Supplement to
"Statistical calibration of the SEQUEST XCorr function"

Aaron A. Klammer
Department of Genome Sciences
University of Washington
Seattle, WA, USA

Christopher Y. Park
Department of Genome Sciences
University of Washington
Seattle, WA, USA

William Stafford Noble
Department of Genome Sciences
Department of Computer Science and Engineering
University of Washington
Seattle, WA, USA

1 Results

1.1 Selection of fraction of distribution tail to fit

In the Methods section we described how we calculate p-values from the tail of the score distribution. Here, we describe our exact method for selecting how much of the tail to fit. We enumerated all distribution fractions between 0.1 and 1.0, incremented by 0.1 and evaluated each fraction using two figures of merit. The first is the slope of the QQ plot shown in Figure 3; the closer the p-value slope error is to unity, the more uniform the calculated p-values are, and the better our fraction of fit peptides is considered to be (1). The second is the number of positive PSMs at 5% false discovery rate; for this figure of merit, higher is better. We selected the best fraction separately for Sp and $Xcorr$ using 1000 held out spectra (Figure 1), and validated this selection on two other held out sets of 1000 spectra (Figure 1). The held-out data sets were not used in subsequent analysis. We visually inspected the plots and selected tail fractions of 0.40 and 0.55, for Sp and $Xcorr$, respectively.

References

[1] Bailey, T., Gribskov, M.: Estimating and evaluating the statistics of gapped local-alignment scores. Journal of Computational Biology **9**(3) (2002) 575–593

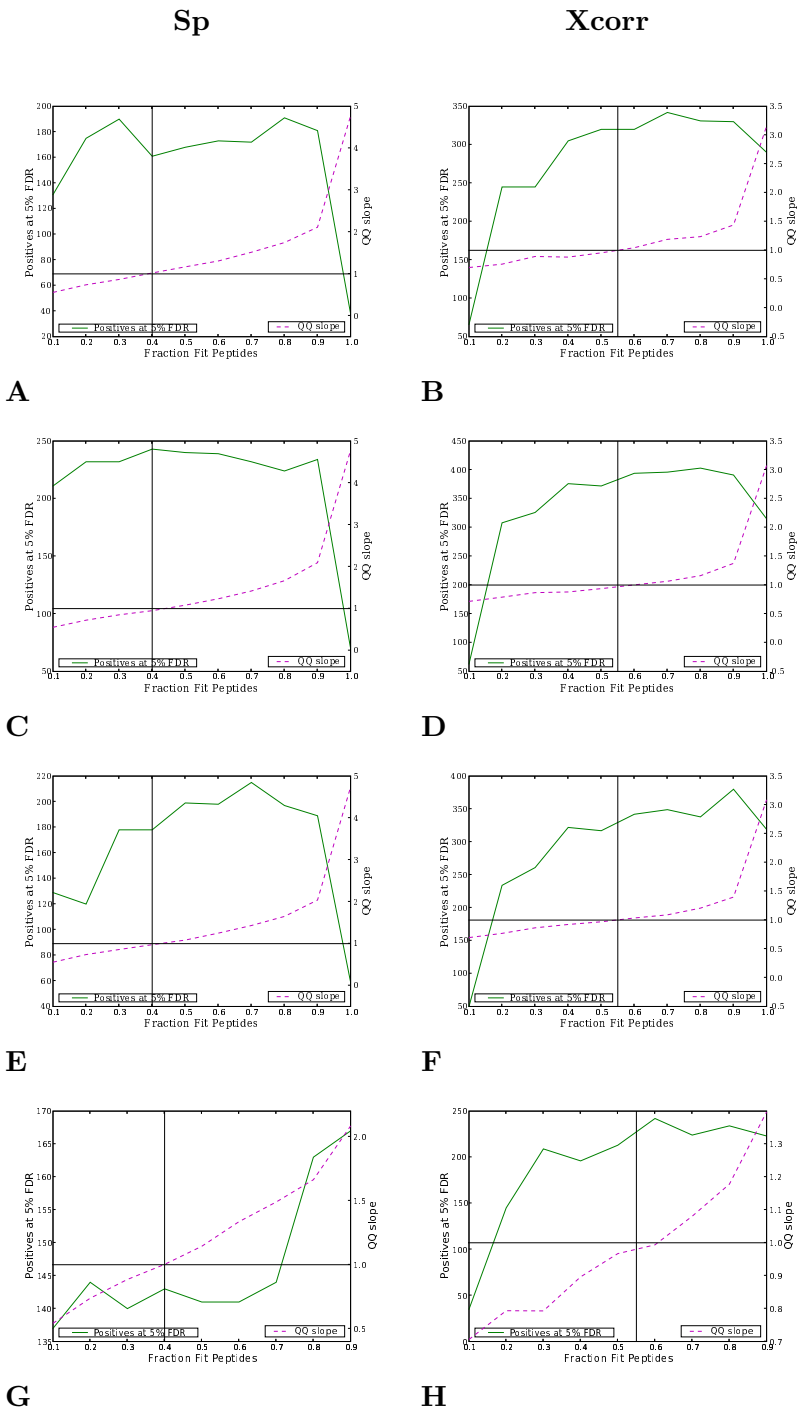

Figure 1: **Evaluation of different fractions of the distribution tail to fit.** Varying the portion of the peptide score distribution tail to fit affects the number of positive peptides at a 5% FDR (green lines) and the slope of the QQ plot (dashed magenta lines) for *Sp* (A) and *Xcorr* (B). A perfect QQ plot has a slope of 1.0 (horizontal black lines). After visual inspection, we selected a fraction of 0.40 for *Sp* and 0.55 for *Xcorr* (vertical black lines). Shown also are the same plots for two other subsets of 1000 spectra (C, D, E and F), demonstrating that the thresholds selected in (A) and (B) generalize. Finally the bottom panels show results on a completely different data set (G, H).