Orthopox Proteomics and Host-Pathogen Interactions

Ryan D. Estep[1], Nathan P. Manes[2], Harold R. Udseth[2], Heather M. Mottaz[2], Joshua N. Adkins[2], Scott W. Wong[1], and Richard D. Smith[2]

[1]Vaccine and Gene Therapy Institute, Beaverton, OR 97007; [2]Pacific Northwest National Laboratory, Richland, WA 99352

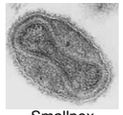
Overview

- Orthopox viruses are the largest and most complex of the animal viruses with >200 protein coding regions.
- Vaccinia (a non-lethal orthopox) is used in the smallpox vaccine.
- Monkeypox (a lethal orthopox) has recently emerged in Africa, and has infected hundreds of Africans and dozens of Americans.
- A comparative proteomic study of monkeypox and vaccinia has been initiated with the goal of identifying proteins required for virulence.
- LC-MS^n bottom-up proteomic analyses of purified virions has resulted in the identification of hundreds of monkeypox, vaccinia, and virion-associated host proteins.

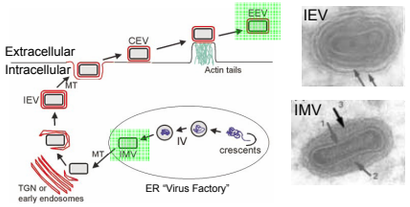
VGTI Biosafety Level 3	1. Virions Grown in He-La Cells 2. Sucrose Gradient Ultracentrifugation 3. RapiGest™ Trypsin Digestion
PNNL Biosafety Level 1	4. SCX Fractionation 5. C-18 LC-MS^n 6. SEQUEST, X!Tandem, AMT Analysis

Introduction

- Double-Stranded DNA Viruses
- Genome – 200 kBp
- Proteome – 220 Predicted Proteins
- 200nm Diameter, 300nm Long
- Most Complex of Animal Viruses
- Host Cell Attachment/Entry and Viral Uncoating are Poorly Understood



Smallpox



IEV

IMV

IMV – Intracellular Mature Virus – Infective Particle
IEV – Intracellular Enveloped Virus – Low Infectivity
CEV – Cell-Associated Enveloped Virus – Low Infectivity
EEV – Extracellular Enveloped Virus – Infective Particle

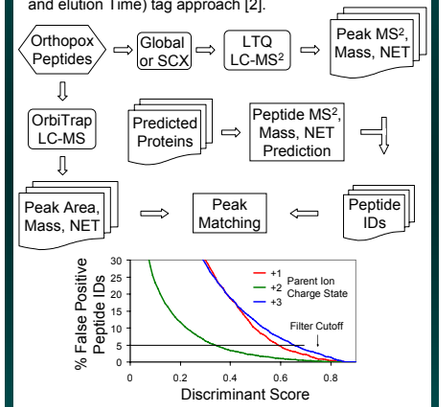
Methods

Sample Preparation and LC-MS^n

Orthopox grown in He-La cells were purified by sucrose gradient ultracentrifugation, denatured using Waters RapiGest™ surfactant, and digested with trypsin. Unfractionated samples and strong cation exchange (SCX) HPLC fractions were analyzed by reversed-phase HPLC / nano-electrospray / Orbitrap MS and LTQ MS^2.

Data Analysis

LC-MS^2 spectra were analyzed by SEQUEST and X!Tandem using normal and scrambled dual organism (viral/host) protein datasets. Peptide identifications were filtered using a discriminant score [1] based on SEQUEST/X!Tandem scores, predicted NETs (Normalized Elution Times), and other factors to reduce false positive peptide identifications to < 5%. LC-MS spectra were analyzed using the AMT (Accurate Mass and elution Time) tag approach [2].



A data directed global normalization removed sample preparation biases, and a nonlinear least squares analysis determined protein abundance values.

$$R^2 = \sum_{i,j}\left(1 - \frac{\text{Efficiency}(i) * \text{ProteinAbundance}(j)}{\text{PeptideAbundance}(i,j)}\right)^2$$

Example:

Protein	Peptide	IMV1	IMV2	IMV3	EEV1	EEV2	EEV3	Efficiency
MPXV-ZRE_087	DFYIPHGNK	1.48	1.42	1.05				0.21
MPXV-ZRE_087	FVDEEYLR	5.72	6.18	6.33	8.54	9.31	10.01	0.82
MPXV-ZRE_087	IILISDVR	7.46	9.24	6.62	31.14	31.78	27.34	1.43
MPXV-ZRE_087	ILFLQQSIFR	1.46	1.09	0.86	2.65	2.64	2.23	0.19
MPXV-ZRE_087	LLSIYTGENMR	16.41	15.07	13.54	26.17	27.86	28.48	2.31
MPXV-ZRE_087	MLQPFAPSYSAEMR	13.06	12.88	11.59	26.36	24.86	20.40	1.94
MPXV-ZRE_087	SDAVNYEK		0.61		1.15	1.99	2.65	0.11
ProteinAbundance		6.58	6.40	5.28	11.75	13.68	12.28	

Two monkeypox preparations (one IMV and one EEV) were analyzed in triplicate by Orbitrap LC-MS. Missing (grey) and outlier (yellow) peptide abundance values were excluded from the sum of squares.

Results

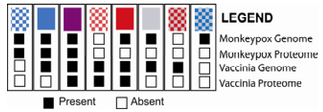
Monkeypox Viral/Host Peptide Identification Overview

Spectra resulting from LC-MS^2 analysis of global (unfractionated) and SCX fractionated samples were analyzed by both SEQUEST and X!Tandem and unique peptide identifications were tallied. LC-MS peptide identifications were analyzed using the AMT tag approach. SEQUEST analyses using monkeypox and vaccinia "Stop-to-Stop" protein datasets failed to identify any unannotated proteins.

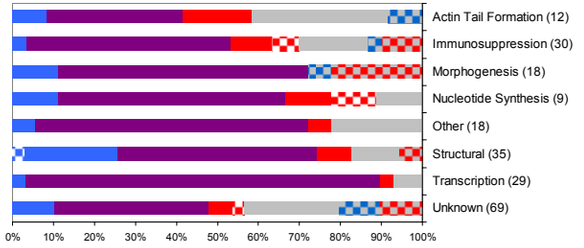
	# Unique Peptide Identifications			
	LC-MS^2			LC-MS
Identified by:	*Global*	*SCX*	*Total*	*Global*
SEQUEST, but not X!Tandem	895	2079	2200	588 (27%)
X!TANDEM, but not SEQUEST	2164	5777	6305	1985 (31%)
(unmodified and modified)	780	3076	3286	1841 (56%)
Both search engines	5816	12,171	13,620	8174 (60%)
Total Peptides	9655	23,103	25,411	12,588 (50%)
Total Proteins	1848	2968	3001	1503 (50%)

Comparison of Monkeypox and Vaccinia by LC-MS^2

Monkeypox (MPV) and vaccinia (VAV) viral proteins were each assigned one of eight functional annotations (listed right with the number of assigned proteins). In addition, each protein was classified as being present in or absent from each orthopox proteome and genome.



For each functional annotation, the percentage of each protein classification is displayed as a stacked bar chart.

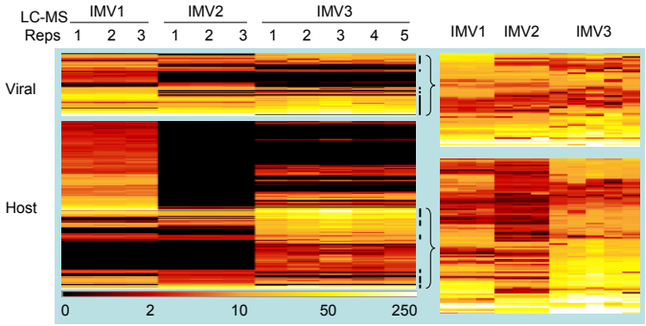


Virion-Associated and Contaminant Host Proteins

10 putative virion-associated proteins and 10 putative contaminants were selected from the top 100 most abundant host-cell proteins determined by LC-MS. Host proteins were putatively identified as virion-associated or as contamination based on their known roles.

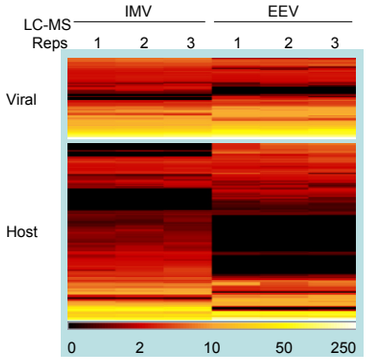
Putative Virion-Associated Host Proteins (Role)	Putative Contaminants (Role)
Actin (Structural)	Alkaline Phosphatase (Dephosphorylation)
Annexin (Membrane Structure)	Elongation Factors (Translation)
Basigin (Structural)	Enolase (Metabolism)
Cofilin (Structural)	Glyceraldehyde-3-Phosphate Dehydrogenase (Metabolism)
Filamin (Structural)	Heat Shock Proteins (Stress Response)
Ion Channels (Ion Transport)	Histones (Chromatin Structure)
Lamin (Structural)	Initiation Factors (Translation)
Moesin (Structural)	Keratin (Structural)
Profilin (Structural)	Peroxiredoxin (Antioxidation)
Tubulin (Structural / Transport)	Pyruvate Kinase (Metabolism)

Comparison of Monkeypox IMV Preps by LC-MS



Three monkeypox IMV (Intracellular Mature Virus) preparations were multiply analyzed by LC-MS (Orbitrap). Rows represent a single protein, and protein abundance is indicated by the color bar (bottom). Unlike the first prep, preps 2 and 3 were purified using a second sucrose gradient. This resulted in moderately different viral protein sub-proteomes, and significantly different host protein sub-proteomes. Prep 2 appears to have the fewest contaminating host proteins. Prep 3 was a scaled-up version of prep 2. Proteins observed in all three preps (smaller heat maps on the right) are putatively virion-associated. Note the good agreement between LC-MS replicate analyses.

Comparison of Monkeypox IMV and EEV by LC-MS



Putative IMV-Specific Host Proteins
B-Cell Receptor-Associated Protein
Calnexin
Heat Shock Protein
Ion Channel
Prohibitin
Ras-Eelated Protein Rab-1A

Putative EEV-Specific Host Proteins
14-3-3 Protein
Brain Acid Soluble Protein
Destrin
Erythrocyte Band 7 Membrane Protein
Nuclear Protein RAN
Liprin
Moesin
Radixin
Receptor Tyrosine Phosphatase
Tubulin
Villin

Monkeypox IMV (Intracellular Mature Virus) was isolated from disrupted host cells and EEV (Extracellular Enveloped Virus) was isolated from the corresponding culture media. The quality of the parallel virion preparations was determined by electron microscopy, and each sample was analyzed by LC-MS in triplicate. Consistent with the current orthopox literature, IMV and EEV virions were found to have very similar viral protein sub-proteomes and significantly different host protein sub-proteomes.

Conclusions

- Orthopox denatured with Waters RapiGest™ surfactant were successfully analyzed by bottom-up LC-MS^n.
- Parallel SEQUEST and X!Tandem analyses of MS^2 spectra produced high confidence peptide identifications.
- Nonlinear least squares analyses reduced raw peptide abundance data into useful protein abundance data.
- LC-MS^n of peptides derived from purified orthopox resulted in the identification and quantification of hundreds of viral and host-cell proteins.
- Comparative proteomic analysis of monkeypox and vaccinia resulted in the identification of 23 monkeypox-unique and 20 vaccinia-unique viral proteins.
- Comparative proteomic analysis of monkeypox IMV and EEV resulted in the identification of 7 IMV-unique and 4 EEV-unique viral proteins.

Future Work

- Study Infection in Host Cells (He-La, THP-1) at 2, 4, and 12 Hours Post-Infection to Identify Early, Mid, and Late Expressing Viral Proteins
- Hypothesis Driven Experiments (e.g. Mutants)
- Validation of Proteomics Results using Orthogonal Experiments (e.g. Immunogold Electron Microscopy)
- Further Development of Nonlinear Least Squares Methodology to Improve Analysis of Proteins with Large Numbers of Missing Peptide Abundance Values.

Acknowledgements

Portions of this research were supported by the NIH National Institute for Allergy and Infectious Diseases (N01-AI-40053). Experimental portions of this research were performed in the Environmental Molecular Sciences Laboratory, a U.S. Department of Energy (DOE) national scientific user facility located at the Pacific Northwest National Laboratory (PNNL) in Richland, Washington. PNNL is a multiprogram national laboratory operated by Battelle for the DOE under Contract No. DE-AC05-76RLO 1830.

References

1. Strittmatter E.F., et. al.; Application of peptide LC retention time information in a discriminant function for peptide identification by tandem mass spectrometry; J Proteome Res 3(4):760-9; 2004
2. Zimmer J.S., et. al.; Advances in proteomics data analysis and display using an accurate mass and time tag approach; Mass Spectrom Rev 25(3):450-82; 2006
3. Orthopox Figures: J Gen Virol 83:2915–31; Trends Microbiol 10:15-24; http://www.zkea.com/archives/archive02001.html

For more information, please visit:
http://ncrr.pnl.gov/ and http://www.proteomicsresource.org/

Contact Information

Nathan Manes, Ph.D.
Biological Sciences Division, K8-98
Pacific Northwest National Laboratory
P.O. Box 999, Richland, WA 99352
E-mail: nathan.manes@pnl.gov
Phone: (509) 376-0480