

1

2 *Prepared for Biorxiv submission 8th October 2019*

3

4 **An experimentally-derived measure of inter-replicate variation in reference samples:**
5 **the same-same permutation methodology**

6

7 **David C. Handler and Paul A. Haynes***

8

9 Department of Molecular Sciences, Macquarie University, North Ryde, NSW 2109,
10 Australia

11

12 Email - david.handler@mq.edu.au

13 Email - paul.haynes@mq.edu.au

14

15 * Correspondence: Professor Paul A. Haynes, Department of Molecular Sciences,
16 Macquarie University, North Ryde, NSW 2109, Australia

17 Phone: 61-2-9850 6258

18 Fax: 61-2-9850 6200

19 Email - paul.haynes@mq.edu.au

20

21

22

23

24

25 **Abstract**

26 The multiple testing problem is a well-known statistical stumbling block in high-
27 throughput data analysis, where large scale repetition of statistical methods introduces
28 unwanted noise into the results. While approaches exist to overcome the multiple testing
29 problem, these methods focus on theoretical statistical clarification rather than incorporating
30 experimentally-derived measures to ensure appropriately tailored analysis parameters. Here,
31 we introduce a method for estimating inter-replicate variability in reference samples for a
32 quantitative proteomics experiment using permutation analysis. This can function as a
33 modulator to multiple testing corrections such as the Benjamini-Hochberg ordered Q value
34 test. We refer to this as a ‘same-same’ analysis, since this method incorporates the use of six
35 biological replicates of the reference sample and determines, through non-redundant triplet
36 pairwise comparisons, the level of quantitative noise inherent within the system. The method
37 can be used to produce an experiment-specific Q value cut-off that achieves a specified false
38 discovery rate at the quantitation level, such as 1%. The same-same method is applicable to
39 any experimental set that incorporates six replicates of a reference sample. To facilitate
40 access to this approach, we have developed a same-same analysis R module that is freely
41 available and ready to use via the internet.

42

43 **Keywords:** Label-free shotgun proteomics, false discovery rates, data quality, data
44 validation, statistics

45

46

47

48

49

50 **1. Introduction**

51

52 Shotgun proteomics experiments that seek to compare ‘reference’ and ‘treated’ states of
53 a given sample will often contain thousands of individual comparisons, each requiring
54 statistical validation. In such circumstances, repeated use of the Student’s t-test will
55 invariably introduce false discoveries into the results. A Student’s t-test with a significance
56 cut-off threshold P value of 0.05 produces 95% confidence i.e. 5% of tests will have an equal
57 likelihood to be attributable to chance instead of experimental factors [1-4]. This is a follow-
58 on conclusion from the probability that at least one test in an experiment will be significant,
59 which is described as $1 - (1 - \alpha)^k$, where α is the significance cut-off and k is the number of
60 tests conducted [5].

61 Multiple Testing Corrections (MTCs) were introduced to help address this limitation [6],
62 including the use of Q values rather than P values [7], Bonferroni correction [8,9], Benjamini-
63 Hochberg (BH) adjusted t-test [10], Bonferroni-Holm test [11], and the Benjamini-Yoav
64 (BY) test [12]. MTCs, however, are often overly conservative and increase the false negative
65 rate by eliminating otherwise valid protein identifications. This is especially a problem at the
66 protein-quantitation level; MTCs, by their nature, contribute to a lessening of the protein
67 quantitation false discovery rate (PQ-FDR) at the expense of otherwise valid protein
68 identifications [13,14]. In the context of this article, PQ-FDR is defined as false discoveries
69 arising from comparative quantitative proteomics calculations between one or more samples.
70 There is always a balance to be struck between stringency and accuracy when controlling
71 false discoveries at the protein quantitation level [15]. A recent study in this area applied
72 Bayesian statistics to great effect, detecting a greater number of relevant protein quantitation
73 changes in previously published data sets [16]. There are also numerous software packages
74 available which incorporate various other MTC approaches, including Proteus [17], DAPAR

75 and ProStaR [18], MSqRob [19], UbiA-MS [20], ProteoSign [21], msVolcano [22], FDRtool
76 [23], MSstats [24], and limma [25].

77 Although the use of MTC correction methods in the proteomics field is not standardized
78 [9,26-28], MTCs are an important tool that researchers can employ for extracting the best
79 results from their dataset i.e. finding the balance between reducing noise without losing
80 signal. This desire to reduce the noise in the system led us to ask the question: is there a better
81 way to quantify variability between replicate analyses of a reference sample ?

82 One established approach for assessing variability across a sample set is to use
83 permutation analysis, based upon the Significance Analysis of Micro-array (SAM)
84 permutation methodology [29,30]. This is a similar theoretical framework to that used for
85 permutation analysis within Perseus [31], a well-established data analysis program in the
86 MaxQuant environment [32]. The SAM permutation analysis method assigns a score to each
87 gene on the basis of change in gene expression relative to the standard deviation of repeated
88 measurements. SAM then uses redundant permutations of repeated measurements to estimate
89 the percentage of genes identified by chance, which is used to calculate the false discovery
90 rate. The permutations are performed across all of the ‘reference’ and ‘treated’ sample
91 replicates within a given experimental data set. Those genes with scores higher than the
92 specified threshold are deemed potentially significant, and the threshold can be adjusted to
93 identify smaller or larger sets of genes, with FDR calculated for each set.

94 The same – same method introduced in this study, in contrast, employs non-redundant
95 permutations of experimentally repeated measurements of protein abundance in replicate
96 analysis of a defined reference sample. The permutations are performed on data from the
97 reference samples only, isolated from the ‘treated’ samples This is used to generate a single
98 average Q value indicative of the degree of variation of abundance across the reference
99 sample replicates. Proteins which reach a defined statistical significance threshold are
100 deemed to be false discoveries at the protein quantitation level, since comparing a reference

101 sample against itself should theoretically yield no changes in protein abundance. It is
102 important to emphasize that the underlying assumption is that the biological variability
103 between reference samples is zero, so this approach is accounting for the technical variability.
104 This facilitates subsequent assessment of induced biological variation between reference
105 samples and treated samples.

106 A specified false discovery rate in the same – same analysis of replicates of the reference
107 sample is used to generate a Q value threshold, and that value can then be carried forward to
108 the subsequent analysis of a reference sample versus ‘treated’ sample within the same larger
109 experimental data set. One of main the applications of this method for determining an
110 experimentally-derived measure of reference sample variability is that it can subsequently be
111 used to modify an existing MTC protocol for downstream analysis, thus minimising the PQ-
112 FDR without introducing false negatives. By performing a specific permutation analysis to
113 measure the variability inherent within reference sample replicates, we can produce an
114 experimentally modulated Q value threshold for use with MTCs when comparing the
115 reference sample to treated samples. In essence, rather than using a default Q value of .05, or
116 choosing a more stringent value, we are employing a Q value threshold that is experimentally
117 determined for each set of samples analyzed. The same-same method represents another tool
118 in the proteomics toolbox, and can be used to enable the extraction of additional biological
119 knowledge from large-scale datasets.

120

121

122 **2. Materials and Methods**

123

124 ***2.1 Label free quantitative proteomics data sets***

125 To demonstrate the utility of the sam-same approach we reanalyzed two sets of previously
126 published label free quantitative shotgun proteomics data. Protein identification and
127 Normalized Spectral Abundance Factor (NSAF) values [33,34] were sourced from

128 previously published studies from our laboratory on cultured Cabernet Sauvignon grape cells
129 grown at different temperatures [35] (ProteomeXchange identifier PXD000977) and leaf
130 tissue of IAC1131 rice plants exposed to drought stress [36] (ProteomeXchange identifier
131 PXD004096). The cultured Cabernet Sauvignon grape cell data consists of six biological
132 replicates of cell cultures maintained at 26°C, as the optimum, or control, temperature, and
133 biological triplicate analysis of cells maintained at 18°C and 10°C as moderate and extreme
134 cold stress conditions, and 34°C and 42°C as moderate and extreme heat stress conditions.
135 The rice leaf data consists of two sets of three biological replicates each of unstressed plants
136 as controls, and biological triplicate analysis of plants exposed to moderate drought stress,
137 extreme drought stress, and extreme drought stress followed by recovery.

138 *2.2 Same – same permutation analysis of reference samples*

139 For analysis using the same-same workflow, six replicates of a reference sample are run
140 through a PSM (peptide-to-spectrum matching) engine and protein identification lists are
141 exported as csv files. Next, these six replicates are grouped into two sets of triplicates by the
142 use of inner joins (dummy state ‘control’, dummy state ‘treatment’), and a test array is formed
143 through a full join of the states [37]. One hundred Student’s t-tests are conducted on each
144 identified protein comparison with significance cut-off values from 0.01 to 1, stepped at 0.01
145 intervals. All proteins found at different quantitation levels are considered false discoveries,
146 since comparison between two data sets of the same sample type would theoretically give
147 identical quantitation with no observed changes. This process is repeated for all ten
148 combinations of non-redundant triplet pairs that six replicates can form. The MTC analysis
149 then begins by iterating over this 10x t-test array and applying one of five user-specified
150 MTC methods (BH, Benjamini-Yoav, Bonferroni, Hommel [38], and Bioconductor Q [39]).
151 The program then averages the MTC test results from all arrays examined, and reports the
152 point at which the significance cut-off corresponds to a user-specified PQ-FDR.

153 The same-same methodology is automated through an R script. Source code is available
154 from <https://bitbucket.org/peptidewitch/samesame/>, and a freely accessible working web
155 version can be found at <https://peptidewitch.shinyapps.io/samesame>. The R Shiny web-app
156 provides three distinct outputs from the same-same analysis:

- 157 1) A series of Q value vs FDR bar plots (x axis 0.01 to 1, stepped at 0.01) from all ten
158 triplet paired combinations,
- 159 2) A series of P value histograms of these same combinations, and,
- 160 3) A numerical value that corresponds to the user-specified MTC cut-off that produces
161 the desired PQ-FDR (default 1%).

162 Input data types are not constrained to spectral counts, as in theory any data type that
163 consists of protein identifications coupled with abundance or intensity value measurements
164 can be used. However, the first generation of the analysis tool was designed and tested using
165 spectral counting-based data, so it is recommended that spectral counts or spectral abundance
166 factors be used initially.

167 ***2.3 Perseus permutation analysis of reference samples***

168 To serve as a comparison against the same-same process, the same NSAF data from both
169 Grape and Rice samples as above were reanalyzed using Perseus software [31]. Spreadsheet
170 files containing NSAF values for each set of samples were uploaded to Perseus through a
171 generic matrix upload. Using the two-sample module, we applied the Perseus permutation
172 method as a form of truncation using ungrouped (no grouping preserved), 250-count
173 permutation analysis on two-tailed Student's t-testing arrays with BH correction, comparing
174 six reference replicates with three replicates from each of the 'treated' sample states, with the
175 specified FDR thresholds ranging from 1-5%.

176
177
178
179

180 3. Results

181

182 The following section details how same-same approach was applied to the published data
183 from Grape and Rice cells. Figure 1 displays the outputs described above for the same-same
184 analysis conducted on the grape cell culture label-free data when specifying BH correction
185 and 1% PQ-FDR. The end-point of the same-same process is the modulated Q value, in this
186 example 0.054 (Figure 1C), produced from averaging the threshold values in Figure 1A at
187 the desired PQ-FDR value. This value can be used for downstream analysis on subsequent
188 control vs treatment samples as a modulator for the chosen MTC.

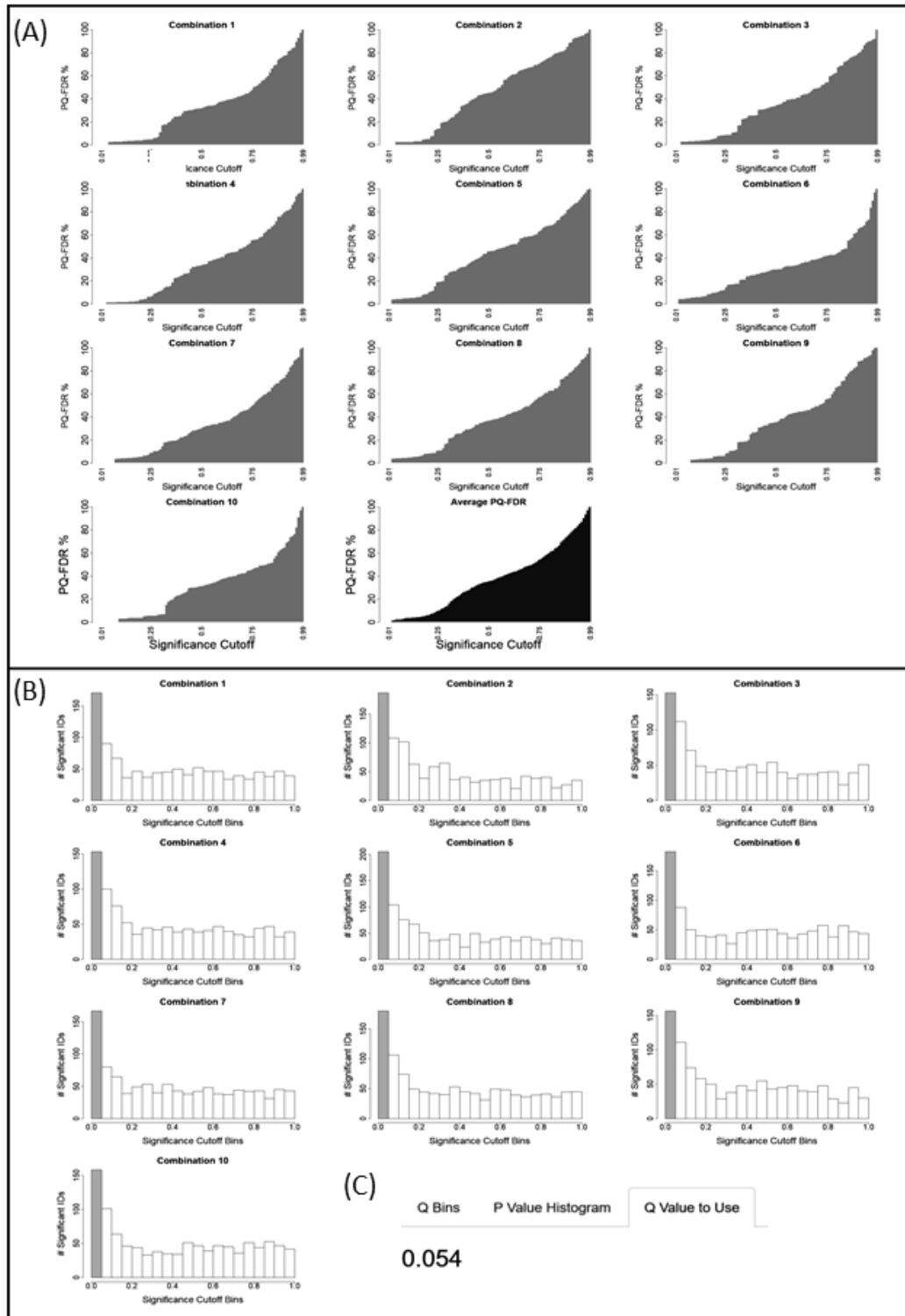
189 Figure 2 presents the subsequent downstream analysis of the grape cell cultures grown at
190 different temperatures. Figure 2A shows the number of proteins found to be significantly
191 differentially expressed in terms of protein fold change when compare the set of six reference
192 replicates to the set of three replicates of cells grown at each temperature. These are analysed
193 using different statistical measures of significance: P values of 0.05 and 0.01 for a student's
194 t-test, BH Q value of 0.05, and BH using the same-same derived Q value (SS-Q), and
195 specifying PQ – FDR of 1%, 2% or 3%. It is evident that the same-same derived Q values at
196 1% PQ-FDR produce results very similar to the use of default BH Q values, which is expected
197 given that the SS-Q value used is very close to the 0.05 BH-Q value threshold. The two
198 approaches give similar results, although it is noticeable that at a specified PQ – FDR of 3%,
199 the comparison with the largest effect size (Figure 2E) shows significantly more differentially
200 expressed proteins.

201

202

203

Figure 1

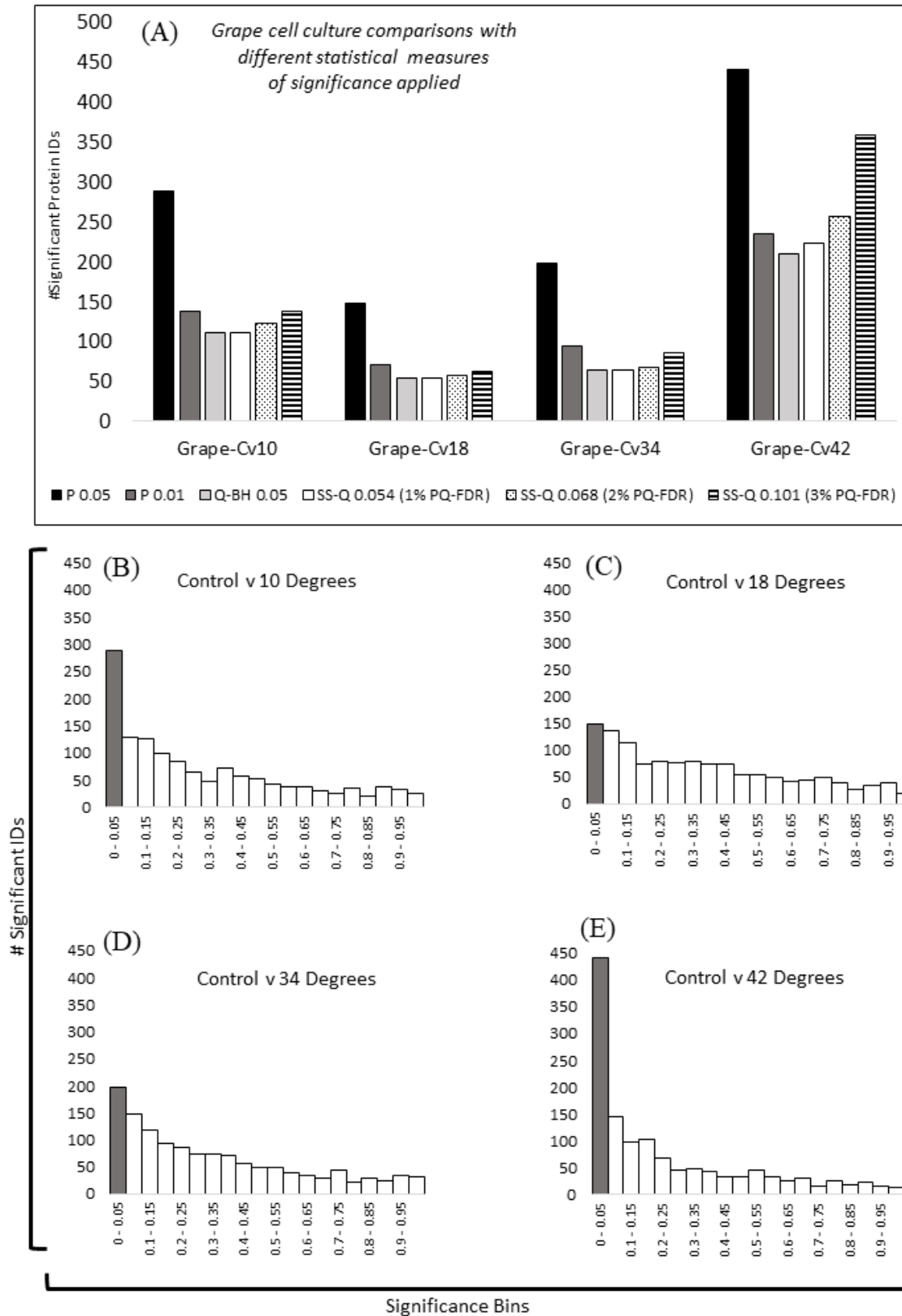


204

205

206

207 **Figure 2**



208

209

210 Figure 3 presents the same type of analyses as shown in Figure 1 for the data derived
211 from comparative analysis of leaf tissue from IAC1131 rice plants exposed to different levels
212 of drought stress. Interestingly, in contrast to figure 2, it is clear that in this case there is a
213 direct correlation between observed effect size and number of differentially expressed
214 proteins identified using the SS-Q approach. In comparisons with greater effect size as
215 observed in P value histograms (Figure 3E,3F,3G), the same-same derived Q Values are able
216 to identify a greater number of differentially expressed proteins than were found using the
217 default BH Q value, and at 3% PQ-FDR are approaching the number of differentially
218 expressed proteins found using uncorrected P values.

219 Table 1 presents the results of analyzing the grape and rice cell NSAF data referred to
220 above using different analysis approaches, including Student t-tests with and without BH
221 correction, application of same – same derived Q values to a BH corrected t-test at specified
222 PQ-FDR values ranging from 1% to 5%, and t-tests using Perseus permutations at specified
223 PQ-FDR values ranging from 1% to 5%. The table shows the number of proteins which are
224 reported to be significantly differentially expressed when comparing the reference samples
225 against the grape cells grown at four different temperatures, and the rice cells grown under
226 three different watering regimes. It is clear from these comparisons that, as expected, the
227 uncorrected student's t-test gives a much greater number than any sort of correction. The BH
228 correction reduces the number of significant proteins by approximately 95%. The Perseus
229 permutation processing is even more strict and, for example, produces zero significant
230 identifiers in more than half of the grape sample comparisons. In contrast, the same-same-

231 modulated BH test is able to detect significantly differentially regulated proteins for every
 232 test case for both tissue types while always remaining well below the results reported from
 233 uncorrected Student's t-testing P values. Multiple testing correction still takes place, but the
 234 experimentally derived Q value thresholds allow for the recovery of a greater number of
 235 significant differences at the protein quantitation level.

236

237

238 **Table 1 – Comparison of number of protein identifications retained using different**
 239 **analysis approaches to assess protein quantitation false discovery rate**

Comparisons	t-test ^a		Perseus permutations BH					Same-Same BH				
	.05	BH	1% ^b	2%	3%	4%	5%	1%	2%	3%	4%	5%
Grape												
Cont vs 10 C	735	42	0	0	0	0	0	48	54	72	91	99
Cont vs 18 C	560	29	0	0	0	9	9	29	34	43	51	65
Cont vs 34 C	432	21	0	0	0	0	6	24	28	36	42	48
Cont vs 42 C	775	83	0	0	0	0	44	108	116	133	166	199
Rice												
Cont vs Ext	543	12	0	0	0	0	0	15	16	17	17	17
Cont vs Mod	481	13	0	0	0	0	0	13	17	20	20	20
Cont vs Recov	568	21	0	0	0	0	0	22	33	40	40	40

240 ^a .05 = standard 2-sample t-test, BH = Benjamin-Hochberg corrected 2 sample t-test

241 ^b protein quantitation false discovery rate assessed at 1% -5% using the approaches indicated

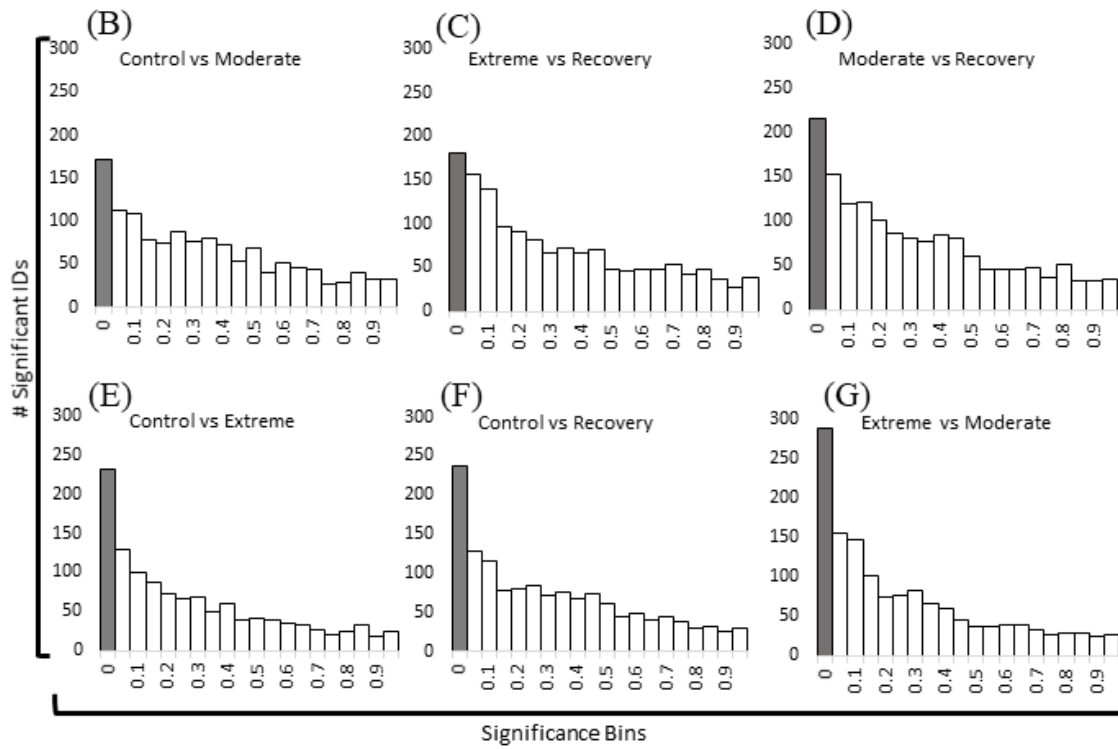
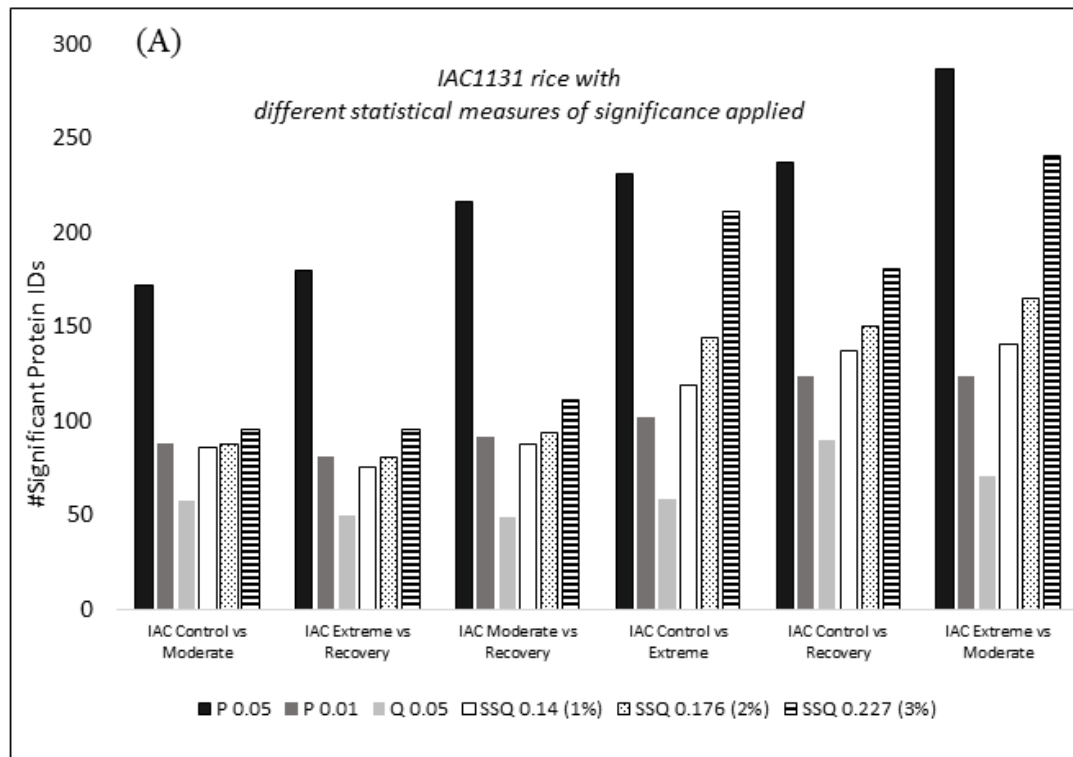
242

243

244

245

Figure 3



246

247

248

249

4. Discussion

250

251

252

253

254

255

256

257

258

259

260

261

262

263

264

265

266

267

268

269

270

271

272

273

274

The correlation observed between effect size and number of differentially expressed proteins found in the dataset presented in Figure 3 has also been found in numerous other datasets we have analysed. In general, SS-Q values are generally better suited to those datasets that show a larger effect size. This may be due to the fact that not all quantitatively different proteins in a small effect sample are false positives, or may be a consequence of NSAFs overstating expression change ratios for protein identifications based on lower spectral counts, which can help to increase the effect size [4]. While the use of higher Q value thresholds raises the implicit question of whether or not the dataset contains too much noise, it is important to remember why the same-same experiment is conducted in the first place. If, in an experiment where we expect there to be minimal noise, we demonstrate that there is a SS-Q threshold value that produces 1% PQ-FDR between sets of control or reference replicates, then in a closely related experiment with the same reference sample using the same threshold value, we can infer experimentally that the specified PQ-FDR has been achieved.

It is important to stress, however, that this method is suited more towards initial discovery, and that follow-up experimentation must employ orthogonal validation protocols. In order to obtain an experimentally-derived PQ-FDR of 1%, or other specified value, the same-same method is a very useful tool, because inferring the PQ-FDR based on the Q value cut-off alone does not yield corresponding PQ-FDR levels (i.e. a Q value of 0.05 does not specifically produce either 5% or 1% PQ-FDR). Modifying the MTC significance value cut-off so that it takes into account the experimental variability inherent within the replicates helps to produce a more tailored list of differentially expressed protein identifications whilst controlling for PQ-FDR. Also, compounding the same-same technique with another method of filtering, such as fold change cut-offs, can reduce the number of false positives included in the final dataset, further reducing the PQ-FDR [30,31].

275 In this research article, we have demonstrated a revised method for statistical analysis for
276 shotgun proteomics datasets. The same-same method facilitates the construction of post
277 analysis P value histograms and aids the researcher in choosing an appropriate statistical
278 testing protocol for their analysis. We have shown that in the right circumstances, using BH
279 Q value cut-offs derived from the same-same analysis yields a set of results that provide more
280 significantly differentially expressed proteins from a given dataset, while also determining
281 PQ-FDR at the experimental level. In the future, we hope to expand on this methodology so
282 that it can be applied equally well to other quantitative proteomics data types, and also
283 develop new tests to build onto the existing same-same architecture to further improve the
284 statistical rigour for all shotgun proteomics results.

285

286 -----||-----

287

288 **5. Acknowledgments**

289 The authors acknowledge support from BioPlatforms Australia through the Australian
290 Government's National Collaborative Research Infrastructure Scheme, and the Australian
291 Research Council through Discovery Project DP190103140. The authors declare no conflict
292 of interest.

293

294 **6. Author contributions**

295 DCH and PAH contributed equally to the conceptualization and design of this study.
296 DCH was responsible for writing the software described. DCH prepared the original draft
297 of the manuscript, and PAH was responsible for revision and editing.

298

299 **7. Figure Legends**

300 **Figure 1.** Screenshots from the same/same shiny apps module
301 (<https://peptidewitch.shinyapps.io/SameSame>), using the grape cell control samples and
302 specifying BH correction at 1% PQ-FDR. (A) Q value vs PQ-FDR bar plots (x axis 0.01 to
303 1, stepped at 0.01) for all ten triplet paired permutations generated from six replicate analyses
304 of a reference sample (see Figure 1) (B) P value histograms for each permutation, showing
305 number of significantly expressed protein identifications sorted into P values bins in
306 increments of 0.05. (C) displays a single numerical value which produces the desired PQ-
307 FDR value (default BH at 1%, can be user specified).

308 **Figure 2.** Grape cell culture comparisons with application of different statistical
309 significance measures. Cells grown at 26°C were designated as the reference sample, and
310 compared with cells grown at 18°C (moderate cold), 10°C (extreme cold), 34°C (moderate

311 heat), and 42°C (extreme heat). Panel A displays the number of significantly differentially
312 expressed protein identifications found for each comparison using P values at .05 and .01,
313 Benjamini-Hochberg adjusted values at 0.05, and BH using the same-same derived Q value
314 (SS-Q), and specifying PQ – FDR of 1%, 2% or 3%. Panels B to E contain P value histograms
315 showing the number of significantly expressed protein identifications sorted into P value bins
316 in increments of 0.05, for each of the four experimental comparisons performed, as indicated.

317 **Figure 3.** IAC1131 rice samples drought stress comparisons with application of different
318 statistical significance measures. Control plants were unstressed, and compared with plants
319 exposed to moderate drought stress, extreme drought stress, or extreme drought stress
320 followed by recovery. Panel A displays the number of significantly differentially expressed
321 protein identifications found for each comparison using P values at .05 and .01, Benjamini-
322 Hochberg adjusted values at 0.05, and BH using the same-same derived Q value (SS-Q), and
323 specifying PQ – FDR of 1%, 2% or 3%. Panels B to G are P value histograms showing the
324 number of significantly expressed protein identifications sorted into P value bins in
325 increments of 0.05, for each of the six experimental comparisons performed, as indicated.

326

327

328

329

330

331

332

333 **8. References**

- 334 1. Colquhoun, D. An investigation of the false discovery rate and the misinterpretation
335 of p-values. *R Soc Open Sci* **2014**, *1*, 140216.
- 336 2. Diz, A.P.; Carvajal-Rodriguez, A.; Skibinski, D.O. Multiple hypothesis testing in
337 proteomics: A strategy for experimental work. *Mol Cell Proteomics* **2011**, *10*, M110
338 004374.
- 339 3. Karp, N.A.; Lilley, K.S. Design and analysis issues in quantitative proteomics studies.
340 *Proteomics* **2007**, *7* 42-50.
- 341 4. Sullivan, G.M.; Feinn, R. Using effect size-or why the p value is not enough. *J Grad*
342 *Med Educ* **2012**, *4*, 279-282.
- 343 5. Streiner, D.L.; Norman, G.R. Correction for multiple testing: Is there a resolution?
344 *Chest* **2011**, *140*, 16-18.
- 345 6. Bender, R.; Lange, S. Adjusting for multiple testing--when and how? *J Clin*
346 *Epidemiol* **2001**, *54*, 343-349.
- 347 7. Armstrong, R.A. When to use the bonferroni correction. *Ophthalmic Physiol Opt*
348 **2014**, *34*, 502-508.
- 349 8. Choi, H.; Nesvizhskii, A.I. False discovery rates and related statistical concepts in
350 mass spectrometry-based proteomics. *J Proteome Res* **2008**, *7*, 47-50.
- 351 9. Benjamini, Y.; Hochberg, Y. Controlling the false discovery rate: A practical and
352 powerful approach to multiple testing. *J Royal Stat Society, Series B* **1995**, 289-300.
- 353 10. Holm, S. A simple sequentially rejective multiple test procedure. *Scandinavian*
354 *journal of statistics* **1979**, *6*, 65-70.
- 355 11. Benjamini, Y.; Yekutieli, D. The control of the false discovery rate in multiple testing
356 under dependency. *Annals of statistics* **2001**, *29*, 1165-1188.
- 357 12. Carvajal-Rodriguez, A.; de Una-Alvarez, J. Assessing significance in high-
358 throughput experiments by sequential goodness of fit and q-value estimation. *PLoS*
359 *One* **2011**, *6*, e24700.
- 360 13. Pascovici, D.; Handler, D.C.; Wu, J.X.; Haynes, P.A. Multiple testing corrections in
361 quantitative proteomics: A useful but blunt tool. *Proteomics* **2016**, *16*, 2448-2453.
- 362 14. Noble, W.S. How does multiple testing correction work? *Nat Biotechnol* **2009**, *27*,
363 1135-1137.
- 364 15. The, M.; Kall, L. Integrated identification and quantification error probabilities for
365 shotgun proteomics. *Mol Cell Proteomics* **2019**, *18*, 561-570.
- 366 16. Gierlinski, M.; Gastaldello, F.; Barton, G.J. Proteus: An r package for downstream
367 analysis of maxquant output. *BioRxiv* **2018**, doi.org/doi10.1101/416511.
- 368 17. Wiczorek, S.; Combes, F.; Lazar, C.; Gianetto, Q.G.; Gatto, L.; Dorffer, A.; Hesse,
369 A.M.; Coute, Y.; Ferro, M.; Bruley, C., *et al.* Dapar & prostar: Software to perform
370 statistical analyses in quantitative discovery proteomics. *Bioinformatics* **2017**, *33*,
371 135-136.
- 372 18. Goeminne, L.J.E.; Gevaert, K.; Clement, L. Experimental design and data-analysis in
373 label-free quantitative lc/ms proteomics: A tutorial with msqrob. *J Proteomics* **2018**,
374 *171*, 23-36.

- 375 19. Zhang, X.; Smits, A.H.; van Tilburg, G.B.; Ovaa, H.; Huber, W.; Vermeulen, M.
376 Proteome-wide identification of ubiquitin interactions using ubia-ms. *Nat Protoc*
377 **2018**, *13*, 530-550.
- 378 20. Efstathiou, G.; Antonakis, A.N.; Pavlopoulos, G.A.; Theodosiou, T.; Divanach, P.;
379 Trudgian, D.C.; Thomas, B.; Papanikolaou, N.; Aivaliotis, M.; Acuto, O., *et al.*
380 Proteosign: An end-user online differential proteomics statistical analysis platform.
381 *Nucleic Acids Res* **2017**, *45*, W300-W306.
- 382 21. Singh, S.; Hein, M.Y.; Stewart, A.F. Msvolcano: A flexible web application for
383 visualizing quantitative proteomics data. *Proteomics* **2016**, *16*, 2491-2494.
- 384 22. Strimmer, K. A unified approach to false discovery rate estimation. *BMC*
385 *Bioinformatics* **2008**, *9*, 303.
- 386 23. Choi, M.; Chang, C.Y.; Clough, T.; Broudy, D.; Killeen, T.; MacLean, B.; Vitek, O.
387 Msstats: An r package for statistical analysis of quantitative mass spectrometry-based
388 proteomic experiments. *Bioinformatics* **2014**, *30*, 2524-2526.
- 389 24. Ritchie, M.E.; Phipson, B.; Wu, D.; Hu, Y.; Law, C.W.; Shi, W.; Smyth, G.K. Limma
390 powers differential expression analyses for rna-sequencing and microarray studies.
391 *Nucleic Acids Res* **2015**, *43*, e47.
- 392 25. Gelman, A.; Hill, J.; Yajima, M. Why we (usually) don't have to worry about multiple
393 comparisons. *J Res Educ Effect* **2012**, *5*, 189-211.
- 394 26. Kall, L.; Storey, J.D.; MacCoss, M.J.; Noble, W.S. Posterior error probabilities and
395 false discovery rates: Two sides of the same coin. *J Proteome Res* **2008**, *7*, 40-44.
- 396 27. Rothman, K.J. No adjustments are needed for multiple comparisons. *Epidemiology*
397 **1990**, *1*, 43-46.
- 398 28. Gierlinski, M.; Cole, C.; Schofield, P.; Schurch, N.J.; Sherstnev, A.; Singh, V.;
399 Wrobel, N.; Gharbi, K.; Simpson, G.; Owen-Hughes, T., *et al.* Statistical models for
400 rna-seq data derived from a two-condition 48-replicate experiment. *Bioinformatics*
401 **2015**, *31*, 3625-3630.
- 402 29. Tusher, V.G.; Tibshirani, R.; Chu, G. Significance analysis of microarrays applied to
403 the ionizing radiation response. *Proc Natl Acad Sci U S A* **2001**, *98*, 5116-5121.
- 404 30. Tyanova, S.; Temu, T.; Sinitcyn, P.; Carlson, A.; Hein, M.Y.; Geiger, T.; Mann, M.;
405 Cox, J. The perseus computational platform for comprehensive analysis of
406 (prote)omics data. *Nat Methods* **2016**, *13*, 731-740.
- 407 31. Tyanova, S.; Temu, T.; Cox, J. The maxquant computational platform for mass
408 spectrometry-based shotgun proteomics. *Nat Protoc* **2016**, *11*, 2301-2319.
- 409 32. Neilson, K.A.; Ali, N.A.; Muralidharan, S.; Mirzaei, M.; Mariani, M.; Assadourian,
410 G.; Lee, A.; van Sluyter, S.C.; Haynes, P.A. Less label, more free: Approaches in
411 label-free quantitative mass spectrometry. *Proteomics* **2011**, *11*, 535-553.
- 412 33. Neilson, K.A.; Keighley, T.; Pascovici, D.; Cooke, B.; Haynes, P.A. Label-free
413 quantitative shotgun proteomics using normalized spectral abundance factors.
414 *Methods Mol Biol* **2013**, *1002*, 205-222.
- 415 34. George, I.S.; Pascovici, D.; Mirzaei, M.; Haynes, P.A. Quantitative proteomic
416 analysis of cabernet sauvignon grape cells exposed to thermal stresses reveals

- 417 alterations in sugar and phenylpropanoid metabolism. *Proteomics* **2015**, *15*, 3048-
418 3060.
- 419 35. Wu, Y.; Mirzaei, M.; Pascovici, D.; Chick, J.M.; Atwell, B.J.; Haynes, P.A.
420 Quantitative proteomic analysis of two different rice varieties reveals that drought
421 tolerance is correlated with reduced abundance of photosynthetic machinery and
422 increased abundance of clpd1 protease. *J Proteomics* **2016**, *143*, 73-82.
- 423 36. Mishra, P.; Eich, M. Join processing in relational databases. *CM Computing Surveys*
424 **1992**, *24*, 63-113.
- 425 37. Hommel, G.; Krummenauer, F. Improvements and modifications of tarone's multiple
426 test procedure for discrete data. *Biometrics* **1998**, *54*, 673-681.
- 427 38. Qian, H.R.; Huang, S. Comparison of false discovery rate methods in identifying
428 genes with differential expression. *Genomics* **2005**, *86*, 495-503.
- 429