

OVERVIEW

Purpose:

1. Benchmark peptide quantitation accuracy of our quantitative proteomics methodology with serial dilution samples of *Rhodopseudomonas palustris*
2. Profile *R. palustris* global protein expression change between anaerobic and aerobic growth conditions

Summary of Results:

1. In the quantitative comparison of serial dilution samples, some of the peptide ratios were clearly in error. Protein ratio measurements are less susceptible to errors since multiple peptide ratios are averaged.
2. 1318 proteins have been quantified between anaerobic and aerobic growth conditions of *R. palustris*

INTRODUCTION

Metabolic stable isotope labeling:

1. Stable isotope labeled analytes are excellent control for unlabeled analytes due to their identical biological activity.
2. Metabolic labeling is relatively straightforward, cheap and precise in mixing ratio compared with chemical labeling and enzymatic labeling.

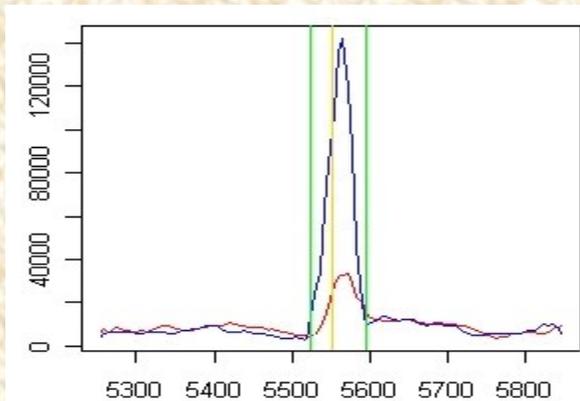


R. palustris growing under aerobic conditions

INTRODUCTION

Quantitative shotgun proteomics:

1. Quantitative shotgun proteomics can use identical LC-MS methodology developed for qualitative shotgun proteomics.
2. Peptides can be identified from MS² scans by database-searching algorithms. Peptides can be quantified from selected ion chromatograms by comparing their isotopic variants' abundance.



Selected ion chromatograms for two isotopic variants of a peptide (coded blue and red). Green lines delimit peak quantification boundary.

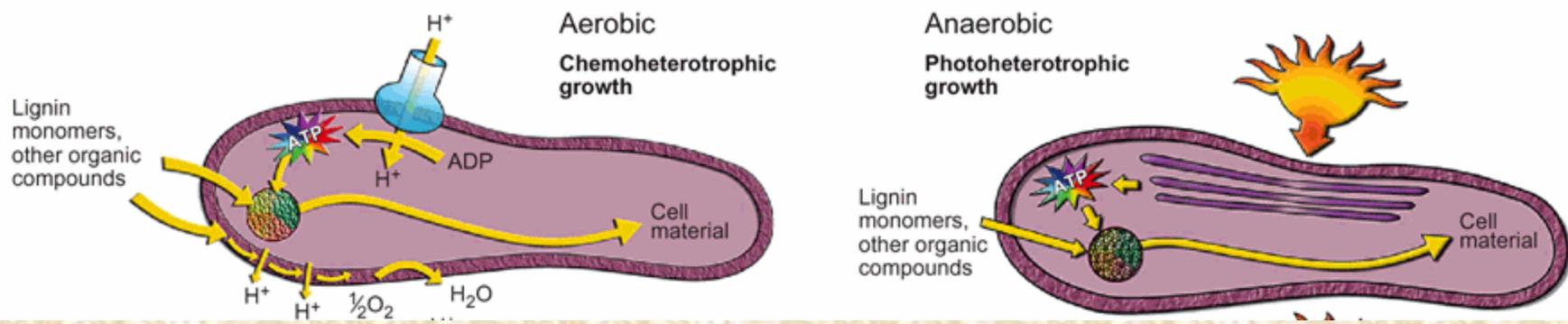
Rhodopseudomonas palustris:

1. A purple nonsulfur anoxygenic phototrophic bacterium that is ubiquitous in soil and water samples.
2. High metabolic diversity, capable of photoheterotrophic and photoautotrophic growth, as well as chemoheterotrophic and chemoautotrophic growth.
3. Its 5.459Mb genome has been sequenced and annotated with ~4836 potential proteins ⁽¹⁾.



Anaerobic vs. aerobic growth:

1. In aerobic growth condition, *R. palustris* utilize oxygen as the electron donor resource
2. In anaerobic growth condition, *R. palustris* utilize carbon or nitrogen as the electron donor resource
3. A wide range of proteins' expression will be different in these two growth conditions to switch between different electron donor resources.



Taken from ref #1

EXPERIMENTAL

Metabolic stable isotope labeling:

1. Culture cells on normal media with $^{14}(\text{NH}_4)_2\text{SO}_4$ versus on N15-enriched media with $^{15}(\text{NH}_4)_2\text{SO}_4$ (98% ^{15}N)
2. Harvest, wash and lyse cells
3. Mix labeled and unlabeled cell lysates at defined ratio of the total protein amount
4. Digest mixed cell lysate with trypsin
5. Clean up digestion product with solid phase extraction

Split-phase MudPIT on LTQ-MS:

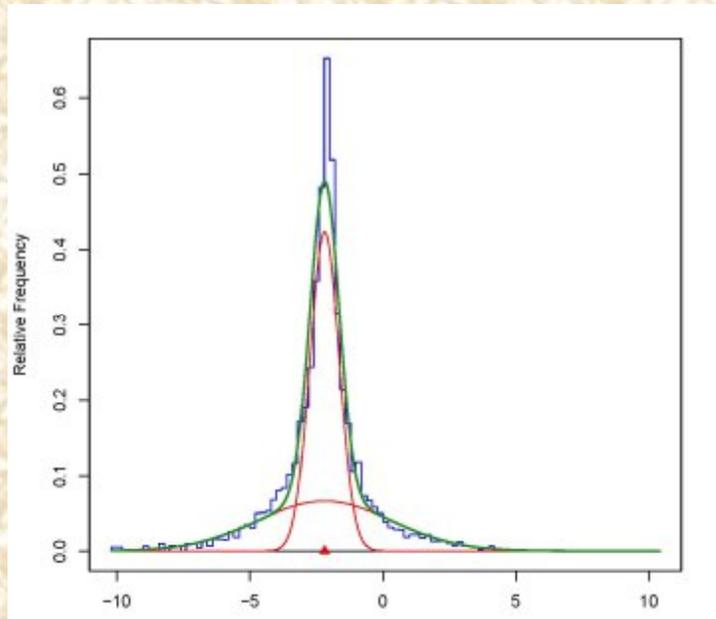
1. Split-phase column was packed as described (2)
2. 24-hour twelve-salt-step MudPIT elution
3. LTQ-MS:
 - I A MS¹ scan is followed by five data-dependent MS² scans with dynamic exclusion.
 - I Each scan is averaged from two microscans



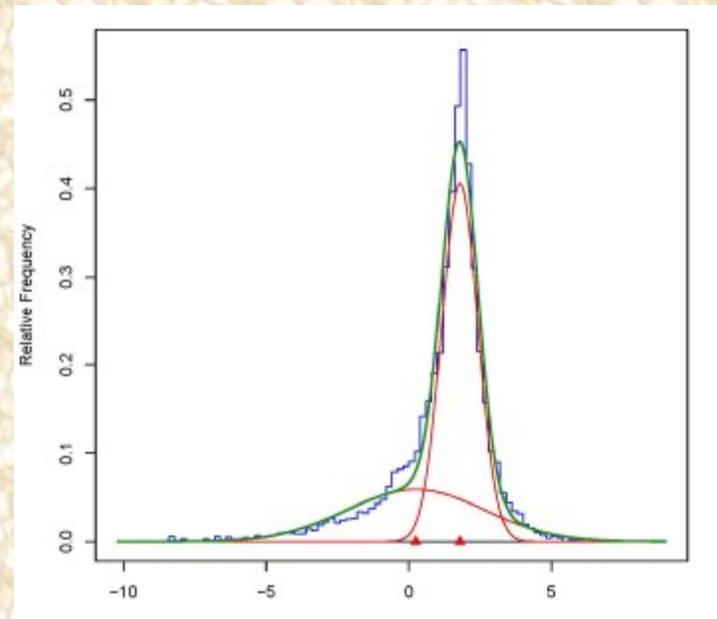
Quantitation Accuracy Benchmark Results

- I Serial dilution samples were prepared by mixing N14 and N15 *R. palustris* cell lysates at ratios of 10:1, 5:1, 1:5 and 1:10.
- I For a mixing ratio, histogram of logarithm of calculated peptide ratios was plotted in R (blue line in figure below)
- I Two gaussian curves were fit into the ratio distribution (red lines) and their sum was compared with the observed distribution (green line)

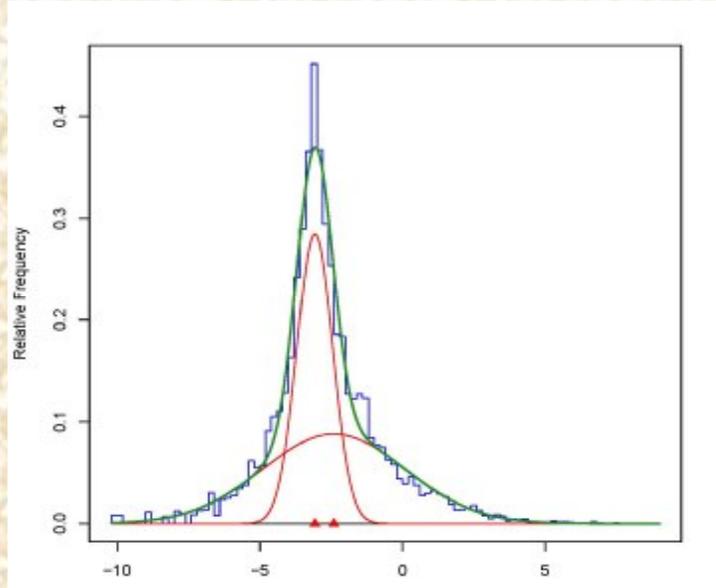
$N_{14}:N_{15}=1:5$



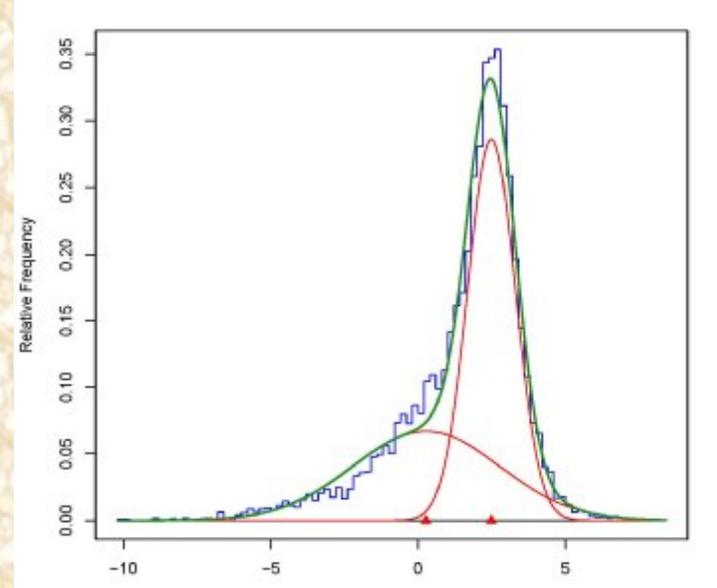
$N_{14}:N_{15}=5:1$



$N_{14}:N_{15}=1:10$



$N_{14}:N_{15}=10:1$



- False quantification: calculated ratio indicates wrong direction of abundance change
- Random error in true quantification: the variance of calculated ratios
- Systematic error in true quantification: the difference between the expected ratio and the mean calculated ratios.
- Dynamic range: extreme ratios suffers from large fraction of false quantification, large random error and large systematic error.

Aerobic vs. Anaerobic Growth Results

- | Bacteria were cultured aerobically on normal media and anaerobically on N15-enriched media.
- | Labeled and unlabeled cell lysates were mixed at equal ratio
- | To minimize random error of quantification, protein ratios were averaged from the ratios of at least three peptides

Ratio: Aerobic to Anaerobic

SD: standard deviation of peptide ratios

Pep #: number of quantified peptides

Selected down-regulated proteins; from aerobic to anaerobic

Protein Name	Ratio	SD	Pep. #
putative porphobilinogen deaminase 1671137:1672120 forward MW35601	0.204	0.157	12
putative 6-phosphofructokinase 413011:413949 forward MW32563	0.238	0.239	5
probable antioxidant protein 2771641:2772330 reverse MW25740	0.247	0.139	15
putative beta-glucosidase 1941645:1943021 reverse MW51738	0.254	0.193	13
possible branched-chain amino acid transport system substrate-binding protein 3717617:3718969	0.264	0.065	20
heat shock protein HslU proteasome-related ATPase subunit 335995:337296 reverse MW47489	0.27	0.339	5
unknown protein 2621691:2622218 reverse MW19205	0.287	0.147	5
fructose-bisphosphate aldolase 5225636:5226721 reverse MW39341	0.31	0.357	16
Ycel like family 4833016:4833657 forward MW22807	0.316	0.283	4
unknown protein 1575534:1575932 reverse MW13914	0.332	0.527	4
uroporphyrinogen decarboxylase 1676073:1677050 reverse MW35798	0.352	0.407	4
conserved unknown protein 2389539:2389871 forward MW12587	0.363	0.865	8
delta-aminoevulinic acid dehydratase 3084385:3085437 reverse MW38416	0.364	0.327	5
acyl-CoA synthetase 2609019:2610710 forward MW60776	0.376	0.592	7
cytochrome c556 2622816:2623256 forward MW15200	0.382	0.094	5
putative coproporphyrinogen oxidase III 1850616:1851965 reverse MW49338	0.398	0.528	10
5-aminoevulinic acid synthase (ALAS) 948749:949978 reverse MW44754	0.408	0.36	6
alcohol dehydrogenase 2273049:2274071 reverse MW36217	0.436	0.433	5
ribulose-bisphosphate carboxylase form II 5224192:5225577 reverse MW50485	0.446	0.933	16
possible GTP cyclohydrolase I 3832523:3833212 reverse MW25702	0.458	0.337	5
50S ribosomal protein L1 3690510:3691199 reverse MW24010	0.466	0.391	5
cysteine-tRNA ligase 2246922:2248313 forward MW51727	0.468	0.191	5
possible DNA-binding stress protein 3128515:3128997 reverse MW18247	0.473	0.555	5
5-aminoevulinic acid synthase (ALAS) 1724111:1725322 forward MW43799	0.478	0.123	8
conserved unknown protein 1843683:1844273 reverse MW23016	0.479	0.802	9
UDP-N-acetylglucosamine pyrophosphorylase 3026154:3027512 forward MW47047	0.48	0.41	5
beta-ketothiolase, acetoacetyl-CoA reductase 592255:592980 forward MW25466	0.488	0.451	5
conserved unknown protein 1886304:1886744 reverse MW15237	0.508	0.289	6
conserved unknown protein 4820269:4820985 forward MW26110	0.518	0.488	4
glucose-1-phosphate adenyltransferase 416649:417911 forward MW47026	0.519	0.426	5
possible lipoprotein 1582970:1583743 forward MW27813	0.52	0.479	6
putative chromosome segregation SMC protein 5063290:5066823 forward MW129213	0.522	0.612	8

Selected up-regulated proteins; from aerobic to anaerobic

Protein Name	Ratio	SD	Pep. #
acetyl-CoA synthetase 5081847: 5083757 reverse MW69381	3.582	12.317	21
probable alcohol dehydrogenase 4127608: 4128660 reverse MW36512	3.602	2.396	7
phosphoribosylaminoimidazole succinocarboxamide synthetase 4316467: 4317234 reverse MW29171	3.646	3.334	4
transcriptional regulator, LysR family 1981608: 1982498 reverse MW31727	3.672	1.165	9
possible branched-chain amino acid transport system substrate-binding protein 1566319: 1567563 forward	3.681	4.179	11
conserved unknown protein 5042658: 5043479 forward MW30383	3.688	3.068	9
putative trehalose synthase 4108522: 4111824 forward MW125271	3.699	3.731	9
conserved hypothetical protein 1282367: 1283224 reverse MW31371	3.716	3.018	4
putative coproporphyrinogen III oxidase precursor 1682058: 1682969 forward MW33661	3.723	2.731	4
sulfur/thiosulfate oxidase protein SoxB 5039800: 5041497 reverse MW61530	3.866	4.338	15
glycogen phosphorylase 5336312: 5338858 forward MW94541	3.922	4.639	5
putative acyl-CoA dehydrogenase 1792376: 1793548 reverse MW43105	3.924	4.718	6
putative DEAD-box protein, ATP-independent RNA helicase 4525204: 4527288 forward MW76129	4.037	4.893	6
SufB, needed for hnfF Fe-S center stability 2802346: 2803827 forward MW54960	4.077	3.928	5
TrapT family, dctP subunit, C4-dicarboxylate periplasmic binding protein 1997376: 1998386 forward	4.313	2.816	5
putative branched-chain amino acid transport system substrate-binding protein 1955272: 1956519 forward	4.434	4.48	13
pineloyl-CoA dehydrogenase (large subunit) 4193874: 4195064 reverse MW44541	4.458	5.759	6
response regulator receiver: histidine kinase 5407468: 5409120 forward MW60377	4.533	4.006	5
putative cell division protein FtsY 232613: 233560 reverse MW33143	4.64	6.419	4
possible methionyl-CoA small subunit 2064252: 2066123 reverse MW65276	4.702	11.652	12
conserved unknown protein 4679786: 4680169 reverse MW14022	4.89	3.005	4
possible branched-chain amino acid binding protein, ABC transport system 643359: 644579 reverse MW	4.924	3.904	11
cold shock DNA binding protein 3456174: 3456377 reverse MW7526	5.017	2.468	8
4-diphosphocytidyl-2C-methyl-D-erythritol synthase: YgbB 2944754: 2945950 reverse MW41660	5.042	6.559	4
magnesium coproporphyrin O-methyltransferase BchH subunit 1712857: 1716603 forward MW135985	5.172	6.28	6
bacterioferritin 4064442: 4064924 reverse MW18850	5.265	4.976	5
acetolactate synthase (small subunit) 2288912: 2289454 forward MW19942	5.31	5.679	4
putative aldolase 2080588: 2081661 reverse MW39645	5.564	2.557	7
ABC transporter, substrate-binding protein 981450: 982448 reverse MW35236	5.581	5.565	4
possible dipeptide ABC transporter (dipeptide-binding protein) 1602948: 1604501 forward MW56790	5.636	8.724	7
TrapT family, dctP subunit, C4-dicarboxylate periplasmic binding protein 2883903: 2884922 reverse	5.748	3.035	4
possible branched-chain amino acid transport system substrate-binding protein 1947510: 1948760 re	5.779	3.669	18

CONCLUSIONS

- 1. Peptide quantification results suffer from significant amount of error.**
- 2. Multiple peptide ratios are required to give good protein ratio estimation**
- 3. In total, 1318 proteins have been quantified between anaerobic and aerobic growth conditions of *R. palustris*.**
- 4. The expression of many genes for *R. palustris* is influenced strongly by the availability of oxygen.**

FUTURE DIRECTIONS

- 1. Better LC-MS methods and peptide quantitation algorithms are needed to improve the peptide quantitation accuracy.**
- 2. Community consensus criteria are needed to filter protein quantification results.**
- 3. Measuring protein expression changes between other growth conditions will increase the proteome coverage and provide more information on the dynamics of *R. palustris* proteome**

ACKNOWLEDGEMENTS

Research support was provided by the U.S. Department of Energy Office of Biological and Environmental Research Genomes-to-Life Program.

C. Pan acknowledges financial support from the ORNL-UTK Genome Science and Technology Graduate School.

Oak Ridge National Laboratory is managed and operated by the University of Tennessee-Battelle, L.L.C., under contract DE-AC05-00OR22725 with the U.S. Department of Energy.

Reference:

1. Larimer, F.W., P. Chain, L. Hauser, J. Lamerdin, S. Malfatti, L. Do, M.L. Land, D.A. Pelletier, J.T. Beatty, A.S. Lang, F.R. Tabita, J.L. Gibson, T.E. Hanson, C. Bobst, J.L. Torres, C. Peres, F.H. Harrison, J. Gibson, and C.S. Harwood, *Complete genome sequence of the metabolically versatile photosynthetic bacterium Rhodospseudomonas palustris*. Nat Biotechnol, 2004. **22**(1): p. 55-61.
2. McDonald, W.H. and J.R. Yates, 3rd, *Shotgun proteomics and biomarker discovery*. Dis Markers, 2002. **18**(2): p. 99-105.
3. Tabb, D.L. et al, Anal Chem, 2005 (in press)
4. Tabb, D.L., W.H. McDonald, and J.R. Yates, 3rd, *DTASelect and Contrast: tools for assembling and comparing protein identifications from shotgun proteomics*. J Proteome Res, 2002. **1**(1): p. 21-6.
5. MacCoss, M.J., C.C. Wu, H. Liu, R. Sadygov, and J.R. Yates, 3rd, *A correlation algorithm for the automated quantitative analysis of shotgun proteomics data*. Anal Chem, 2003. **75**(24): p. 6912-21.