



**Figure S8.** ROC curves when highly homologous proteins [1] are also counted as true positive proteins. Plots done this way are analogous to the ROC plots obtained using a decoy database to estimate the number of false positives. Each panel displays the results of a scoring function. The resulting ROC curves from using RAID<sub>aPS</sub> implementation and the implementation in the original search program are both shown. The results from profile data (NHLBI data set [2]) are shown in solid curves, while the results from centroid data (A1-A4 of ISB data set [3]) are shown in long-dash curves. Panels (A,B,C,D) respectively display the results from using RAID score, K-score, XCorr, and Hyperscore. Except for RAID score, the RAID<sub>aPS</sub> implemented scoring functions performs comparably to the original implementation in other search methods.

## References

1. Alves G, Ogurtsov AY, Yu YK (2007) RAID\_DbS: Peptide identification using database searches with realistic statistics. *Biology Direct* 2: 25.
2. Alves G, Ogurtsov AY, Wu WW, Wang G, Shen RF, et al. (2007) Calibrating E-values for MS<sup>2</sup> library search methods. *Biology Direct* 2: 26.
3. Keller A, Samuel P, Nesvizhskii AI, Stolyar S, Goodlett DR, et al. (2002) Experimental protein mixture for validating tandem mass spectral analysis. *OMICS* 6: 207-212.