statistical power. Notably, switching from PSM-only to PSM-and-peptide is relatively straightforward, as long as it is possible to match each target peptide to its corresponding decoy.

2 Methods

2.1 Methods for FDR control

2.1.1 PSM-level TDC

The goal of TDC carried out at the PSM level is to estimate and control the false discovery rate among a collection of PSMs produced by a database search engine. We are given a set S of n spectra, and we assume that the spectra have already been searched against a target database $\mathcal{T}$ and a decoy database $\mathcal{D}$. For each spectrum we retain the top-scoring PSM, breaking any ties randomly. We refer to the scores of the PSMs that involve target peptides as $t_1, t_2, \ldots, t_{m_t}$ and the scores of decoy PSMs as $d_1, d_2, \ldots, d_{m_d}$, where $m_t + m_d = n$. We can then estimate the FDR among all PSMs that score greater than a specified score threshold τ (assuming that larger scores are better) as

$$\widehat{\mathrm{FDR}}(\tau) = \min\left(1, \frac{|\{d_i \geq \tau; i = 1, \ldots, m_d\}| + 1}{|\{t_i \geq \tau; i = 1, \ldots, m_t\}|}\right) \tag{1}$$

Intuitively, the denominator represents the number of discoveries of interest (the target PSMs), and the numerator is our decoy-based estimate of the number of false positives among those discoveries.

In practice, rather than specifying a priori a score threshold τ, we typically specify the desired FDR threshold α and choose our rejection threshold, $\tau(\alpha)$, as the smallest τ for which the estimated FDR is $\leq \alpha$:

$$\tau(\alpha) = \min\{t_i \,:\, \widehat{\mathrm{FDR}}(t_i) \leq \alpha\}. \tag{2}$$

The estimated FDR in Equation (1) differs from the one offered by Elias and Gygi [6] in three ways. First, their formulation includes a factor of 2 in the numerator and includes both target and decoy PSMs in the denominator. This approach thus controls the FDR among the combined set of target and decoy PSMs. In practice, it typically is of more interest to control the FDR only among the targets, as in Equation (1). Second, the numerator in our formulation includes a +1 that is missing from the Elias and Gygi formulation. This +1 correction is required in order to achieve valid FDR control [11, 17]. In practice, this +1 correction will have a negligible effect except in

the presence of very few discoveries or a very stringent FDR threshold. Finally, our formulation includes an enclosing min operation, which simply ensures that we do not report an estimated FDR > 1.

2.1.2 PSM-only competition for peptide-level FDR control

The most commonly used method for estimating FDR at the peptide level, which we refer to as "PSM-only" is quite straightforward [14]. This procedure starts by carrying out PSM-level target-decoy competition, retaining only the top-scoring PSM per peptide sequence. Thereafter, the FDR among the target peptides is estimated and controlled using the analogs of Equations (1) and (2), where the scores t_i and d_i refer to the peptide rather than PSM scores.

2.1.3 Peptide-only and PSM-and-peptide competition for peptide-level FDR control

The next two procedures for peptide-level FDR control procedure mimic a crucial step of the "picked protein" procedure for protein-level FDR control [28]. Specifically, both the "peptide-only" and "PSM-and-peptide" procedures employ a head-to-head competition between each target peptide and its paired decoy so that only the higher scoring of the two is kept.

The difference between the two methods is that the PSM-and-peptide approach creates PSMs by scanning each spectrum against the concatenated target-decoy database, keeping only the best matching peptide for the scanned spectrum. In contrast, the peptide-only approach does not employ a PSM-level competition; instead, it separately scans each spectrum against the target and the decoy databases, keeping the corresponding two best matches for the scanned spectrum.

Both methods assign each target or decoy peptide a score that is the maximum of all PSMs involving the peptide. Both methods then carry out peptide-level competition for each target-decoy peptide pair. The rest follows exactly as in the PSM-only approach, where the FDR among the target peptides is estimated and controlled using Equations (1) and (2), where the scores refer to the peptide scores.

2.2 Datasets

To evaluate our various FDR control methods, we used runs from six different mass spectrometry datasets. These six datasets consisted of data from the ISB18 mix [16], castor plant [22], *E. coli* [5], human [5], mouse [13], and yeast samples [10]. All nine runs from the ISB18 dataset were downloaded from `http://regis-web.systemsbiology.net/PublicDatasets`. Two runs from each of the five other datasets were downloaded from PRIDE [24] (Table 1). Raw mass spectrometry runs were converted to MS2 format using MSConvert version 3.0 [2]. The protein sequence databases corresponding to the castor plant, *E. coli*, human, mouse, and yeast proteomes were downloaded from Uniprot (`www.uniprot.org`) in October 2021 or January 2022. The sequences corresponding to the ISB18 mix were downloaded from `https://regis-web.systemsbiology.net/PublicDatasets/database`.

2.3 Database searching

A database search was conducted on spectra from the castor plant, *E. coli*, human, mouse, and yeast runs using the Tide search engine [4] as implemented in Crux version 4.1 [21, 23] against a target-decoy database containing the proteome of the sample being analyzed. Each file was searched using four different score functions: XCorr [7], XCorr p-value [12], combined p-value [19], and Tailor [26]. The precursor mass tolerances were estimated using Param-Medic [20] and set

File name	Species	PRIDE ID
Rcom_9_M4_AM_R1_7Mar16_Samwise_15-08-55	castor plant	PXD007933
Rcom_Zanz_2_1_03Jun16_Samwise_16-03-32	castor plant	PXD007933
134_2018_ZBS6_Ecoli_SP3_2	*E. coli*	PXD011189
141_2018_ZBS6_Ecoli_SPEED_3	*E. coli*	PXD011189
228_2018_ZBS6_HeLa_FASP_3	human	PXD011189
235_2018_ZBS6_HeLa_SP3_1	human	PXD011189
HF1_003796	mouse	PXD028550
HF1_003800	mouse	PXD028550
Tre1	yeast	PXD009420
Tre4	yeast	PXD009420

Table 1: **Datasets used in this study.**

to be 85, 40, 35, 40, and 25 ppm for the castor plant, *E. coli*, human, mouse, and yeast runs, respectively. For these searches, all other parameters were set to their default values, except that `--top-match=1` and one missed cleavage was allowed. For the XCorr p-value and combined p-value score functions, `--exact-p-value=T` and `--mz-bin-width=1.0005079`. For the Tailor score function `--use-tailor-calibration=T`. We note that `--concat=F` by default, and therefore target and decoy results are reported separately.

For the database search of the ISB18 data, we used an entrapment setup where all scans from the nine runs were searched against a database containing the ISB18 proteins and the castor plant proteome using the exact p-value score function. The precursor mass tolerances was set to 50 ppm, and we set `--exact-p-value=T` and `--mz-bin-width=1.0005079`. In addition, the spectra were searched against 10 different sets of generated decoys. Peptides that were found in both ISB18 proteins and the castor plant proteome were removed from the database.

The false discovery proportion (FDP) of the FDR controlled set of ISB18 PSMs was calculated by dividing the number of incorrect PSMs (*i.e.* ISB18 spectra that match with a castor plant peptide) by the total number of detections.

3 Results

3.1 PSM-level FDR control with TDC can be liberally biased

As we described above, in order for TDC to be valid when controlling the PSM-level FDR, the incorrect PSMs must be independent from each other. However, in practice we expect spectra that are generated from the same peptide to be highly correlated. Dynamic exclusion helps to reduce the magnitude of this problem but does not remove the problem, especially if, as is often done, multiple runs are combined before applying TDC. The problem with these highly correlated spectra is that their corresponding optimal PSMs tend to involve the same peptide. If this match happens to be an incorrect match to a target peptide, then it can significantly inflate the actual FDP in a way that TDC cannot account for. In such cases, TDC may fail to control the FDR.

Indeed, imagine an extreme case where the dataset consists of 100 spectra that were all generated by the same peptide and this peptide does not appear in the target database. In this case it is likely that all or most of the spectra will match to the same database peptide, and there is a 50% chance that this peptide will be a target. In that scenario, having observed no decoy matches, we would accept these 100 target PSMs at 1% FDR threshold, whereas the reality is that all of them are incorrect.

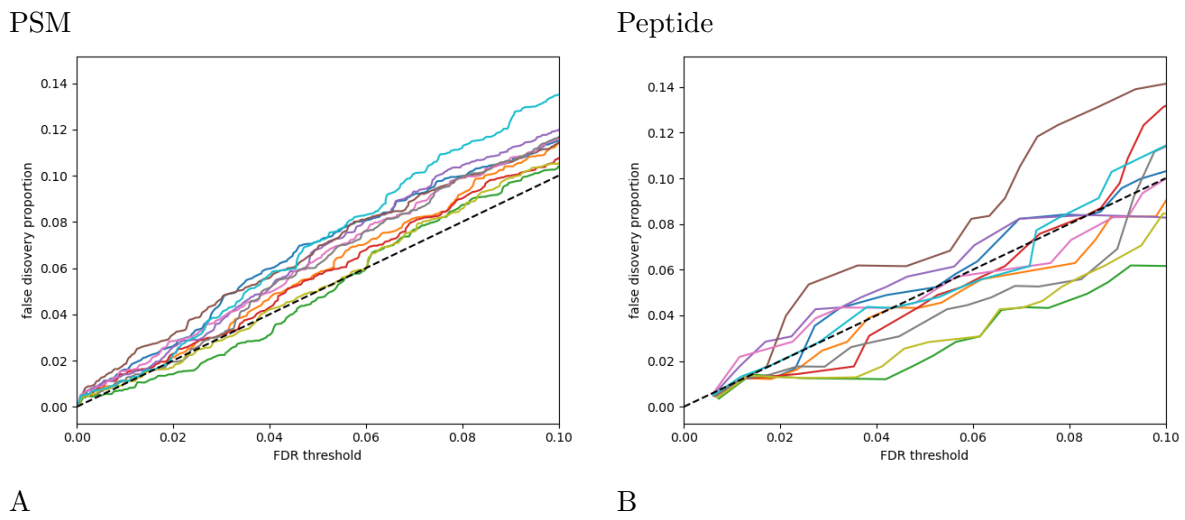

Figure 1: **PSM-level FDR control with TDC can be liberally biased.** The figure plots the false discovery proportion (y-axis) as a function of the FDR threshold (x-axis) for a database search carried out using an entrapment setup for PSM-level (A) and peptide-level (B) analysis. Each color represents a different set of decoys, and the colors are consistent between the two plots. Valid FDR control would correspond to lines whose average falls below the line $y = x$ (dashed line). The higher variance in the peptide-level FDR plot is related to the fewer number of discoveries at each threshold. We observe an average of 3084.5 PSM-level and 156.8 peptide-level discoveries at 1% FDR.

To demonstrate that such a failure can happen in practice, we conducted an entrapment experiment using the ISB18 data. In an entrapment experiment, spectra are searched against a database containing the sequences in the sample in addition to a set of sequences not present in the sample. Importantly, the number of entrapment sequences is set to be much larger than the number of sequences that are present in the sample, and hence PSMs involving entrapment sequences are most likely incorrect. Therefore, we can estimate the false discovery proportion (FDP) among the set of reported target PSMs and compare that FDP to the selected FDR threshold. If an FDR control procedure is valid, then we expect that the FDP among its reported discoveries should not be consistently larger than the threshold.

In our setting, we jointly searched the spectra from nine ISB18 runs, as an example of a dataset with moderately high spectrum multiplicity. We searched the spectra against a database containing the ISB18 proteins plus the castor plant proteome, where the castor plant proteome was considered the entrapment database. After removing peptide sequences in common, the castor plant and ISB18 proteomes contained 571,319 and 449 peptides, respectively. We repeated this process 10 times with 10 different sets of decoys.

Our results suggest that TDC fails to control the PSM-level FDR in this entrapment setup. Specifically, our analysis shows that for the entire plotted range of FDR thresholds (0–10%) the FDP among the reported PSMs was consistently larger than the given threshold across all 10 searches (Figure 1A). The deviation between the FDP and the FDR threshold increased as the threshold increased. For example, at FDR thresholds of 5% and 10% the average FDP was 0.062 and 0.1148, respectively. On the other hand, our results suggest that using TDC to control peptide-level FDR is valid because the peptide-level FDP values center around the $y = x$ line across the entire plotted threshold range (Figure 1B). Note that, because the FDR is defined as the expectation

of the FDP, we expect the FDP to occasionally exceed the threshold, as observed.

3.2 Peptide-level FDR control benefits from peptide-level competition

Next, we focused on comparing the statistical power of our three different methods for estimating peptide-level FDR: PSM-only, peptide-only, and PSM-and-peptide. For this analysis, we searched runs from castor plant, *E. coli*, human, mouse, and yeast against their respective databases, and we counted the number of reported discoveries at various FDR thresholds (0–10%) for each of the three FDR procedures. Each run was searched four times against the same protein database but with a different score function each time (XCorr, XCorr p-value, Tailor, and combined p-value).

Empirically, we found that PSM-and-peptide yielded the best performance across all runs and score functions (Figure 2). Conversely, we found that the PSM-only procedure had the overall worst performance, while the peptide-only procedure always had middling performance. For example, in the *E. coli* run at a 1% FDR, PSM-and-peptide outperformed PSM-only by 1050 (11.33%), 2397 (25.85%), 3040 (41.02%), and 881 (10.38%) peptide detections for the combined p-value, Tailor, XCorr, and XCorr p-value score functions, respectively. Similarly, at the same 1% threshold PSM-and-peptide outperformed peptide-only by 589 (5.84%), 544 (5.35%), 1338 (14.68%), and 364 (4.04%) peptide detections. An analysis of a second run from each of the five datasets showed similar results (Supplementary Figure 1)

We observed that the performance boost of PSM-and-peptide, when compared to peptide-only and PSM-only, was generally greatest when using a non-calibrated score. In this setting, a score is calibrated if a particular value X has the same significance regardless of the peptide that was scored [15]. Specifically, we noted that in nine out of 10 runs, the performance boost when using PSM-and-peptide was greatest when using the non-calibrated XCorr score. The one exception was a castor plant run where the heuristically-calibrated Tailor score had the greatest boost in performance instead of the XCorr score. This observation is not surprising, because the additional peptide-level competition in PSM-and-peptide confers a form of calibration.

4 Discussion

We argued here, as did He et al. before us [11], that multiplicity, where multiple spectra are generated by the same peptide species, can impair TDC's ability to control the PSM-level FDR. To avoid this problem, we therefore suggest that mass spectrometrists should employ FDR control exclusively at the peptide or protein levels.

An alternative way to circumvent the problem with PSM-level FDR control is to cluster the spectra in the hope that each cluster of spectra corresponds to a single peptide species [8]. However, it is not clear whether this approach—with its added complexity of how to define the clusters—offers any advantage over simply switching to peptide-level analysis in the first place.

One question that intrigued us is whether we can use the multiplicity to improve our peptide-level analysis. Specifically, we postulated that if, in addition to its score, we assign to each peptide its multiplicity, i.e., the number of PSMs that were optimally matched to the peptide, then we could better distinguish between peptides that were truly present in the samples and false detections. To test this hypothesis, we used our recently developed Group-walk [9] to divide the peptides into groups according to their multiplicity and compared that approach to a single-group approach that ignores the multiplicity. While we did see a fairly consistent power gain when using the multiplicity in this way, that gain was unfortunately marginal (about 0.5% at 1% FDR), so we concluded that it does not justify the added complexity of the procedure.

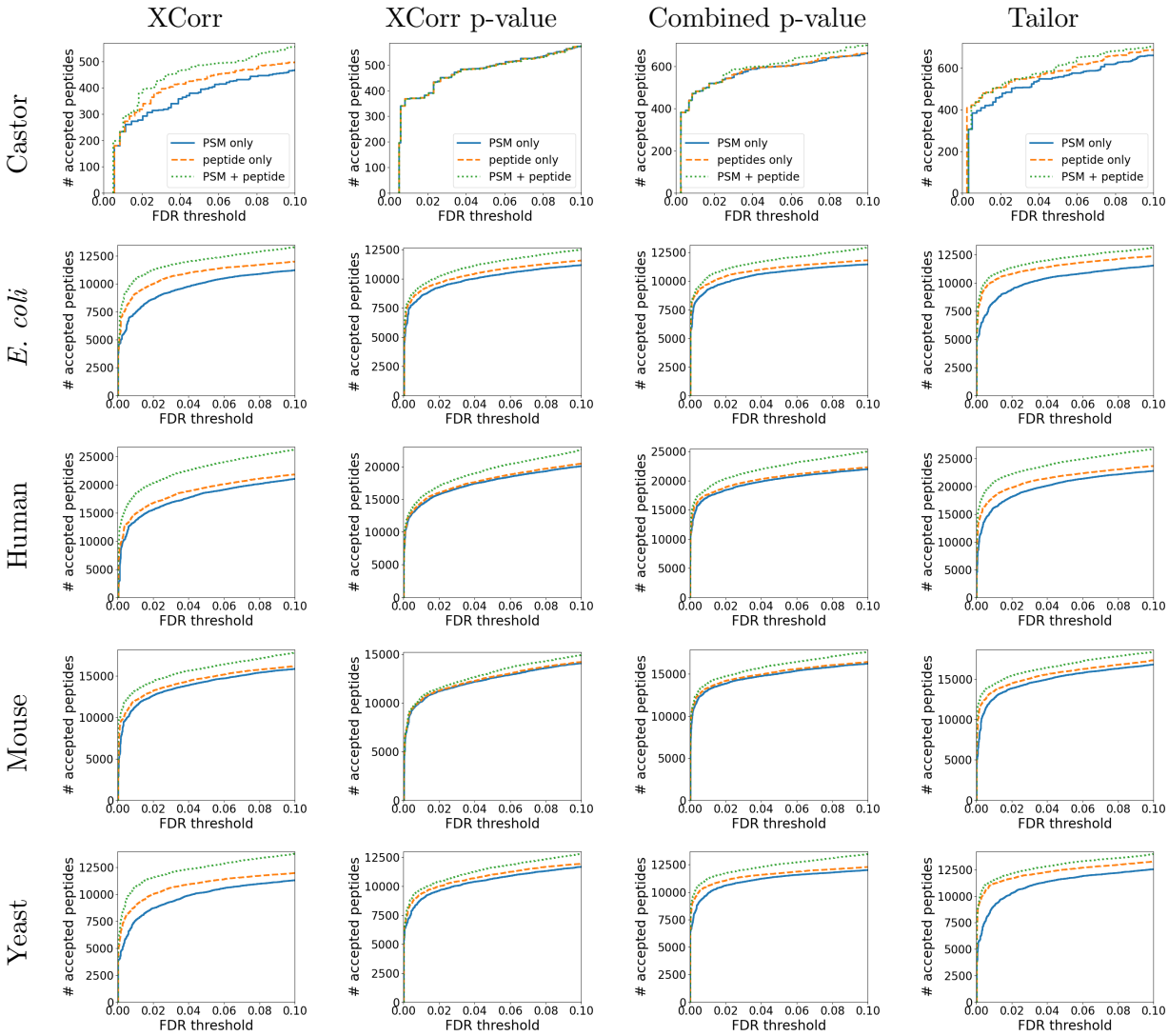

Figure 2: **Peptide-level FDR procedures.** Each plot compares the performance of three different peptide-level FDR estimation and control procedures. Each row of panels represents a run from a different species, and each column represents a different score function (XCorr, XCorr p-value, combined p-value, and Tailor score). The PSM-and-peptide method generally has the best performance for all score functions and runs across the plotted range of FDR thresholds (0–10%). Supplementary Figure 1 shows similar results using a different run from each of the five datasets.

One of our main conclusions is that the commonly used method for estimating peptide-level FDR using the PSM-only approach is conservative. To get get some intuition behind why this is the case, consider the following example: suppose our target database consists only of peptides that are in the sample. In this case there cannot be any false target discoveries, so any false discovery must be a decoy discovery. However, this scenario violates one of the the two critical assumptions of TDC (the first being the independence of the false identifications): that conditioned on the discovery being false, every win is equally likely to be a target or a decoy. In particular, this example implies that the PSM-only procedure will be conservative, which is consistent with our observation that this procedure reports significantly fewer peptides than the peptide-only and PSM-and-peptide methods.

One practical challenge in switching from the PSM-only to the PSM-and-peptide procedure is the requirement that the search engine retain the exact matching from target to decoy peptides. If the decoy database is generated by reversing each target peptide, then matching targets to corresponding decoys is trivial. On the other hand, for a shuffled database, this pairing information must be retained and reported to the user for use during FDR control. The Tide search engine supports this type of analysis by reporting, for each decoy PSM, the corresponding target peptide and by optionally generating a list of targets and corresponding decoys during the database indexing step.

Finally, in this work, we have focused on TDC for two reasons. First, TDC is the most commonly used approach to controlling the FDR in the analysis of mass spectrometry data. Second, it is a fairly flexible method that in principle can work with any score function as long as competing decoy scores can be computed. However, some tools are reporting p-values which can then alternatively be used in combination with canonical p-value based FDR controlling procedures [3].

Acknowledgments Some of research described in this paper was conducted under the Laboratory Directed Research and Development Program at Pacific Northwest National Laboratory, a multiprogram national laboratory operated by Battelle for the U.S. Department of Energy. Andy Lin is grateful for the support of the Linus Pauling Distinguished Postdoctoral Fellowship program. Pacific Northwest National Laboratory is a multiprogram national laboratory operated by Battelle Memorial Institute for the United States Department of Energy under contract DE-AC06-76RLO.

References

[1] R. F. Barber and Emmanuel J. Candès. Controlling the false discovery rate via knockoffs. *The Annals of Statistics*, 43(5):2055–2085, 2015.

[2] M. C. Chambers, B. Maclean, R. Burke, D. Amodei, D. L. Ruderman, S. Neumann, L. Gatto, B. Fischer, B. Pratt, J. Egertson, K. Hoff, D. Kessner, N. Tasman, N. Shulman, B. Frewen, T. A. Baker, M. Y. Brusniak, C. Paulse, D. Creasy, L. Flashner, K. Kani, C. Moulding, S. L. Seymour, L. M. Nuwaysir, B. Lefebvre, F. Kuhlmann, J. Roark, P. Rainer, S. Detlev, T. Hemenway, A. Huhmer, J. Langridge, B. Connolly, T. Chadick, K. Holly, J. Eckels, E. W. Deutsch, R. L. Moritz, J. E. Katz, D. B. Agus, M. J. MacCoss, D. L. Tabb, and P. Mallick. A cross-platform toolkit for mass spectrometry and proteomics. *Nature Biotechnology*, 30(10):918–920, 2012.

[3] Yohann Couté, Christophe Bruley, and Thomas Burger. Beyond target–decoy competition: Stable validation of peptide and protein identifications in mass spectrometry-based discovery proteomics. *Analytical Chemistry*, 92(22):14898–14906, 2020.

[4] B. Diament and W. S. Noble. Faster SEQUEST searching for peptide identification from tandem mass spectra. *Journal of Proteome Research*, 10(9):3871–3879, 2011.

[5] J. Doellinger, A. Schneider, M. Hoeller, and P. Lasch. Sample preparation by easy extraction and digestion (speed) - a universal, rapid, and detergent-free protocol for proteomics based on acid extraction. *Molecular & Cellular Proteomics*, 19(1):209–222, 2020.

[6] J. E. Elias and S. P. Gygi. Target-decoy search strategy for increased confidence in large-scale protein identifications by mass spectrometry. *Nature Methods*, 4(3):207–214, 2007.

[7] J. K. Eng, B. Fischer, J. Grossman, and M. J. MacCoss. A fast SEQUEST cross correlation algorithm. *Journal of Proteome Research*, 7(10):4598–4602, 2008.

[8] A. Frank, N. Bandeira, Z. Shen, S. Tanner, S. P. Briggs, R. D. Smith RD, and P. A. Pevzner. Clustering millions of tandem mass spectra. *Journal of Proteome Research*, 7(1):113–121, 2008.

[9] J. Freestone, T. Short, W. S. Noble, and U. Keich. Group-walk, a rigorous approach to group-wise false discovery rate analysis by target-decoy competition. *bioRxiv*, 2022. 10.1101/2022.01.30.478144.

[10] M. Garcia-Albornoz, S. W. Holman, T. Antonisse, P. Daran-Lapujade, B. Teusink, R. J. Beynon, and S. J. Hubbard. A proteome-integrated, carbon source dependent genetic regulatory network in *saccharomyces* cerevisiae. *Moleculr Omics*, 16:59–72, 2020.

[11] K. He, Y. Fu, W.-F. Zeng, L. Luo, H. Chi, C. Liu, L.-Y. Qing, R.-X. Sun, and S.-M. He. A theoretical foundation of the target-decoy search strategy for false discovery rate control in proteomics. *arXiv*, 2015. https://arxiv.org/abs/1501.00537.

[12] J. J. Howbert and W. S. Noble. Computing exact p-values for a cross-correlation shotgun proteomics score function. *Molecular and Cellular Proteomics*, 13(9):2467–2479, 2014.

[13] A. Imbert, M. Rompais, M. Selloum, F. Castelli, E. Mouton-Barbosa, M. Brandolini-Bunlon, E. Chu-Van, C. Joly, A. Hirschler, P. Roger, T. Burger, S. Leblanc, T. Sorg, S. Ouzia, Y. Vandenbrouck, C. Médigue, C. Junot, M. Ferro, E. Pujos-Guillot, A. G. de Peredo, F. Fenaille, C. Carapito, Y. Herault, and E. A. Thévenot. ProMetIS, deep phenotyping of mouse models by combined proteomics and metabolomics analysis. *Sci Data*, 8(1):311, 12 2021.

[14] K. Jeong, S. Kim, and N. Bandeira. False discovery rates in spectral identification. *BMC Bioinformatics*, 13(Suppl. 16):S2, 2012.

[15] U. Keich and W. S. Noble. On the importance of well calibrated scores for identifying shotgun proteomics spectra. *Journal of Proteome Research*, 14(2):1147–1160, 2015.

[16] J. Klimek, J. S. Eddes, L. Hohmann, J. Jackson, A. Peterson, S. Letarte, P. R. Gafken, J. E. Katz, P. Mallick, H. Lee, A. Schmidt, R. Ossola, J. K. Eng, R. Aebersold, and D. B. Martin. The standard protein mix database: a diverse data set to assist in the production of improved peptide and protein identification software tools. *Journal of Proteome Research*, 7(1):96–1003, 2008.

[17] L. I. Levitsky, M V. Ivanov, A. A. Lobas, and M. V. Gorshkov. Unbiased false discovery rate estimation for shotgun proteomics based on the target-decoy approach. *Journal of Proteome Research*, 16(2):393–397, 2017.

[18] H. Li, J. Park, H. Kim, K.-B. Hwang, and E. Paek. Systematic comparison of false-discover-rate-controlling strategies for proteogenomic search using spike-in experiments. *Journal of Proteome Research*, 16:2231–2239, 2017.

[19] A. Lin, J. J. Howbert, and W. S. Noble. Combining high-resolution and exact calibration to boost statistical power: A well-calibrated score function for high-resolution ms2 data. *Journal of Proteome Research*, 17:3644–3656, 2018.

[20] D. H. May, K. Tamura, and W. S. Noble. Param-Medic: A tool for improving MS/MS database search yield by optimizing parameter settings. *Journal of Proteome Research*, 16(4):1817–1824, 2017.

[21] S. McIlwain, K. Tamura, A. Kertesz-Farkas, C. E. Grant, B. Diament, B. Frewen, J. J. Howbert, M. R. Hoopmann, L. Käll, J. K. Eng, M. J. MacCoss, and W. S. Noble. Crux: rapid open source protein tandem mass spectrometry analysis. *Journal of Proteome Research*, 13(10):4488–4491, 2014.

[22] E. D. Merkley, S. C. Jenson, J. S. Arce, A. M. Melville, O. P. Leiser, D. S. Wunschel, and K. L. Wahl. Ricin-like proteins from the castor plant do not influence liquid chromatography-mass spectrometry detection of ricin in forensically relevant samples. *Toxicon*, 140:18–31, 2017.

[23] C. Y. Park, A. A. Klammer, L. Käll, M. P. MacCoss, and W. S. Noble. Rapid and accurate peptide identification from tandem mass spectra. *Journal of Proteome Research*, 7(7):3022–3027, 2008.

[24] Y. Perez-Riverol, A. Csordas, J. Bai, M. Bernal-Llinares, S. Hewapathirana, D. J. Kundu, A. Inuganti, J. Griss, G. Mayer, M. Eisenacher, E. Pérez, J. Uszkoreit, J. Pfeuffer, T. Sachsenberg, S. Yilmaz, S. Tiwary, J. Cox, E. Audain, M. Walzer, A. F. Jarnuczak, T. Ternent, A. Brazma, and J. A. Vizcaíno. The PRIDE database and related tools and resources in 2019: improving support for quantification data. *Nucleic Acids Res*, 47(D1):D442–D450, 01 2019.

[25] M. M. Savitski, M. Wilhelm, H. Hahne, B. Kuster, and M. Bantscheff. A scalable approach for protein false discovery rate estimation in large proteomic data sets. *Molecular & Cellular Proteomics*, 14(9):2394–2404, 2015.

[26] P. Sulimov and A. Kertész-Farkas. Tailor: A nonparametric and rapid score calibration method for database search-based peptide identification in shotgun proteomics. *Journal of Proteome Research*, 19(4):1481–1490, 2020.

[27] M. The, A. Tasnim, and L. Käll. How to talk about protein-level false discovery rates in shotgun proteomics. *Proteomics*, 16(18):2461–2469, 2016.

[28] M. Wilhelm, J. Schlegl, H. Hahne, A. Moghaddas Gholami, M. Lieberenz, M. M Savitski, E. Ziegler, L. Butzmann, S. Gessulat, H. Marx, T. Mathieson, S. Lemeer, K. Schnatbaum, U. Reimer, H. Wenschuh, M. Mollenhauer, J. Slotta-Huspenina, J. Boese, M. Bantscheff, A. Gerstmair, F. Faerber, and B. Kuster. Mass-spectrometry-based draft of the human proteome. *Nature*, 509(7502):582–587, 2014.