



PeptideProphet Explained

Brian C. Searle

Proteome Software Inc.

www.proteomesoftware.com

1336 SW Bertha Blvd, Portland OR 97219

(503) 244-6027

An explanation of the Peptide Prophet algorithm developed by Keller, A., Nesvizhskii, A. I., Kolker, E., and Aebersold, R. (2002) *Anal. Chem.* **74**, 5383 –5392



Threshold model

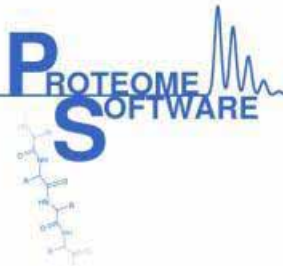
Before PeptideProphet was developed, a threshold model was the standard way of evaluating the peptides matched by a search of MS/MS spectra against a protein database.

The threshold model sorts search results by a match score.

| # | File | MH+ | XCorr | dCn | Sp | RSp | Ions | Ref | Sequence |
|------|---------------------------------------|---------------|--------|-------|--------|-----|--------|-----------------------|-------------------------------|
| 325 | ./sergei_digest_A_full_01.1001.1003.3 | 2511.7 (-0.6) | 7.5063 | 0.450 | 4546.1 | 1 | 42/ 84 | sp P00921 CAH2_BOVIN | R.MVNHGHSFNVEYDDSDQKAVLK.D |
| 462 | ./sergei_digest_A_full_01.1239.1241.3 | 2504.7 (+2.0) | 6.1800 | 0.302 | 3126.2 | 1 | 39/ 84 | sp P00921 CAH2_BOVIN | R.LVGFHFHWGSSBBOGSEHTVDR.K |
| 1070 | ./sergei_digest_A_full_01.2335.2337.3 | 2254.5 (+0.5) | 6.0682 | 0.491 | 2166.9 | 1 | 37/ 84 | sp P00921 CAH2_BOVIN | K.YGDFCTAAQQPDGLAVVGVFLK.V |
| 1105 | ./sergei_digest_A_full_01.2405.2407.3 | 2254.5 (+0.8) | 6.0041 | 0.511 | 1873.6 | 1 | 35/ 84 | sp P00921 CAH2_BOVIN | K.YGDFCTAAQQPDGLAVVGVFLK.V |
| 510 | ./sergei_digest_A_full_01.1317.1325.3 | 2584.7 (+1.5) | 5.9521 | 0.403 | 2488.3 | 1 | 38/ 84 | sp P00921 CAH2_BOVIN | R.LVGFHFHWGSSBBOGSEHTVDR.K |
| 1219 | ./sergei_digest_A_full_01.2617.2619.3 | 2107.6 (-0.1) | 5.7343 | 0.502 | 2282.1 | 1 | 33/ 72 | sp P02666 CASH_BOVIN | R.DHFIQAFLLYGEFVLGPPV.G |
| 894 | ./sergei_digest_A_full_01.2013.2015.3 | 2314.7 (+0.0) | 5.5636 | 0.410 | 1260.3 | 1 | 33/ 76 | sp P02754 LACB_BOVIN | R.VYVEELKFTPEGDLLEILLQK.W |
| 1142 | ./sergei_digest_A_full_01.1873.1875.3 | 2314.7 (+0.5) | 5.5466 | 0.428 | 1407.6 | 1 | 35/ 76 | sp P02754 LACB_BOVIN | R.VYVEELKFTPEGDLLEILLQK.W |
| 1142 | ./sergei_digest_A_full_01.2471.2475.3 | 2254.5 (-0.8) | 5.5372 | 0.526 | 771.0 | 1 | 24/ 42 | sp P00921 CAH2_BOVIN | K.YGDFCTAAQQPDGLAVVGVFLK.V |
| 836 | ./sergei_digest_A_full_01.1943.1945.3 | 2314.7 (+0.9) | 5.4579 | 0.426 | 1581.3 | 1 | 34/ 76 | sp P02754 LACB_BOVIN | R.VYVEELKFTPEGDLLEILLQK.W |
| 1209 | ./sergei_digest_A_full_01.2765.2771.2 | 2709.1 (+0.4) | 5.3678 | 0.495 | 1654.1 | 1 | 24/ 50 | sp P02754 LACB_BOVIN | K.VAGTWSLAMAASDLSLLDAQSAPLR.V |
| 1220 | ./sergei_digest_A_full_01.2621.2623.2 | 2107.6 (-0.5) | 5.3391 | 0.461 | 1646.5 | 1 | 22/ 36 | sp P02666 CASH_BOVIN | R.DHFIQAFLLYGEFVLGPPV.G |
| 1153 | ./sergei_digest_A_full_01.2491.2493.3 | 2219.5 (+2.1) | 5.3167 | 0.276 | 1640.4 | 1 | 31/ 72 | sp P00921 CAH2_BOVIN | R.TLNFNAEGEPELLNLANNR.P |
| 1102 | ./sergei_digest_A_full_01.2401.2403.2 | 2254.5 (-0.6) | 5.1675 | 0.495 | 1009.0 | 1 | 24/ 42 | sp P00921 CAH2_BOVIN | K.YGDFCTAAQQPDGLAVVGVFLK.V |
| 1067 | ./sergei_digest_A_full_01.2329.2332.2 | 2254.5 (-0.9) | 5.1492 | 0.546 | 600.2 | 1 | 23/ 42 | sp P00921 CAH2_BOVIN | K.YGDFCTAAQQPDGLAVVGVFLK.V |
| 125 | ./sergei_digest_A_full_01.0681.0681.2 | 2100.2 (+0.8) | 4.9146 | 0.481 | 1779.2 | 1 | 22/ 34 | sp P00921 CAH2_BOVIN | R.MVNHGHSFNVEYDDSDQK.A |
| 1020 | ./sergei_digest_A_full_01.2237.2239.2 | 1568.7 (-0.1) | 4.8921 | 0.413 | 1764.0 | 1 | 19/ 24 | sp P02769 ALBU_BOVIN | K.DAFLGSLFYEYSR.R |
| 981 | ./sergei_digest_A_full_01.2163.2165.2 | 2147.3 (-0.7) | 4.8738 | 0.452 | 1515.4 | 1 | 26/ 38 | sp P02666 CASH_BOVIN | E.LNVFGEIVSLSSESESIIR.I |
| 533 | ./sergei_digest_A_full_01.1361.1367.3 | 2906.0 (+0.7) | 4.8712 | 0.301 | 655.0 | 1 | 29/ 92 | sp P02666 CASH_BOVIN | K.FQSEKQOQTEDELDKIMFFAQTL.S |
| 015 | ./sergei_digest_A_full_01.1079.1092.2 | 2314.7 (-0.6) | 4.7027 | 0.419 | 430.7 | 1 | 19/ 39 | sp P02754 LACB_BOVIN | R.VYVEELKFTPEGDLLEILLQK.W |
| 760 | ./sergei_digest_A_full_01.1771.1773.2 | 2034.2 (-0.1) | 4.7587 | 0.465 | 1047.9 | 1 | 26/ 36 | sp P02666 CASH_BOVIN | L.NVFGIVEVLSSESESIIR.I |
| 1048 | ./sergei_digest_A_full_01.2289.2291.2 | 2804.0 (+0.5) | 4.7564 | 0.433 | 1497.0 | 1 | 26/ 49 | sp P02666 CASH_BOVIN | A.DELEKLVNFGIVEVLSSESESIIR.I |
| 157 | ./sergei_digest_A_full_01.0735.0737.3 | 2100.2 (-0.5) | 4.7471 | 0.300 | 1496.2 | 1 | 32/ 68 | sp P00921 CAH2_BOVIN | R.MVNHGHSFNVEYDDSDQK.A |
| 91 | ./sergei_digest_A_full_01.0607.0609.2 | 1983.0 (+0.4) | 4.6602 | 0.358 | 3224.6 | 1 | 24/ 30 | sp P02666 CASH_BOVIN | K.FQSEKQOQTEDELDKIMFFAQTL.S |
| 625 | ./sergei_digest_A_full_01.1523.1525.3 | 2046.3 (-0.7) | 4.6539 | 0.307 | 1901.4 | 1 | 29/ 60 | sp P02769 ALBU_BOVIN | R.PHFPTPAPELLYIANK.Y |
| 117 | ./sergei_digest_A_full_01.0663.0665.3 | 2100.2 (+0.0) | 4.5901 | 0.286 | 1627.3 | 1 | 36/ 68 | sp P00921 CAH2_BOVIN | R.MVNHGHSFNVEYDDSDQK.A |
| 712 | ./sergei_digest_A_full_01.1675.1681.2 | 1832.0 (-0.3) | 4.5772 | 0.393 | 1148.8 | 1 | 21/ 34 | sp Q029443 TRFE_BOVIN | K.GEADAMSLDGGYLYIAGK.C |
| 651 | ./sergei_digest_A_full_01.1569.1571.3 | 3106.3 (+0.4) | 4.5737 | 0.362 | 537.5 | 1 | 26/100 | sp P02666 CASH_BOVIN | K.FQSEKQOQTEDELDKIMFFAQTL.S |
| 942 | ./sergei_digest_A_full_01.2085.2089.3 | 2314.7 (+1.4) | 4.5707 | 0.299 | 1348.2 | 1 | 32/ 76 | sp P02754 LACB_BOVIN | R.VYVEELKFTPEGDLLEILLQK.W |
| 1045 | ./sergei_digest_A_full_01.2277.2279.3 | 2804.0 (-0.5) | 4.5012 | 0.352 | 1236.7 | 1 | 31/ 96 | sp P02666 CASH_BOVIN | A.DELEKLVNFGIVEVLSSESESIIR.I |
| 861 | ./sergei_digest_A_full_01.1951.1953.2 | 2314.7 (-0.4) | 4.4874 | 0.348 | 482.4 | 1 | 20/ 38 | sp P02754 LACB_BOVIN | R.VYVEELKFTPEGDLLEILLQK.W |
| 626 | ./sergei_digest_A_full_01.1529.1531.2 | 1480.7 (-0.5) | 4.4192 | 0.315 | 1731.6 | 1 | 20/ 24 | sp P02769 ALBU_BOVIN | K.LGEYFQNALIVR.Y |
| 176 | ./sergei_digest_A_full_01.0767.0767.2 | 1983.0 (+0.7) | 4.4109 | 0.414 | 1966.2 | 1 | 22/ 30 | sp P02666 CASH_BOVIN | K.FQSEKQOQTEDELDKIMFFAQTL.S |
| 320 | ./sergei_digest_A_full_01.1009.1011.3 | 1440.7 (+0.0) | 4.4102 | 0.302 | 2269.3 | 1 | 29/ 44 | sp P02769 ALBU_BOVIN | R.PHFPTPAPELLYIANK.Y |

sort by match score

spectrum scores protein peptide



Set some threshold

Next, a threshold value was set. Different programs have different scoring schemes, so SEQUEST, Mascot, and X!Tandem use different thresholds.

Different thresholds may also be needed for different charge states, sample complexity, and database size.

HTML-SUMMARY v.8 (rev 0), Copyright 1996 - Mozilla Firefox

file:///Snapserver/proteomecof/users/brian/sergei_digest_A_full_01.html

PSWebMail Gmail - Inbox Slashdot Google News PSWebPage JDK 1.4 API

HTML-SUMMARY v.8 (rev 0), Copyright 1996

Molecular Biotechnology, Univ. of Washington, J.Eng/J.Yates
Compiled for use by the Abersold lab @ Univ. of Washington
03/19/02, 07:20 AM, /data/search/akeller/databases/new_hum_plus_proteinai.x.db, AVG/AVG

| # | File | MH+ | XCorr | dCn | Sp | RSp | Ions | Ref | Sequence |
|------|---------------------------------------|---------------|--------|-------|--------|-----|--------|-----------------------|---------------------------------|
| 325 | ./sergei_digest_A_full_01.1001.1003.3 | 2511.7 (-0.6) | 7.5063 | 0.450 | 4546.1 | 1 | 42/ 84 | sp P00921 CAH2_BOVIN | R.MVNHGHSFNVEYDDSDQKAVLK.D |
| 462 | ./sergei_digest_A_full_01.1239.1241.3 | 2504.7 (+2.0) | 6.1800 | 0.302 | 3126.2 | 1 | 39/ 84 | sp P00921 CAH2_BOVIN | R.LVGFHFVGGSSBBGGSEITVDR.K |
| 1070 | ./sergei_digest_A_full_01.2335.2337.3 | 2254.5 (+0.5) | 6.0682 | 0.491 | 2166.9 | 1 | 37/ 84 | sp P00921 CAH2_BOVIN | K.YGDFCTAAQQPDGLAVVGVFLK.V |
| 1105 | ./sergei_digest_A_full_01.2405.2407.3 | 2254.5 (+0.8) | 6.0041 | 0.511 | 1873.6 | 1 | 35/ 84 | sp P00921 CAH2_BOVIN | K.YGDFCTAAQQPDGLAVVGVFLK.V |
| 510 | ./sergei_digest_A_full_01.1317.1325.3 | 2584.7 (+1.5) | 5.9521 | 0.403 | 2488.3 | 1 | 38/ 84 | sp P00921 CAH2_BOVIN | R.LVGFHFVGGSSBBGGSEITVDR.K |
| 1219 | ./sergei_digest_A_full_01.2617.2619.3 | 2107.6 (-0.1) | 5.7343 | 0.502 | 2202.1 | 1 | 33/ 72 | sp P02666 CASB_BOVIN | R.DHFIQAFLLYGEFVLGPPV.G |
| 894 | ./sergei_digest_A_full_01.2013.2015.3 | 2314.7 (+0.0) | 5.5636 | 0.410 | 1260.3 | 1 | 33/ 76 | sp P02754 LACB_BOVIN | R.VVYEELKFTPEGDLEILLQK.W |
| 812 | ./sergei_digest_A_full_01.1873.1875.3 | 2314.7 (+0.5) | 5.5466 | 0.420 | 1407.6 | 1 | 35/ 76 | sp P02754 LACB_BOVIN | R.VVYEELKFTPEGDLEILLQK.W |
| 1142 | ./sergei_digest_A_full_01.2471.2473.2 | 2254.5 (-0.8) | 5.5372 | 0.526 | 771.0 | 1 | 24/ 42 | sp P00921 CAH2_BOVIN | K.YGDFCTAAQQPDGLAVVGVFLK.V |
| 836 | ./sergei_digest_A_full_01.1943.1945.3 | 2314.7 (+0.9) | 5.4579 | 0.426 | 1581.3 | 1 | 34/ 76 | sp P02754 LACB_BOVIN | R.VVYEELKFTPEGDLEILLQK.W |
| 1209 | ./sergei_digest_A_full_01.2765.2771.2 | 2709.1 (+0.4) | 5.3670 | 0.495 | 1654.1 | 1 | 24/ 50 | sp P02754 LACB_BOVIN | K.VAGTMYSLAMAASDLSLLDAQSAFLR.V |
| 1220 | ./sergei_digest_A_full_01.2621.2623.2 | 2107.6 (-0.5) | 5.3391 | 0.461 | 1646.5 | 1 | 22/ 36 | sp P02666 CASB_BOVIN | R.DHFIQAFLLYGEFVLGPPV.G |
| 1153 | ./sergei_digest_A_full_01.2491.2493.3 | 2219.5 (+2.1) | 5.3167 | 0.276 | 1640.4 | 1 | 31/ 72 | sp P00921 CAH2_BOVIN | R.TLMEAEQPELILMLANRR.P |
| 1102 | ./sergei_digest_A_full_01.2401.2403.2 | 2254.5 (-0.6) | 5.1675 | 0.495 | 1009.0 | 1 | 24/ 42 | sp P00921 CAH2_BOVIN | K.YGDFCTAAQQPDGLAVVGVFLK.V |
| 1067 | ./sergei_digest_A_full_01.2329.2332.2 | 2254.5 (-0.9) | 5.1492 | 0.546 | 600.2 | 1 | 23/ 42 | sp P00921 CAH2_BOVIN | K.YGDFCTAAQQPDGLAVVGVFLK.V |
| 125 | ./sergei_digest_A_full_01.0681.0681.2 | 2100.2 (+0.8) | 4.9146 | 0.481 | 1779.2 | 1 | 22/ 34 | sp P00921 CAH2_BOVIN | R.MVNHGHSFNVEYDDSDQK.A |
| 1020 | ./sergei_digest_A_full_01.2237.2239.2 | 1568.7 (-0.1) | 4.8921 | 0.413 | 1764.0 | 1 | 19/ 24 | sp P02769 ALBU_BOVIN | K.DAFLGSLFYYSR.R |
| 981 | ./sergei_digest_A_full_01.2163.2165.2 | 2127.3 (-0.7) | 4.8738 | 0.457 | 1515.4 | 1 | 26/ 38 | sp P02666 CASB_BOVIN | R.LVGFHFVGGSSBBGGSEITVDR.K |
| 533 | ./sergei_digest_A_full_01.1361.1367.3 | 2906.0 (+0.7) | 4.8712 | 0.301 | 855.0 | 1 | 29/ 92 | sp P02666 CASB_BOVIN | K.FQSEKQQQTEDELQDKTHPFAQTSL.V |
| 015 | ./sergei_digest_A_full_01.1079.1092.3 | 2314.7 (-0.6) | 4.7027 | 0.419 | 430.7 | 1 | 19/ 39 | sp P02754 LACB_BOVIN | R.VVYEELKFTPEGDLEILLQK.W |
| 760 | ./sergei_digest_A_full_01.1771.1773.2 | 2034.2 (-0.1) | 4.7587 | 0.465 | 1047.9 | 1 | 26/ 36 | sp P02666 CASB_BOVIN | L.NVFGIVEVLSLSSSEISITR.I |
| 1048 | ./sergei_digest_A_full_01.2289.2291.2 | 2804.0 (+0.5) | 4.7564 | 0.433 | 1497.0 | 1 | 26/ 49 | sp P02666 CASB_BOVIN | A.DLEELKLVVFGIVEVLSLSSSEISITR.I |
| 157 | ./sergei_digest_A_full_01.0735.0737.3 | 2100.2 (-0.5) | 4.7471 | 0.300 | 1496.2 | 1 | 32/ 68 | sp P00921 CAH2_BOVIN | R.MVNHGHSFNVEYDDSDQK.A |
| 91 | ./sergei_digest_A_full_01.0607.0609.2 | 1983.0 (+0.4) | 4.6602 | 0.358 | 3224.6 | 1 | 24/ 30 | sp P02666 CASB_BOVIN | K.FQSEKQQQTEDELQDK.I |
| 625 | ./sergei_digest_A_full_01.1523.1525.3 | 2046.3 (-0.7) | 4.6539 | 0.307 | 1901.4 | 1 | 29/ 60 | sp P02769 ALBU_BOVIN | R.PHFPTPAPELLYIANK.Y |
| 117 | ./sergei_digest_A_full_01.0663.0665.3 | 2100.2 (+0.0) | 4.5901 | 0.286 | 1627.3 | 1 | 36/ 68 | sp P00921 CAH2_BOVIN | R.MVNHGHSFNVEYDDSDQK.A |
| 712 | ./sergei_digest_A_full_01.1675.1681.2 | 1832.0 (-0.3) | 4.5772 | 0.393 | 1148.8 | 1 | 21/ 34 | sp Q029443 TRFE_BOVIN | K.GEADAMSLDGGYVLIAGK.C |
| 651 | ./sergei_digest_A_full_01.1569.1571.3 | 3106.3 (+0.4) | 4.5737 | 0.362 | 537.5 | 1 | 26/100 | sp P02666 CASB_BOVIN | K.FQSEKQQQTEDELQDKTHPFAQTSL.V |
| 942 | ./sergei_digest_A_full_01.2085.2089.3 | 2314.7 (+1.4) | 4.5707 | 0.299 | 1348.2 | 1 | 32/ 76 | sp P02754 LACB_BOVIN | R.VVYEELKFTPEGDLEILLQK.W |
| 1045 | ./sergei_digest_A_full_01.2277.2279.3 | 2804.0 (-0.5) | 4.5012 | 0.352 | 1236.7 | 1 | 31/ 96 | sp P02666 CASB_BOVIN | A.DLEELKLVVFGIVEVLSLSSSEISITR.I |
| 861 | ./sergei_digest_A_full_01.1951.1953.2 | 2314.7 (-0.4) | 4.4874 | 0.348 | 482.4 | 1 | 20/ 38 | sp P02754 LACB_BOVIN | R.VVYEELKFTPEGDLEILLQK.W |
| 626 | ./sergei_digest_A_full_01.1529.1531.2 | 1480.7 (-0.5) | 4.4192 | 0.315 | 1731.6 | 1 | 20/ 24 | sp P02769 ALBU_BOVIN | K.LGQYFQNALIVR.V |
| 176 | ./sergei_digest_A_full_01.0767.0767.2 | 1983.0 (+0.7) | 4.4109 | 0.414 | 1966.2 | 1 | 22/ 30 | sp P02666 CASB_BOVIN | K.FQSEKQQQTEDELQDK.I |
| 320 | ./sergei_digest_A_full_01.1009.1011.3 | 1440.7 (+0.0) | 4.4102 | 0.302 | 2269.3 | 1 | 29/ 44 | sp P02769 ALBU_BOVIN | R.PHFPTPAPELLYIANK.Y |

sort by match score

SEQUEST
XCorr > 2.5
dCn > 0.1

Mascot
Score > 45

X!Tandem
Score < 0.01

spectrum scores protein peptide



Below threshold matches dropped

Peptides that are identified with scores above the threshold are considered "correct" matches. Those with scores below the threshold are considered "incorrect".

There is no gray area where something is possibly correct.

HTML-SUMMARY v.8 (rev 0), Copyright 1996 - Mozilla Firefox

file:///Snapserver/proteomecot/users/brian/sergei_digest_A_full_01.html

PSWebMail Gmail - Inbox Slashdot Google News PSWebPage JXK 1.4 API

HTML-SUMMARY v.8 (rev 0), Copyrig...

HTML-SUMMARY v.8 (rev 0), Copyright 1996
Molecular Biotechnology, Univ. of Washington, J.Eng/J.Yates
Compiled for use by the Abersold lab @ Univ. of Washington
03/19/02, 07:20 AM, /data/search/akeller/databases/new_hum_plus_proteinaix.db, AVG/AVG

| # | File | MH+ | XCorr | dCn | Sp | RSp | Ions | Ref | Sequence |
|------|---------------------------------------|---------------|--------|-------|--------|-----|--------|-----------------------|--------------------------------|
| 325 | ./sergei_digest_A_full_01.1001.1003.3 | 2511.7 (-0.6) | 7.5063 | 0.450 | 4546.1 | 1 | 42/ 84 | sp P00921 CAH2_BOVIN | R.MVNHGHSFNVEYDDSDQKAVLK.D |
| 462 | ./sergei_digest_A_full_01.1239.1241.3 | 2504.7 (+2.0) | 6.1808 | 0.302 | 3126.2 | 1 | 39/ 84 | sp P00921 CAH2_BOVIN | R.LVQFHFHWGSSBBOGSEHTVDR.K |
| 1070 | ./sergei_digest_A_full_01.2335.2337.3 | 2254.5 (+0.5) | 6.0682 | 0.491 | 2166.9 | 1 | 37/ 84 | sp P00921 CAH2_BOVIN | K.YGDFCTAAQQPDGLAVVGVFLK.V |
| 1105 | ./sergei_digest_A_full_01.2405.2407.3 | 2254.5 (+0.8) | 6.0041 | 0.511 | 1873.6 | 1 | 35/ 84 | sp P00921 CAH2_BOVIN | K.YGDFCTAAQQPDGLAVVGVFLK.V |
| 510 | ./sergei_digest_A_full_01.1317.1325.3 | 2584.7 (+1.5) | 5.9521 | 0.403 | 2488.3 | 1 | 38/ 84 | sp P00921 CAH2_BOVIN | R.LVQFHFHWGSSBBOGSEHTVDR.K |
| 1219 | ./sergei_digest_A_full_01.2617.2619.3 | 2107.6 (-0.1) | 5.7343 | 0.502 | 2202.1 | 1 | 33/ 72 | sp P02666 CASB_BOVIN | R.DHFIQAFLLYGFVLPVPR.G |
| 894 | ./sergei_digest_A_full_01.2013.2015.3 | 2314.7 (+0.0) | 5.5636 | 0.410 | 1260.3 | 1 | 33/ 76 | sp P02754 LACB_BOVIN | R.VVYEKLFKTFEGDLEILLQK.W |
| 812 | ./sergei_digest_A_full_01.1873.1875.3 | 2314.7 (+0.5) | 5.5466 | 0.428 | 1407.6 | 1 | 35/ 76 | sp P02754 LACB_BOVIN | R.VVYEKLFKTFEGDLEILLQK.W |
| 1142 | ./sergei_digest_A_full_01.2471.2475.3 | 2254.5 (-0.8) | 5.5372 | 0.526 | 771.0 | 1 | 24/ 42 | sp P00921 CAH2_BOVIN | K.YGDFCTAAQQPDGLAVVGVFLK.V |
| 836 | ./sergei_digest_A_full_01.1943.1945.3 | 2314.7 (+0.9) | 5.4579 | 0.426 | 1581.3 | 1 | 34/ 76 | sp P02754 LACB_BOVIN | R.VVYEKLFKTFEGDLEILLQK.W |
| 1209 | ./sergei_digest_A_full_01.2765.2771.2 | 2709.2 (+0.0) | 5.1492 | 0.546 | 600.2 | 1 | 23/ 42 | sp P02754 LACB_BOVIN | K.VAGTMYSLAMAASDLSLLDAQSAFLR.V |
| 1220 | ./sergei_digest_A_full_01.2621.2623.2 | 2107.6 (-0.1) | 4.9146 | 0.481 | 1779.2 | 1 | 22/ 34 | sp P00921 CAH2_BOVIN | R.MVNHGHSFNVEYDDSDQK.A |
| 1153 | ./sergei_digest_A_full_01.2491.2493.3 | 2219.5 (+2.0) | 4.8738 | 0.457 | 1515.4 | 1 | 26/ 38 | sp P02666 CASB_BOVIN | R.DHFIQAFLLYGFVLPVPR.G |
| 1102 | ./sergei_digest_A_full_01.2401.2403.2 | 2254.5 (-0.6) | 5.1675 | 0.495 | 1009.0 | 1 | 24/ 42 | sp P00921 CAH2_BOVIN | K.YGDFCTAAQQPDGLAVVGVFLK.V |
| 1067 | ./sergei_digest_A_full_01.2329.2332.2 | 2254.5 (-0.9) | 5.1492 | 0.546 | 600.2 | 1 | 23/ 42 | sp P02754 LACB_BOVIN | K.YGDFCTAAQQPDGLAVVGVFLK.V |
| 125 | ./sergei_digest_A_full_01.0681.0681.2 | 2100.2 (+0.8) | 4.9146 | 0.481 | 1779.2 | 1 | 22/ 34 | sp P00921 CAH2_BOVIN | R.MVNHGHSFNVEYDDSDQK.A |
| 1020 | ./sergei_digest_A_full_01.2237.2239.2 | 1568.7 (-0.1) | 4.8921 | 0.413 | 1764.0 | 1 | 19/ 24 | sp P02769 ALBU_BOVIN | K.DAFLGSLFYYSR.R |
| 981 | ./sergei_digest_A_full_01.2163.2165.2 | 2127.3 (-0.7) | 4.8738 | 0.457 | 1515.4 | 1 | 26/ 38 | sp P02666 CASB_BOVIN | R.DHFIQAFLLYGFVLPVPR.G |
| 533 | ./sergei_digest_A_full_01.1361.1367.3 | 2906.0 (+0.7) | 4.8912 | 0.501 | 855.0 | 1 | 29/ 92 | sp P02666 CASB_BOVIN | K.FQSEKQQTTEDELDKTHPFAQTSL.V |
| 015 | ./sergei_digest_A_full_01.1079.1092.2 | 2314.7 (-0.6) | 4.7027 | 0.419 | 430.7 | 1 | 19/ 39 | sp P02754 LACB_BOVIN | R.VVYEKLFKTFEGDLEILLQK.W |
| 760 | ./sergei_digest_A_full_01.1771.1773.2 | 2034.2 (-0.1) | 4.7587 | 0.465 | 1047.9 | 1 | 26/ 36 | sp P02666 CASB_BOVIN | L.NVFGIVESLSSSESIITR.I |
| 1048 | ./sergei_digest_A_full_01.2389.2391.2 | 2804.0 (+0.5) | 4.7564 | 0.433 | 1497.0 | 1 | 26/ 49 | sp P02666 CASB_BOVIN | A.DFLEKLVVFGIVESLSSSESIITR.I |
| 157 | ./sergei_digest_A_full_01.0529.0533.3 | 2100.2 (+0.5) | 4.7471 | 0.500 | 1496.2 | 1 | 32/ 68 | sp P00921 CAH2_BOVIN | R.MVNHGHSFNVEYDDSDQK.A |
| 91 | ./sergei_digest_A_full_01.0607.0609.2 | 2100.2 (+0.4) | 4.6602 | 0.358 | 3224.6 | 1 | 24/ 30 | sp P02666 CASB_BOVIN | K.FQSEKQQTTEDELDKTHPFAQTSL.V |
| 625 | ./sergei_digest_A_full_01.1523.1525.3 | 2046.3 (+0.0) | 4.6539 | 0.307 | 1901.4 | 1 | 29/ 60 | sp P02769 ALBU_BOVIN | R.DHFIQAFLLYGFVLPVPR.G |
| 117 | ./sergei_digest_A_full_01.0663.0665.3 | 2100.2 (+0.0) | 4.7277 | 0.586 | 1627.3 | 1 | 26/100 | sp P00921 CAH2_BOVIN | R.MVNHGHSFNVEYDDSDQK.A |
| 712 | ./sergei_digest_A_full_01.1675.1681.2 | 1832.0 (-0.3) | 4.5772 | 0.500 | 1348.2 | 1 | 21/ 34 | sp I029443 TRFE_BOVIN | K.GEADAMSLDGGYLYIAGK.C |
| 651 | ./sergei_digest_A_full_01.1569.1571.3 | 3106.3 (+0.4) | 4.5737 | 0.499 | 537.5 | 1 | 26/100 | sp P02666 CASB_BOVIN | K.FQSEKQQTTEDELDKTHPFAQTSL.V |
| 942 | ./sergei_digest_A_full_01.2085.2089.3 | 2314.7 (+1.0) | 4.5707 | 0.299 | 1348.2 | 1 | 32/ 68 | sp P02754 LACB_BOVIN | R.VVYEKLFKTFEGDLEILLQK.W |
| 1045 | ./sergei_digest_A_full_01.2277.2279.3 | 2100.2 (+0.5) | 4.5012 | 0.352 | 1236.7 | 1 | 31/ 96 | sp P02666 CASB_BOVIN | A.DFLEKLVVFGIVESLSSSESIITR.I |
| 861 | ./sergei_digest_A_full_01.1057.1059.2 | 2314.7 (+0.4) | 4.4874 | 0.348 | 482.4 | 1 | 20/ 38 | sp P02754 LACB_BOVIN | R.VVYEKLFKTFEGDLEILLQK.W |
| 626 | ./sergei_digest_A_full_01.1529.1531.2 | 1440.7 (-0.1) | 4.4874 | 0.348 | 482.4 | 1 | 20/ 38 | sp P02754 LACB_BOVIN | R.VVYEKLFKTFEGDLEILLQK.W |
| 176 | ./sergei_digest_A_full_01.0767.0767.2 | 1983.0 (-0.1) | 4.4874 | 0.348 | 482.4 | 1 | 20/ 38 | sp P02666 CASB_BOVIN | K.FQSEKQQTTEDELDKTHPFAQTSL.V |
| 320 | ./sergei_digest_A_full_01.1009.1011.3 | 1440.7 (-0.1) | 4.4874 | 0.348 | 482.4 | 1 | 20/ 38 | sp P02769 ALBU_BOVIN | R.DHFIQAFLLYGFVLPVPR.G |

sort by match score

"correct"

"incorrect"

- SEQUEST
- XCorr > 2.5
- dCn > 0.1
- Mascot
- Score > 45
- X!Tandem
- Score < 0.01

spectrum scores protein peptide

There has to be a better way

The threshold model has these problems, which PeptideProphet tries to solve:

- Poor sensitivity/specificity trade-off, unless you consider multiple scores simultaneously.
- No way to choose an error rate ($p=0.05$).
- Need to have different thresholds for:
 - different instruments (QTOF, TOF-TOF, IonTrap)
 - ionization sources (electrospray vs MALDI)
 - sample complexities (2D gel spot vs MudPIT)
 - different databases (SwissProt vs NR)
- Impossible to compare results from different search algorithms, multiple instruments, and so on.



Creating a discriminant score

PeptideProphet starts with a discriminant score. If an application uses several scores, (SEQUEST uses Xcorr, ΔCn , and Sp scores; Mascot uses ion scores plus identity and homology thresholds), these are first converted to a single discriminant score.

HTML-SUMMARY v.8 (rev 0), Copyright 1996
Molecular Biotechnology, Univ. of Washington, J.Eng/J.Yates
Compiled for use by the Abersold lab @ Univ. of Washington
03/19/02, 07:20 AM, /data/search/akeller/databases/new_hum_plus_proteinax.db, AVG/AVG

| # | File | MH+ | Xcorr | dCn | Sp | RPp | Ions | Ref | Sequence |
|------|---------------------------------------|--------|--------|--------|-------|--------|--------|----------------------|--------------------------------|
| 325 | ./sergei_digest_A_full_01.1001.1003.3 | 2511.7 | (-0.6) | 7.5063 | 0.450 | 4546.1 | 42/ 84 | sp P00921 CAH2_BOVIN | R.MVNHGHSFNVEYDDSDQKAVLK.D |
| 462 | ./sergei_digest_A_full_01.1239.1241.3 | 2504.7 | (+2.0) | 6.1808 | 0.302 | 3126.2 | 39/ 84 | sp P00921 CAH2_BOVIN | R.LVGFHFPIWGSBBGGSEHTVDR.K |
| 1070 | ./sergei_digest_A_full_01.2335.2337.3 | 2254.5 | (+0.5) | 6.0682 | 0.491 | 2166.9 | 37/ 84 | sp P00921 CAH2_BOVIN | K.YGDFCTAAQQPDGLAVVGVFLK.V |
| 1105 | ./sergei_digest_A_full_01.2405.2407.3 | 2254.5 | (+0.8) | 6.0041 | 0.511 | 1873.6 | 35/ 84 | sp P00921 CAH2_BOVIN | K.YGDFCTAAQQPDGLAVVGVFLK.V |
| 510 | ./sergei_digest_A_full_01.1317.1325.3 | 2584.7 | (+1.5) | 5.9521 | 0.403 | 2488.3 | 38/ 84 | sp P00921 CAH2_BOVIN | R.LVGFHFPIWGSBBGGSEHTVDR.K |
| 1219 | ./sergei_digest_A_full_01.2617.2619.3 | 2107.6 | (-0.1) | 5.7343 | 0.502 | 2292.1 | 33/ 72 | sp P02666 CASH_BOVIN | R.DHFIQAFLLYQEPVLGPPV.G |
| 894 | ./sergei_digest_A_full_01.2013.2015.3 | 2314.7 | (+0.0) | 5.5636 | 0.410 | 1260.3 | 33/ 76 | sp P02754 LACB_BOVIN | R.VYVEELKFTPEGDLEILLQK.W |
| 1142 | ./sergei_digest_A_full_01.1873.1875.3 | 2314.7 | (+0.5) | 5.5466 | 0.428 | 1407.6 | 35/ 76 | sp P02754 LACB_BOVIN | R.VYVEELKFTPEGDLEILLQK.W |
| 1142 | ./sergei_digest_A_full_01.2471.2475.2 | 2254.5 | (-0.8) | 5.5372 | 0.526 | 771.0 | 24/ 42 | sp P00921 CAH2_BOVIN | K.YGDFCTAAQQPDGLAVVGVFLK.V |
| 836 | ./sergei_digest_A_full_01.1943.1945.3 | 2314.7 | (+0.9) | 5.4579 | 0.426 | 1581.3 | 34/ 76 | sp P02754 LACB_BOVIN | R.VYVEELKFTPEGDLEILLQK.W |
| 1209 | ./sergei_digest_A_full_01.2765.2771.2 | 2709.1 | (+0.4) | 5.3678 | 0.495 | 1654.1 | 24/ 50 | sp P02754 LACB_BOVIN | K.VAGTMYSLAMAASDLSLLDAQSAFLR.V |
| 1220 | ./sergei_digest_A_full_01.2621.2623.2 | 2107.6 | (-0.5) | 5.3391 | 0.461 | 1646.5 | 22/ 36 | sp P02666 CASH_BOVIN | R.DHFIQAFLLYQEPVLGPPV.G |
| 1153 | ./sergei_digest_A_full_01.2491.2493.3 | 2219.5 | (+2.1) | 5.3167 | 0.276 | 1640.4 | 31/ 72 | sp P00921 CAH2_BOVIN | R.TLNFNAEGEPFLNLANLR.P |
| 1102 | ./sergei_digest_A_full_01.2401.2403.2 | 2254.5 | (-0.6) | 5.1675 | 0.495 | 1009.0 | 24/ 42 | sp P00921 CAH2_BOVIN | K.YGDFCTAAQQPDGLAVVGVFLK.V |
| 1067 | ./sergei_digest_A_full_01.2329.2332.2 | 2254.5 | (-0.9) | 5.1492 | 0.546 | 600.2 | 23/ 42 | sp P00921 CAH2_BOVIN | K.YGDFCTAAQQPDGLAVVGVFLK.V |
| 125 | ./sergei_digest_A_full_01.0681.0681.2 | 2100.2 | (+0.8) | 4.9146 | 0.481 | 1779.2 | 22/ 34 | sp P00921 CAH2_BOVIN | R.MVNHGHSFNVEYDDSDQK.A |
| 1020 | ./sergei_digest_A_full_01.2327.2329.2 | 1568.7 | (-0.1) | 4.8921 | 0.413 | 1764.0 | 19/ 24 | sp P02769 ALBU_BOVIN | K.DAFLGSLFYEYSR.R |
| 981 | ./sergei_digest_A_full_01.2163.2165.2 | 2147.3 | (-0.7) | 4.8738 | 0.452 | 1515.4 | 26/ 38 | sp P02666 CASH_BOVIN | E.LNVFGEIVSLSSEESITR.I |
| 533 | ./sergei_digest_A_full_01.1361.1367.3 | 2906.0 | (+0.7) | 4.8712 | 0.301 | 655.0 | 29/ 92 | sp P02666 CASH_BOVIN | K.FQSEKQQQTEDELDKIMFFAQTL.S |
| 015 | ./sergei_digest_A_full_01.1079.1092.2 | 2314.7 | (-0.6) | 4.7027 | 0.419 | 430.7 | 19/ 39 | sp P02754 LACB_BOVIN | R.VYVEELKFTPEGDLEILLQK.W |
| 760 | ./sergei_digest_A_full_01.1771.1773.2 | 2034.2 | (-0.1) | 4.7587 | 0.465 | 1047.9 | 26/ 36 | sp P02666 CASH_BOVIN | L.NVFGIVEVLSSEESITR.I |
| 1048 | ./sergei_digest_A_full_01.2289.2291.2 | 2804.0 | (+0.5) | 4.7564 | 0.433 | 1497.0 | 26/ 49 | sp P02666 CASH_BOVIN | A.DELEKLVNFGIVEVLSSEESITR.I |
| 157 | ./sergei_digest_A_full_01.0735.0737.3 | 2100.2 | (-0.5) | 4.7471 | 0.300 | 1496.2 | 32/ 68 | sp P00921 CAH2_BOVIN | R.MVNHGHSFNVEYDDSDQK.A |
| 91 | ./sergei_digest_A_full_01.0607.0609.2 | 1983.0 | (-0.4) | 4.6602 | 0.358 | 3224.6 | 24/ 30 | sp P02666 CASH_BOVIN | K.FQSEKQQQTEDELDKIMFFAQTL.S |
| 625 | ./sergei_digest_A_full_01.1523.1525.3 | 2046.3 | (-0.7) | 4.6539 | 0.307 | 1901.4 | 29/ 60 | sp P02769 ALBU_BOVIN | R.FHPTFYPELLTYANK.Y |
| 117 | ./sergei_digest_A_full_01.0663.0665.3 | 2100.2 | (+0.0) | 4.5901 | 0.286 | 1627.3 | 36/ 68 | sp P00921 CAH2_BOVIN | R.MVNHGHSFNVEYDDSDQK.A |
| 712 | ./sergei_digest_A_full_01.1675.1681.2 | 1832.0 | (-0.3) | 4.5772 | 0.393 | 1148.8 | 21/ 34 | sp Q29443 TRFE_BOVIN | K.GEADAMSLDGGYLYIAGK.C |
| 651 | ./sergei_digest_A_full_01.1569.1571.3 | 3106.3 | (+0.4) | 4.5737 | 0.362 | 537.5 | 26/100 | sp P02666 CASH_BOVIN | K.FQSEKQQQTEDELDKIMFFAQTL.S |
| 942 | ./sergei_digest_A_full_01.2085.2089.3 | 2314.7 | (+1.4) | 4.5707 | 0.299 | 1348.2 | 32/ 76 | sp P02754 LACB_BOVIN | R.VYVEELKFTPEGDLEILLQK.W |
| 1045 | ./sergei_digest_A_full_01.2277.2279.3 | 2804.0 | (-0.5) | 4.5012 | 0.352 | 1236.7 | 31/ 96 | sp P02666 CASH_BOVIN | A.DELEKLVNFGIVEVLSSEESITR.I |
| 861 | ./sergei_digest_A_full_01.1951.1953.2 | 2314.7 | (-0.4) | 4.4874 | 0.348 | 482.4 | 20/ 38 | sp P02754 LACB_BOVIN | R.VYVEELKFTPEGDLEILLQK.W |
| 626 | ./sergei_digest_A_full_01.1529.1531.2 | 1480.7 | (-0.5) | 4.4192 | 0.315 | 1731.6 | 20/ 24 | sp P02769 ALBU_BOVIN | K.LGEYFQNALIVR.Y |
| 176 | ./sergei_digest_A_full_01.0767.0767.2 | 1983.0 | (+0.7) | 4.4109 | 0.414 | 1966.2 | 22/ 30 | sp P02666 CASH_BOVIN | K.FQSEKQQQTEDELDKIMFFAQTL.S |
| 320 | ./sergei_digest_A_full_01.1009.1011.3 | 1440.7 | (+0.0) | 4.4102 | 0.302 | 2269.3 | 29/ 44 | sp P02769 ALBU_BOVIN | R.FHPTFYPELLTYANK.Y |

sort by match score

spectrum scores protein peptide

Discriminant score for SEQUEST

$$D = \left(\begin{array}{l} +8.4 * \frac{\ln(XCorr)}{\ln(\#AAs)} \\ +7.4 * \Delta Cn \\ -0.2 * \ln(rankSp) \\ -0.3 * \Delta Mass \\ -0.96 \end{array} \right)$$

For example, here's the formula to combine SEQUEST's scores into a discriminant score:

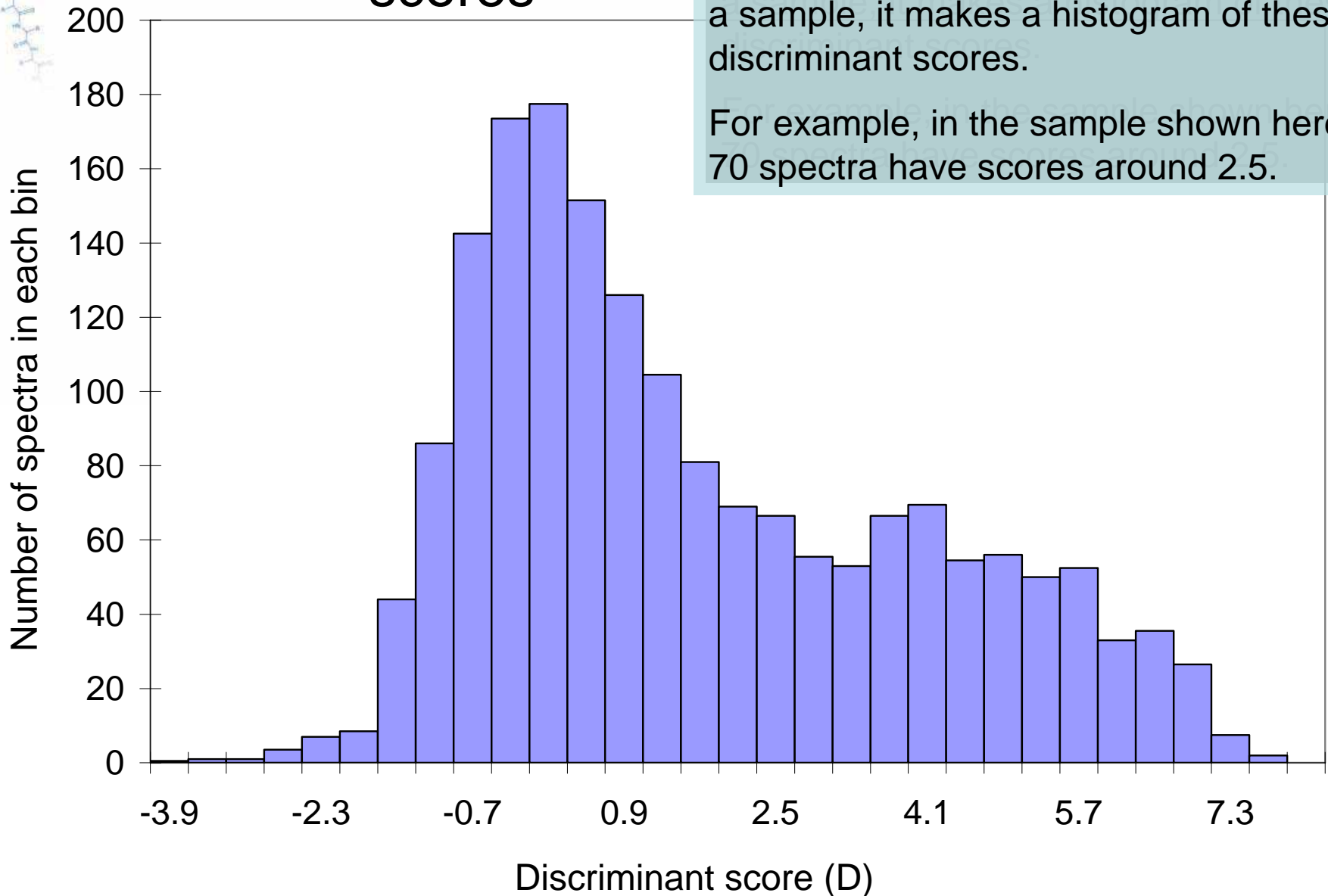
SEQUEST's **XCorr** (correlation score) is corrected for length of the peptide. High correlation is rewarded.

SEQUEST's **ΔCn** tells how far the top score is from the rest. Being far ahead of others is rewarded.

The top ranked by SEQUEST's **Sp** score has **ln(rankSp)=0**. Lower ranked scores are penalized.

Poor mass accuracy (big **ΔMass**) is also penalized.

Histogram of scores



Once Peptide Prophet calculates the discriminant scores for all the spectra in a sample, it makes a histogram of these discriminant scores.

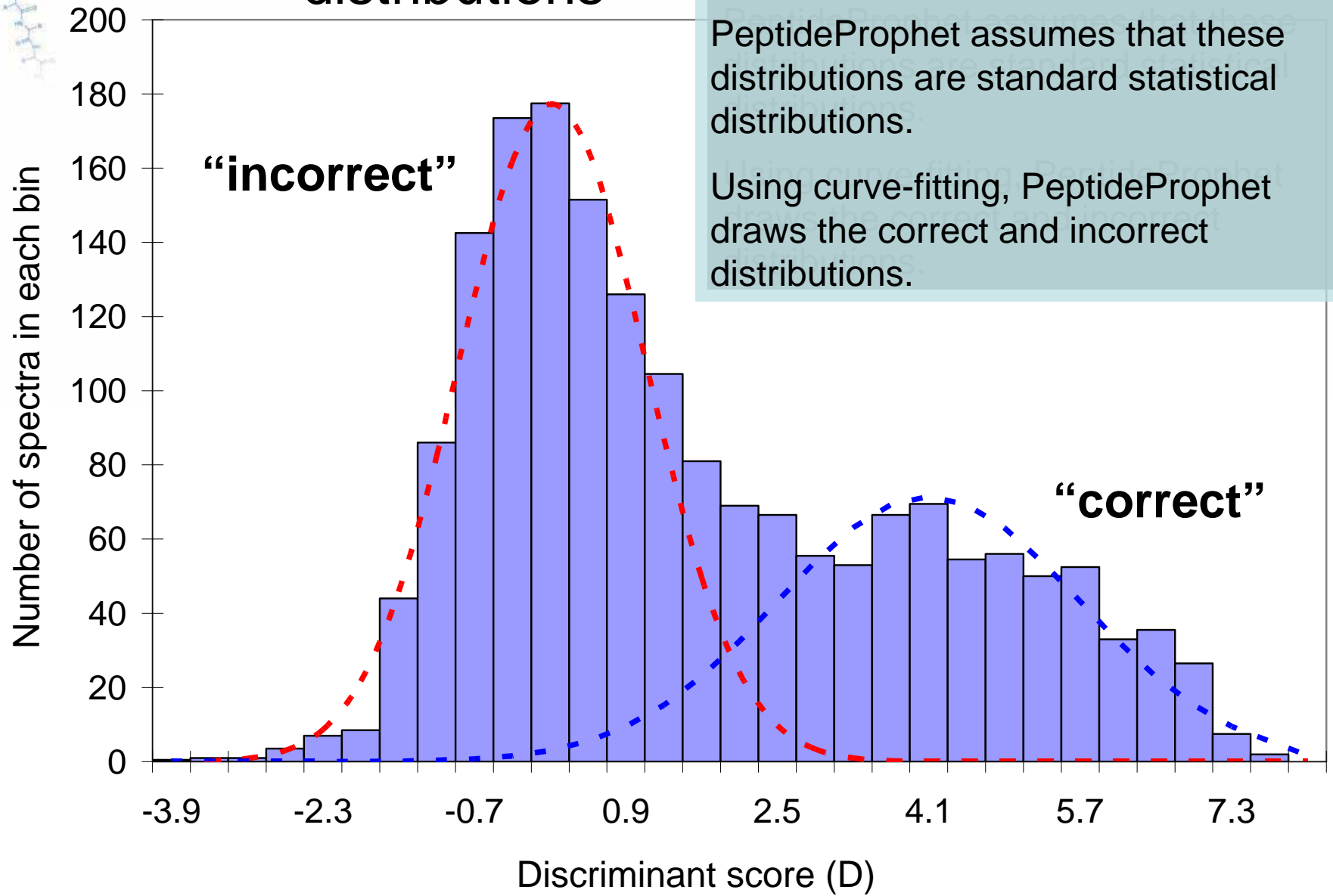
For example, in the sample shown here, 70 spectra have scores around 2.5.

Mixture of distributions

This histogram shows the distributions of correct and incorrect matches.

PeptideProphet assumes that these distributions are standard statistical distributions.

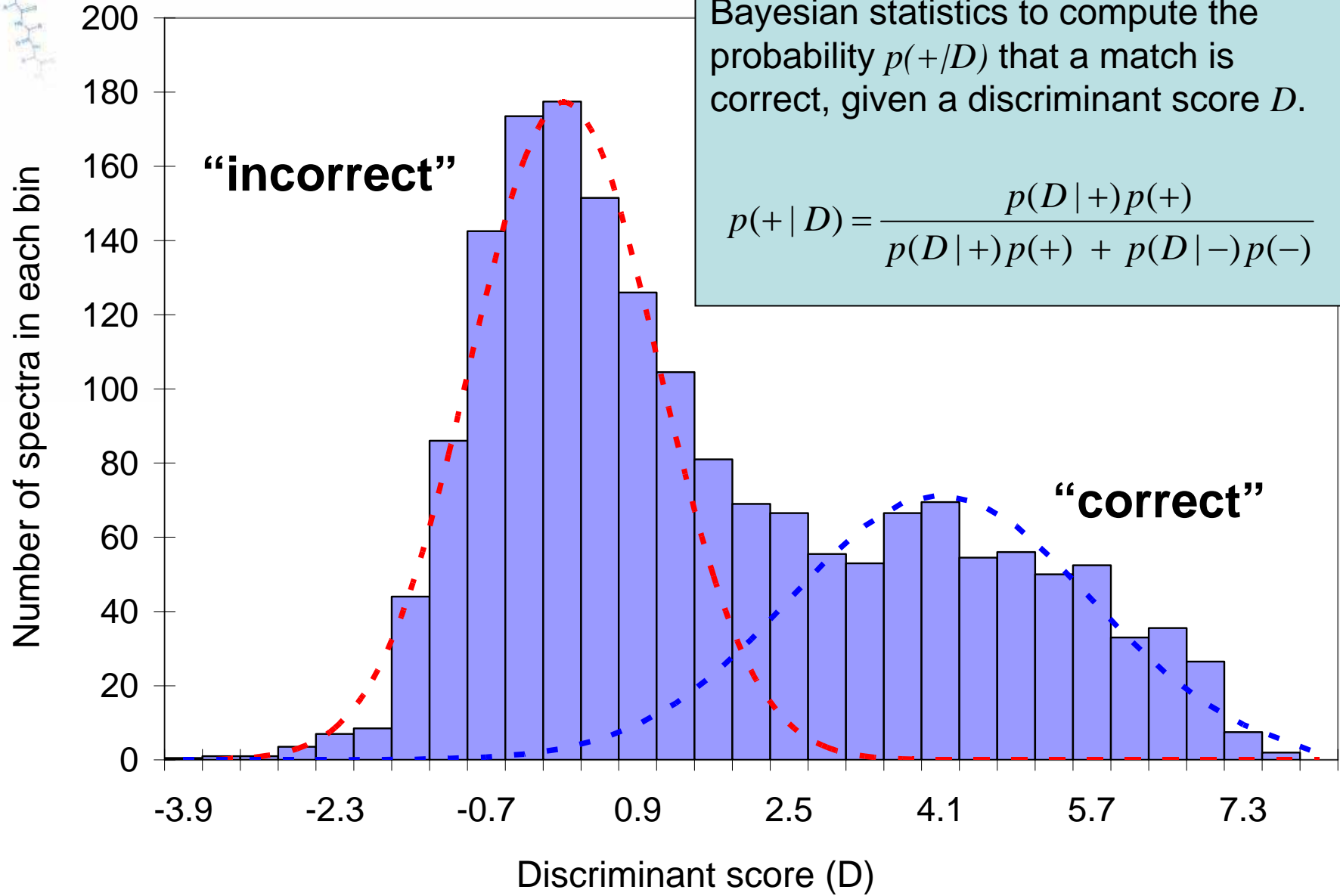
Using curve-fitting, PeptideProphet draws the correct and incorrect distributions.



Bayesian statistics

Once correct and incorrect distributions are drawn, PeptideProphet uses Bayesian statistics to compute the probability $p(+|D)$ that a match is correct, given a discriminant score D .

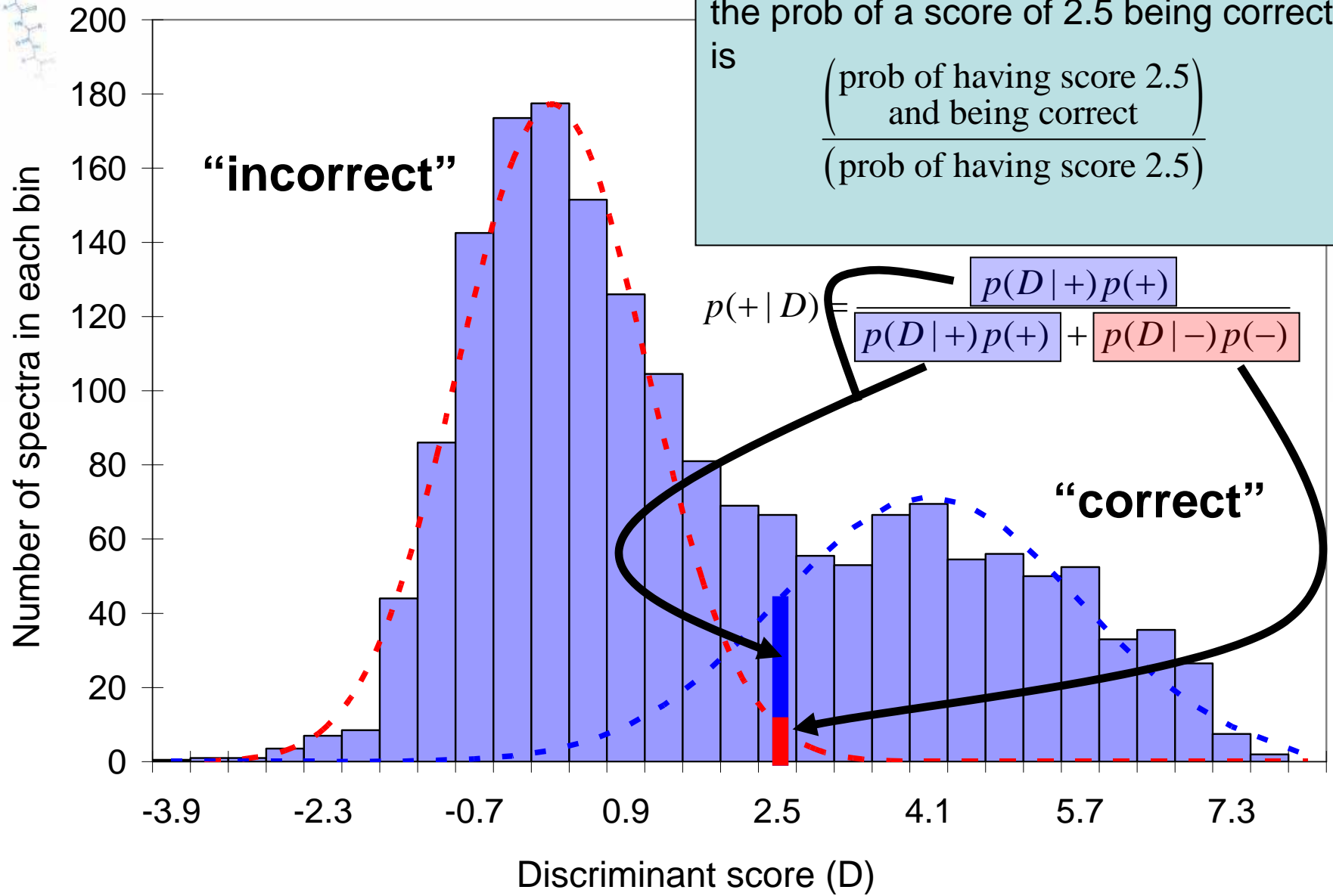
$$p(+|D) = \frac{p(D|+)p(+)}{p(D|+)p(+) + p(D|-)p(-)}$$



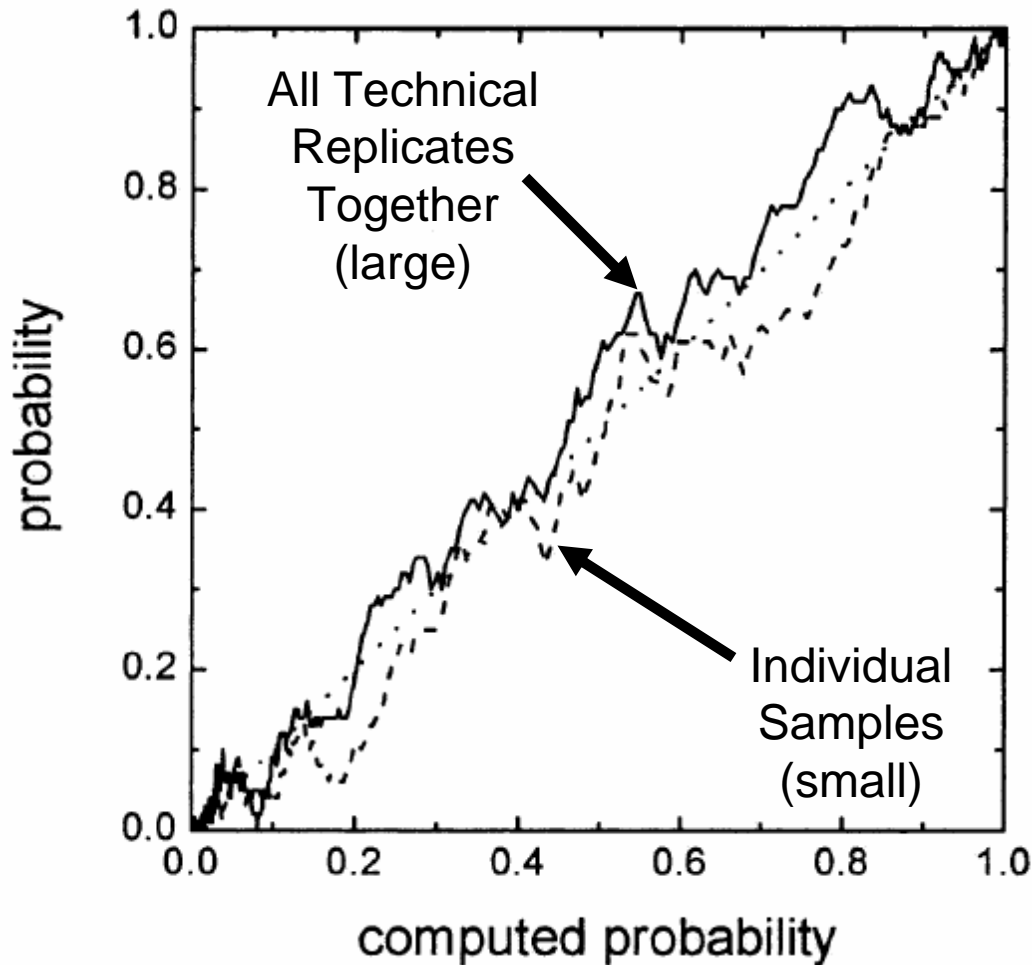
Probability of a correct match

The statistical formula looks fierce, but relating it to the histogram shows that the prob of a score of 2.5 being correct is

$$\frac{\text{(prob of having score 2.5 and being correct)}}{\text{(prob of having score 2.5)}}$$



PeptideProphet model is accurate



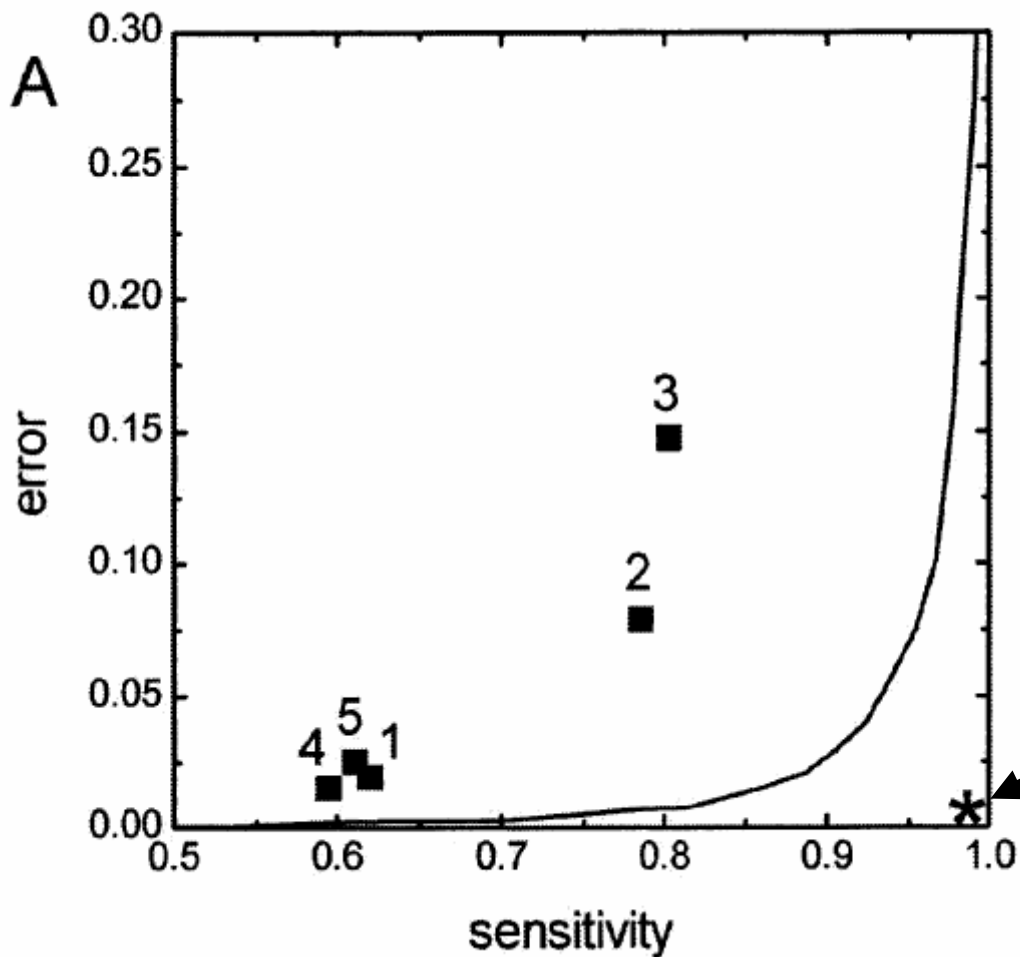
Keller, et al. checked PeptideProphet on a control data set for which they knew the right answer.

Ideally, the PeptideProphet-computed probability should be identical to the actual probability, corresponding to a 45-degree line on this graph.

They tested PeptideProphet with both large and small data sets and found pretty good agreement with the real probability.

Since it was published, the Institute for Systems Biology has used PeptideProphet on a number of protein samples of varying complexity.

PeptideProphet more sensitive than threshold model



This graph shows the trade-offs between the errors (false identifications) and the sensitivity (the percentage of possible peptides identified).

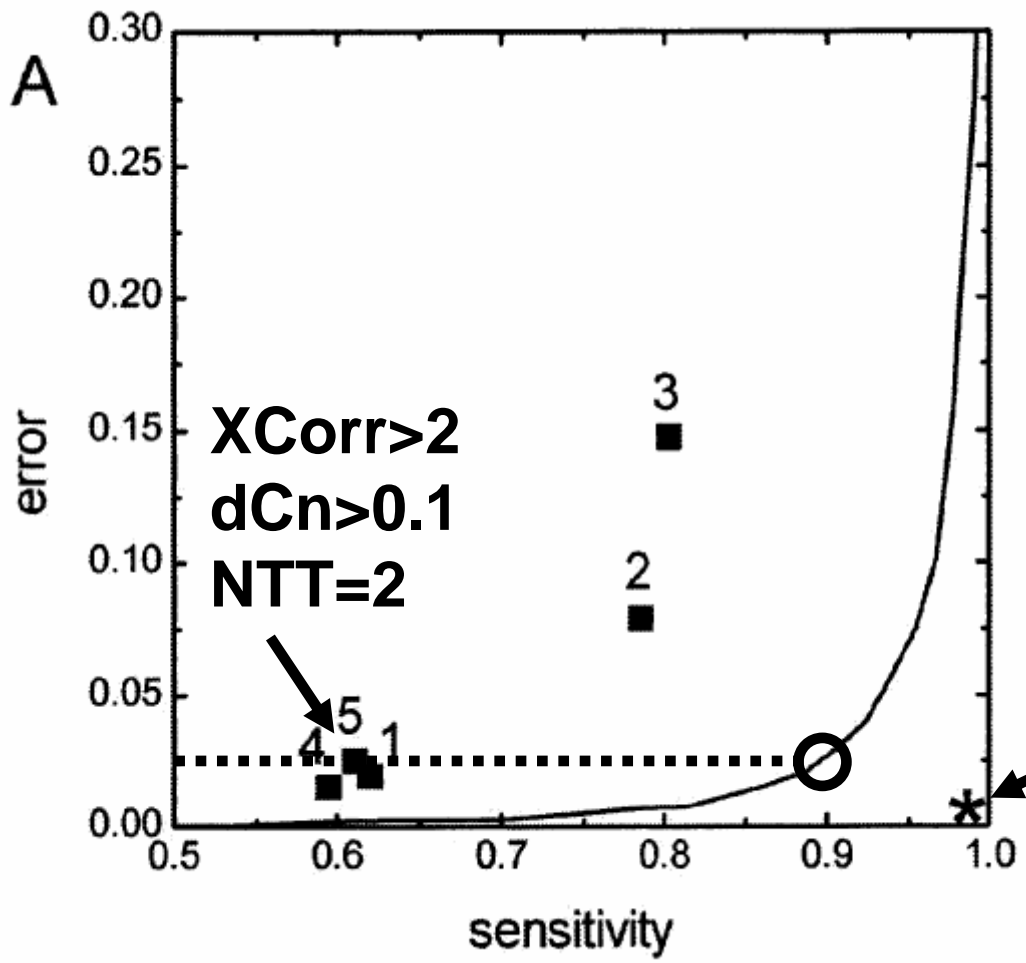
The ideal is zero error and everything identified (sensitivity = 100%).

PeptideProphet corresponds to the curved line. Squares 1–5 are thresholds chosen by other authors.

correctly identifies everything, with no error

PeptideProphet compared to Sequest

Xcorr cutoff of 2

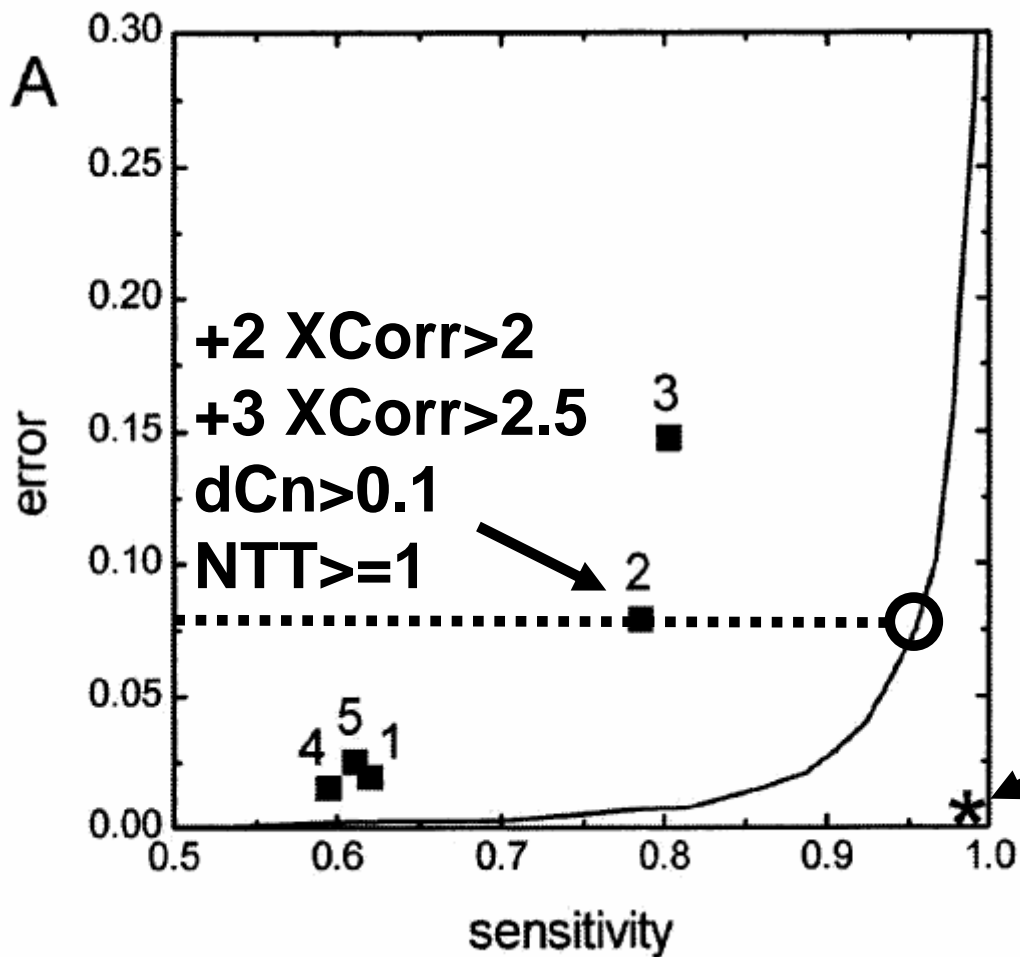


For example, for a threshold of $Xcorr > 2$ and $\Delta Cn > .1$ with only fully tryptic peptides allowed (see square 5 on the graph), Sequest's error rate is only 2%. However, its sensitivity is only 0.6 — that is, only 60% of the spectra are identified.

Using PeptideProphet, the same 2% error rate identifies 90% of the spectra, because the discriminant score is tuned to provide better results.

correctly identifies everything, with no error

Peptide Prophet compared to charge-dependent cutoff



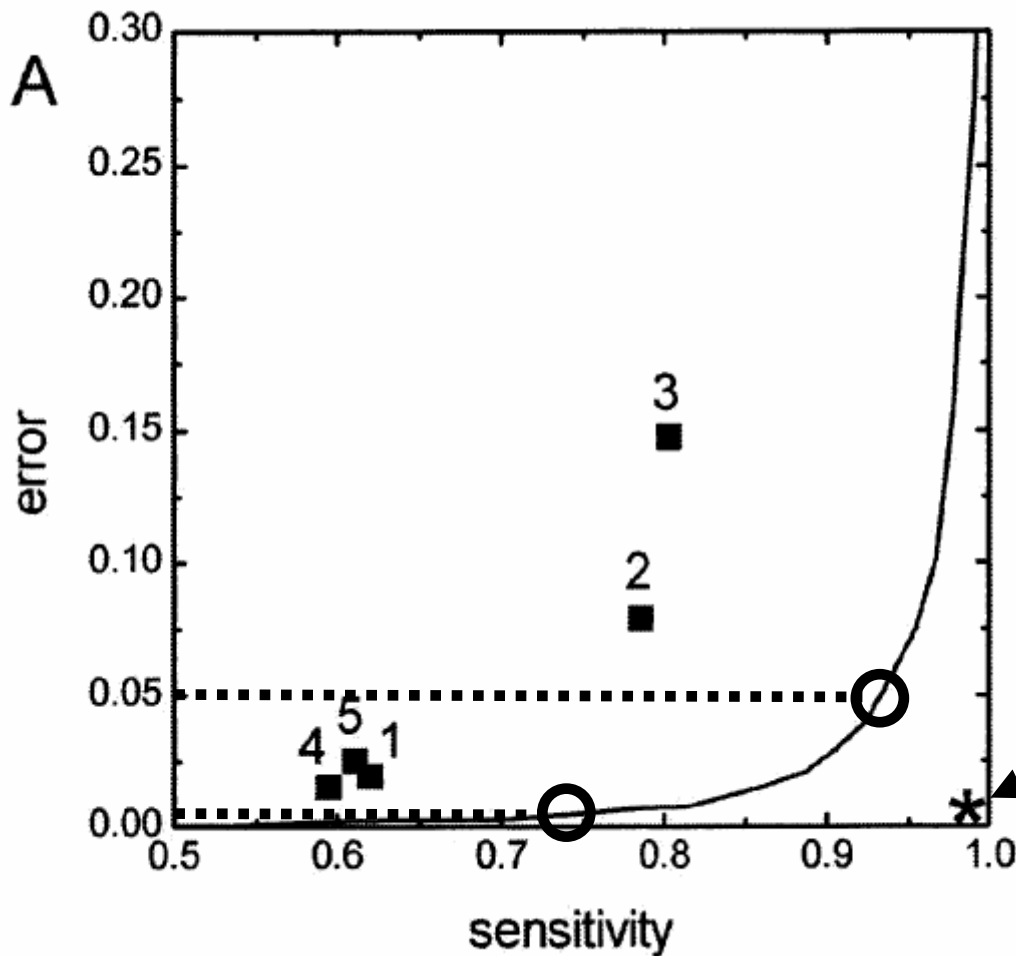
Another example uses a different threshold for charge +2 and charge +3 spectra (see square 2 on the graph). For this threshold, the error rate is 8% and the sensitivity is 80%.

At an error rate of 8%, PeptideProphet identifies 95% of the peptides.

correctly identifies everything, with no error

PeptideProphet allows you to choose an error rate

A big advantage is that you can choose any error rate you like, such as 5% for inclusive searches, or 1% for extremely accurate searches.



correctly identifies everything, with no error

There has to be a better way

Recall the problems that PeptideProphet was designed to fix. How well did it do?

- Poor sensitivity/specificity trade-off unless you consider multiple scores simultaneously.
- No way to choose an error rate ($p=0.05$).
- Need to have different thresholds for:
 - different instruments (QTOF, TOF-TOF, IonTrap)
 - ionization sources (electrospray vs MALDI)
 - sample complexities (2D gel spot vs MudPIT)
 - different databases (SwissProt vs NR)
- Impossible to compare results from different search algorithms, multiple instruments, and so on.



PeptideProphet better scores

The discriminant score combines the various scores into one optimal score.

- ~~• Poor sensitivity/specificity trade off unless you consider multiple scores simultaneously.~~
- **discriminant score**
- No way to choose an error rate ($p=0.05$).
- Need to have different thresholds for:
 - different instruments (QTOF, TOF-TOF, IonTrap)
 - ionization sources (electrospray vs MALDI)
 - sample complexities (2D gel spot vs MudPIT)
 - different databases (SwissProt vs NR)
- Impossible to compare results from different search algorithms, multiple instruments, and so on.



PeptideProphet better control of error rate

The error vs. sensitivity curves derived from the distributions allow you to choose the error rate.

- ~~• Poor sensitivity/specificity trade off unless you consider multiple scores simultaneously.~~ **discriminant score**
- ~~• No way to choose an error rate ($p=0.05$).~~ **estimate error with distributions**
- Need to have different thresholds for:
 - different instruments (QTOF, TOF-TOF, IonTrap)
 - ionization sources (electrospray vs MALDI)
 - sample complexities (2D gel spot vs MudPIT)
 - different databases (SwissProt vs NR)
- Impossible to compare results from different search algorithms, multiple instruments, and so on.



PeptideProphet

better adaptability

Each experiment has a different histogram of discriminant scores, to which the probability curves are automatically adapted.

- ~~• Poor sensitivity/specificity trade off unless you consider multiple scores simultaneously.~~ **discriminant score**
- ~~• No way to choose an error rate ($p=0.05$).~~ **estimate error with distributions**
- ~~• Need to have different thresholds for:~~ **curve-fit distributions to data (EM)**
 - ~~– different instruments (QTOF, TOF-TOF, IonTrap)~~
 - ~~– ionization sources (electrospray vs MALDI)~~
 - ~~– sample complexities (2D gel spot vs MudPIT)~~
 - ~~– different databases (SwissProt vs NR)~~
- Impossible to compare results from different search algorithms, multiple instruments, and so on.



PeptideProphet better reporting

Because results are reported as probabilities, you can compare different programs, samples, and experiments.

- ~~• Poor sensitivity/specificity trade off unless you consider multiple scores simultaneously.~~ **discriminant score**
- ~~• No way to choose an error rate ($p=0.05$).~~ **estimate error with distributions**
- ~~• Need to have different thresholds for:~~ **curve-fit distributions to data (EM)**
 - ~~– different instruments (QTOF, TOF-TOF, IonTrap)~~
 - ~~– ionization sources (electrospray vs MALDI)~~
 - ~~– sample complexities (2D gel spot vs MudPIT)~~
 - ~~– different databases (SwissProt vs NR)~~
- ~~• Impossible to compare results from different search algorithms, multiple instruments, and so on.~~ **report P-values**

PeptideProphet Summary

- Identifies more peptides in each sample.
- Allows trade-offs: wrong peptides against missed peptides.
- Provides probabilities:
 - easy to interpret
 - comparable between experiments
- Automatically adjusts for each data set.
- Has been verified on many real proteome samples: see www.peptideatlas.org/repository.