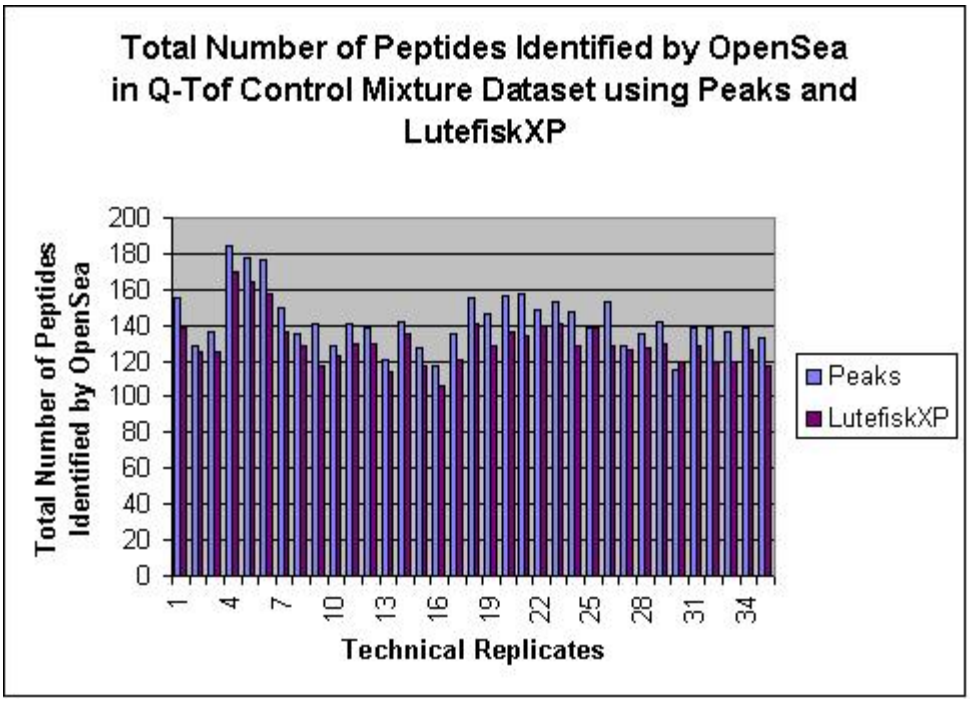
Comparision of Peaks† and LutefiskXP‡ Using OpenSea and Control Mixture Dataset Acquired on Q-Tof

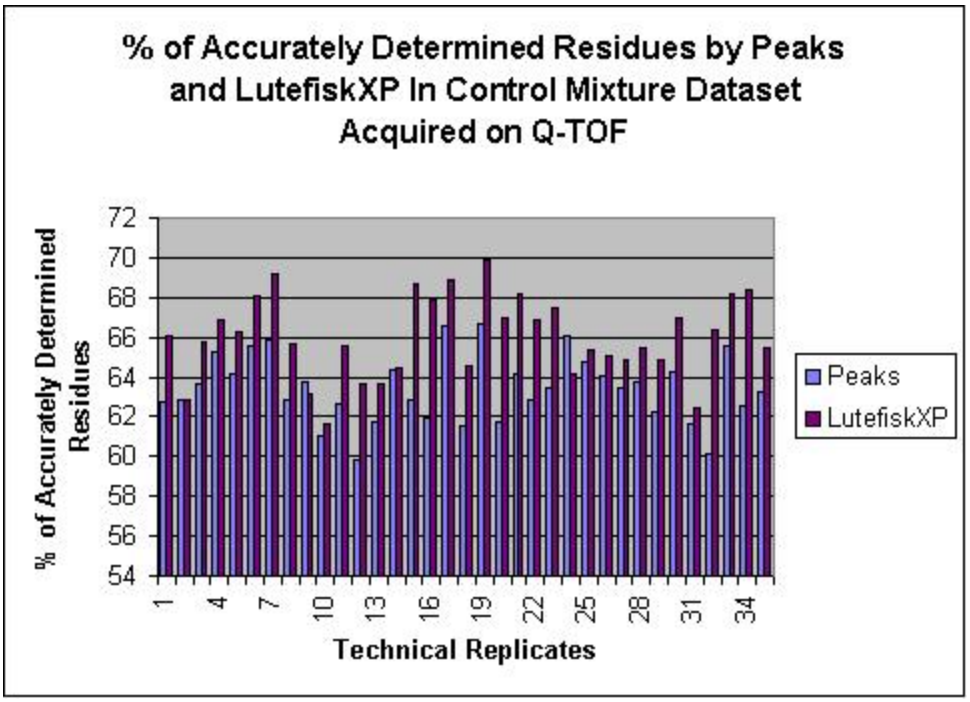
Figure 1



	Peaks	LutefiskXP
Average Number of ID's Made by OpenSea[1]	142.65	130.54

Thirty five technical replicates, each containing 10 control proteins (*Bos Taurus* insulin, ubiquitin, cytochrome C, superoxide dismutase, beta-lactoglobulin A, serum albumin and immunoglobulin G, as well as *Equus Caballus* myoglobin, *Armoracia rusticana* peroxidase, and *Gallus gallus* conalbumin) were tryptically digested and run on a quadrupole-TOF mass spectrometer (Waters, Milford, MA). The dataset was *de novo* sequenced using both LutefiskXP and Peaks, as described in the manuscript. Peaks and LutefiskXP were compared based on the number of peptide identifications made by OpenSea algorithm in each technical replicate. Peptide identifications that belong to the control proteins, which do not have any substitutions, unknown modifications and alignment errors, were considered for the comparision. When a control protein was absent in the search results, then the peptide identifications of the same protein from other species were accepted for the comparision if they don't have any substitutions, unknown modifications and alignment errors. The total number of peptide identifications were manually determined for each of the thirty five technical replicates and plotted in **figure 1**. The average number of peptide identifications[1] made by OpenSea in the control mixture dataset acquired on an Q-TOF instrument was calculated. The average percentage gain in the peptide identifications made by OpenSea algorithm, when Peaks was used to *de novo* sequence the Q-TOF data, rather than LutefiskXP, was determined to be **9.27%**.

Figure 2



	Peaks	LutefiskXP
Average Percentage of Correctly Sequenced Residues[2]	63.27	65.46

Thirty five techical replicates, each containing 10 control proteins (*Bos Taurus* insulin, ubiquitin, cytochrome C, superoxide dismutase, beta-lactoglobulin A, serum albumin and immunoglobulin G, as well as *Equus Caballus* myoglobin, *Armoracia rusticana* peroxidase, and *Gallus gallus* conalbumin) were tryptically digested and run on a quadrupole-TOF mass spectrometer (Waters, Milford, MA). The dataset was *de novo* sequenced using both LutefiskXP and Peaks, as described in the manuscript. Peaks and LutefiskXP were further compared based on the number of residues that were accurately determined by both programs. OpenSea algorithm was used to process the *de novo* sequencing results of both Peaks and LutefiskXP. Peptide identifications that belong to the control proteins, which do not have any substitutions, unknown modifications and alignment errors, were considered for the comparision. When a control protein was absent in the search results, then the peptide identifications of the same protein from other species were accepted for the comparision if they don't have any substitutions, unknown modifications and alignment errors. The number of accurately determined residues was manually calculated for each peptide identification made by OpenSea algorithm. An amino acid residue in a peptide identification was counted as an accurate residue if OpenSea made an one-to-one alignment* between that residue in the *de novo* sequence and the corresponding residue in the database peptide sequence. Isobaric amino acids were treated as accurately determined even if the *de novo* sequencer reported an equivalent residue instead of the correct residue. The percentage of accurately determined residues, by Peaks and LutefiskXP, in each replicate was calculated and plotted in **figure 2**. The average percentage of residues[2] that were accurately determined by Peaks and LutefiskXP, for all thirty five technical replicates of control mixture dataset acquired on an Q-TOF instrument, was also determined.

† **Johnson, R.S. "Lutefisk1900 vs Peaks: A comparision of automated *de novo* sequencing programs", ABRF'04:Integrating technologies in Proteomics and Genomics, Portland,Oregon,February 28-March 2, 2004**

‡ **Peaks Batch Version 2.2 (Bio-informatics solutions, Ontario, Canada).**

* **Searle, B.C; Dasari, S; Turner, M.; Reddy, A.P.; Chio, D.; Wilmarth, P.A.; McCormack, A.L.; David, L.L.; Nagalla, S.R. *Anal. Chem.* 2004,76,2220-2230.**

This site developed by Surendra Dasari and maintained by Matthew J. Rodland
Last updated on Wednesday, January 3, 2005 at 12:24:08 PM