

The Mtb Proteome Library:

Development and application of assays for targeted MS analysis of the complete proteome of *Mycobacterium tuberculosis* by SRM and SWATH-MS



Olga Schubert

Group of Prof. Ruedi Aebersold
ETH Zurich

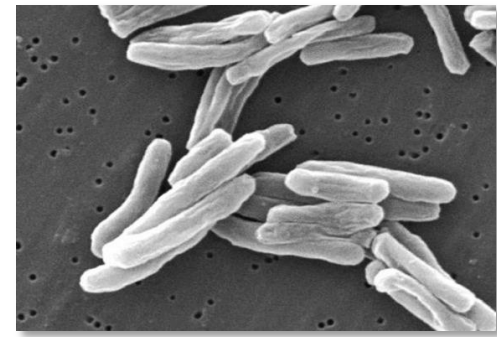
Skyline User Meeting 2013 Minneapolis



ETH

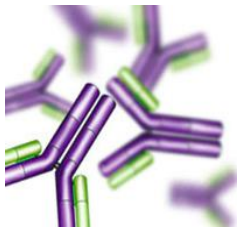
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Mycobacterium tuberculosis (Mtb)



- Mycobacterium tuberculosis is the causative agent of Tuberculosis (TB)
- One third of the world's population latently infected with Mtb
- 1.7 million deaths from TB each year
- More efficient treatments urgently needed
- **Limited availability of techniques to measure proteins with high sensitivity, selectivity and reproducibility**

Traditional approach: Antibodies

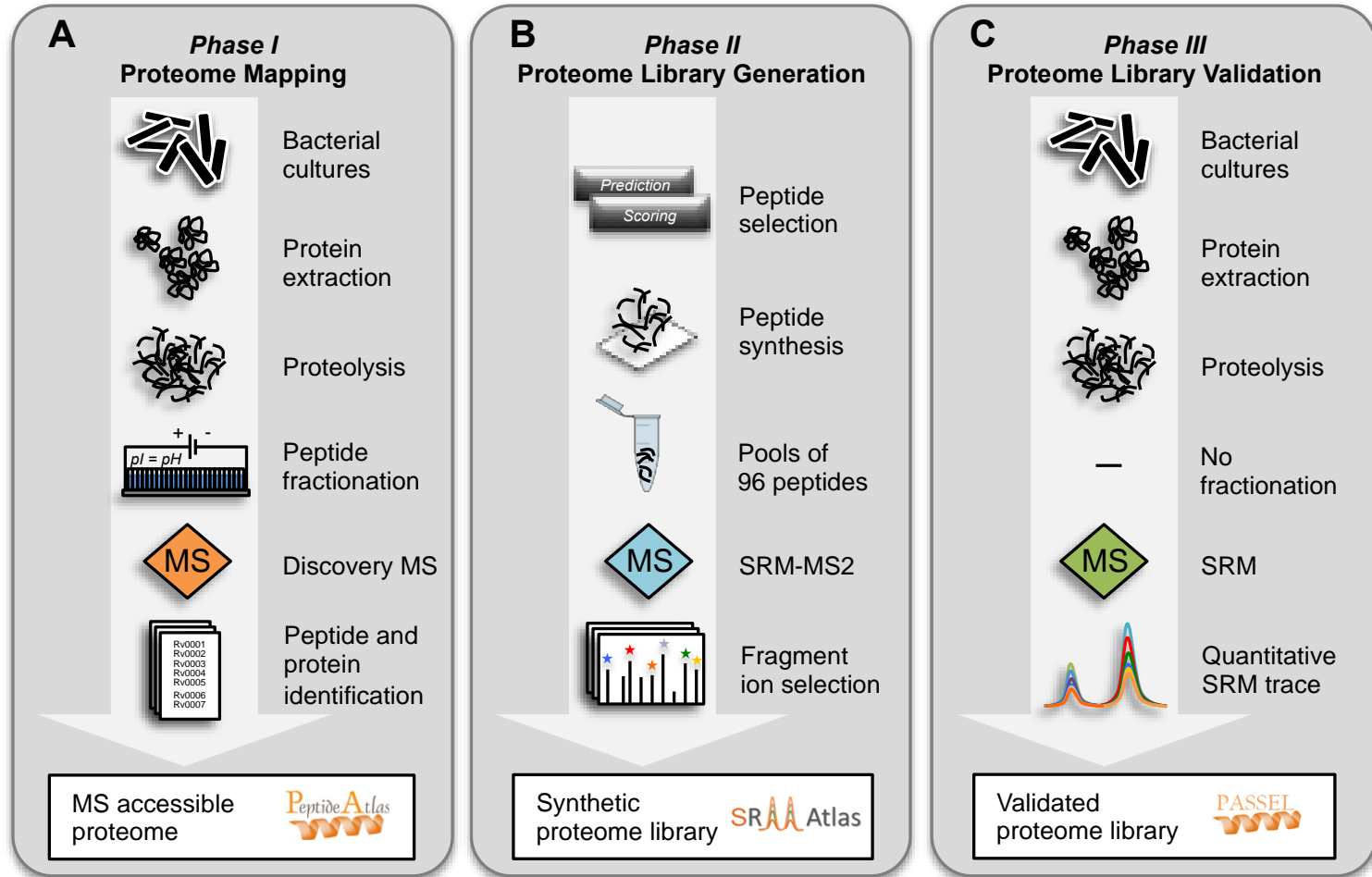


Only for few targets,
high cost, low
throughput

Aim

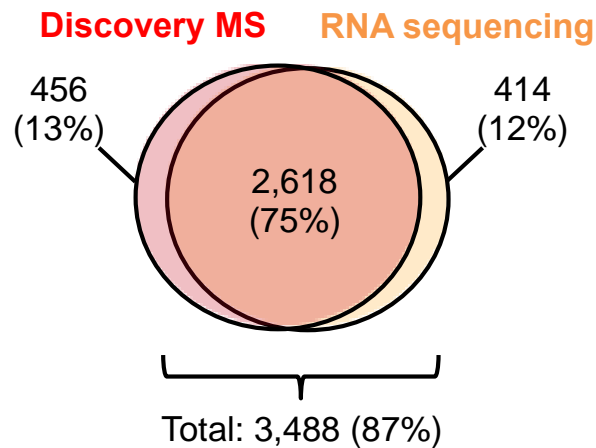
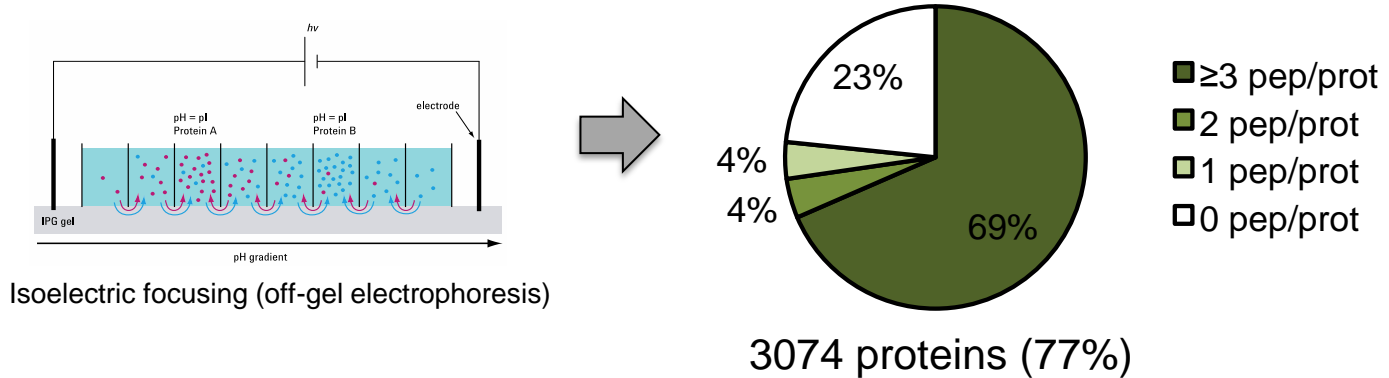
To generate a resource of validated assays for the sensitive detection and accurate quantification of every protein of *Mycobacterium tuberculosis*, even in complex backgrounds.

The Mtb Proteome Library contains SRM assays for the entire proteome of Mtb



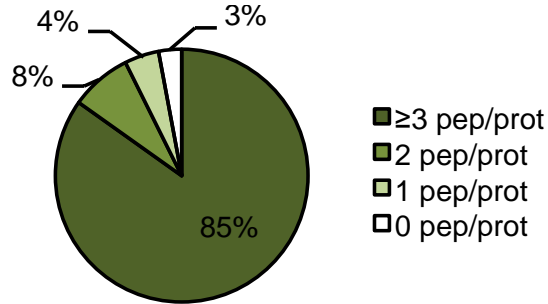
Definition of the MS-accessible Mtb proteome by discovery MS

Extensive fractionation and shotgun MS on Orbitrap XL

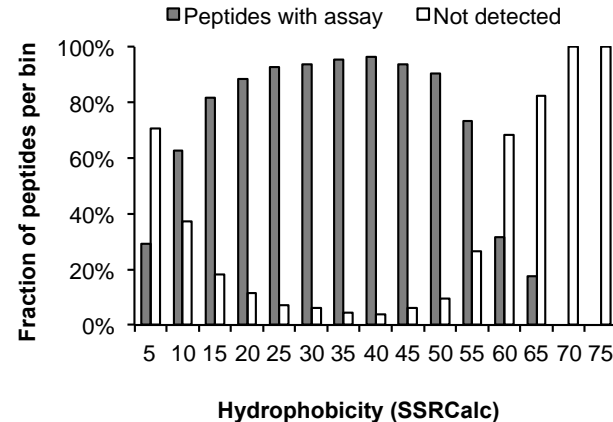
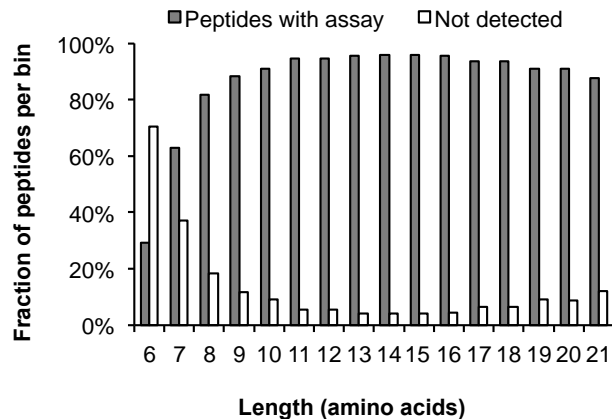
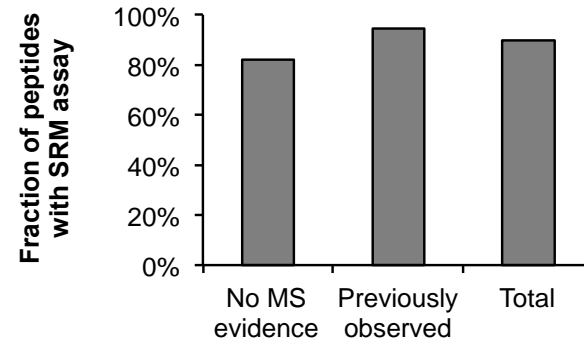


100% corresponds to the 4,012 annotated ORFs in Mtb (TubercuList v2.3)

Generation of SRM assays using crude synthetic peptides

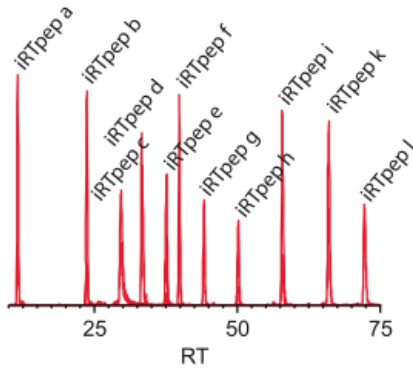


3894 proteins (97%)



17,463 crude synthetic peptides (JPT) measured in pools of ~100 on a Qtrap 4000 in SRM-triggered MS2 mode

Increasing SRM/SWATH assay specificity and throughput by using iRTs and scheduled SRM



Edit iRT Calculator

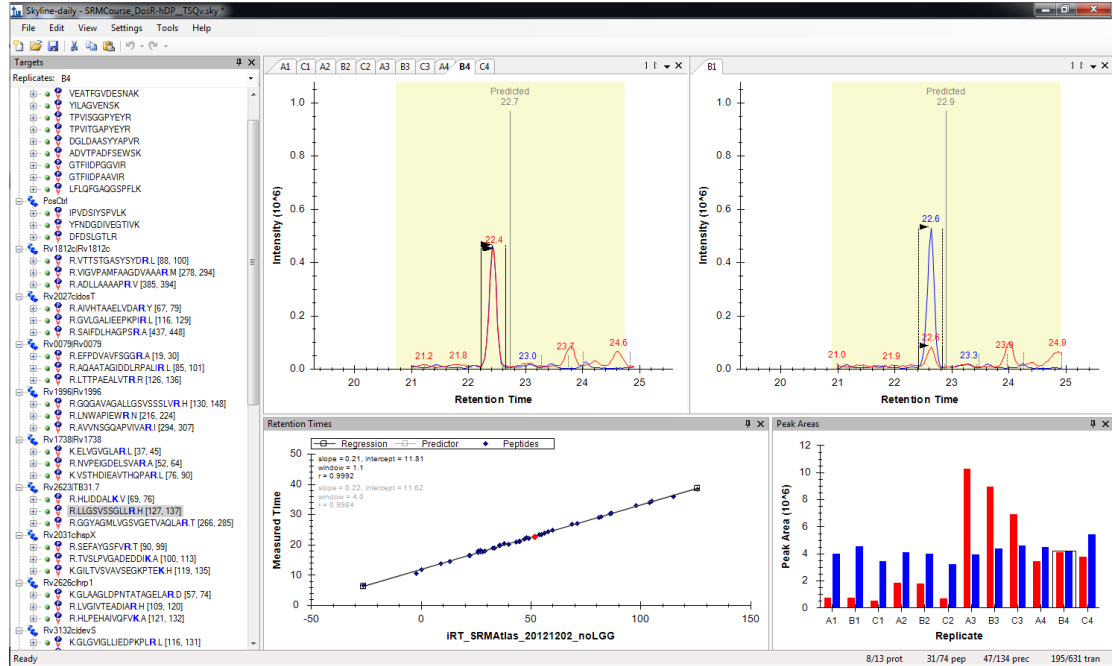
Name: iRT_SRMAtlas_20121202_noLGG

iRT database: C:\Users\user\Desktop\Skyline\iRT_SRMAtlas_201

Modified Sequence	iRT Value
GAGSSEPVTGLDAK	0.23
VEATFGVDESNAK	13.11
YILAGVENS K	22.38
TPVISGGPYEYR	29.00
TPVITGAPYEYR	33.63
DGLDAASYAPV R	43.28

Modified Sequence	iRT Value
SAPHADSI EAVK	-9.00
DPPDPHQPD MTK	-14.00
WGFGLAVC(+57.0)DGEK	73.00
HLVDIDEALNMAR	50.00
DTDIGQP C(+57.0)SPEGAK	-8.00
LWGNPGPYC(+57.0)ER	47.00

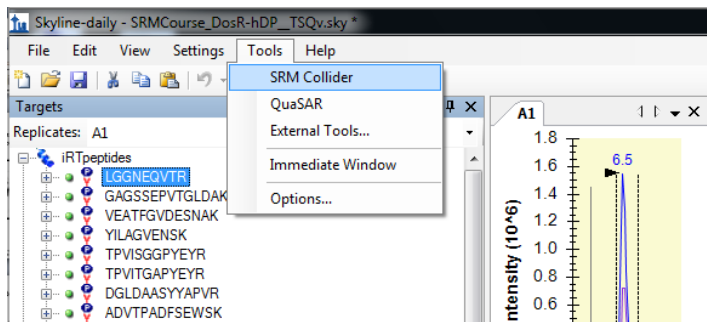
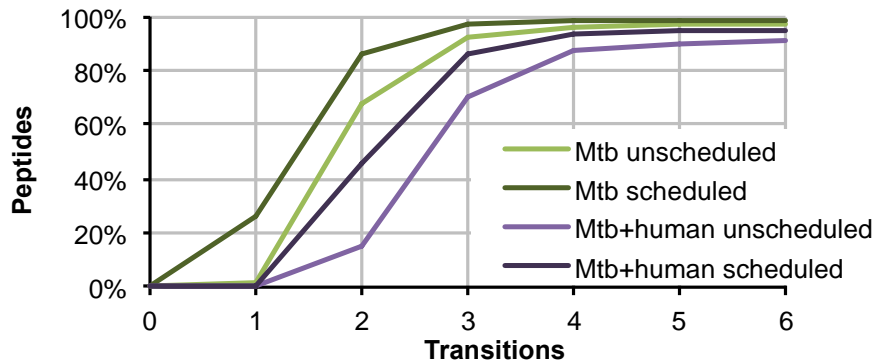
16461 Peptides



iRT peptides spiked into each sample allow to determine a chromatography-independent retention time (iRT) for each peptide.



Theoretical assessment of SRM assay specificity using the SRM Collider



version 1.4
Hannes Röst 2012

SRM Collider

Collider
Download
About
Instructions

The SRM Collider is a program that will take your input transitions and compare them to all other transitions in a given background proteome and find interferences. It will report these interferences on a per-peptide basis, allowing a researcher to identify peptides that share many transitions with the target peptide.

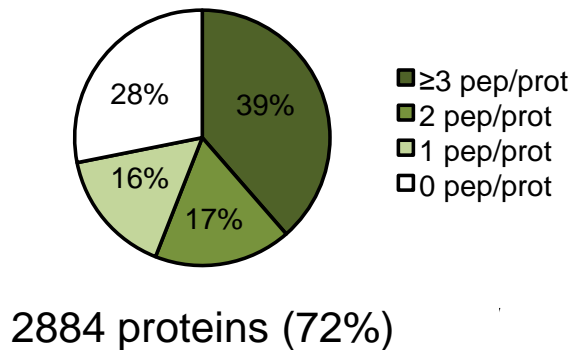
Please enter the peptide sequences here (see [instructions](#) for help):

```

LSDVVDLDQVIQEIR[166]/2, LSDVVDLDQVIQEIR[166], 2, 929.987772, 873.44574, #N/A, 2, y15
LPDNGIELC[160]R/2, LPDNGIELC[160]R, 2, 612.308612, 1031.473119, #N/A, 1, y9
LPDNGIELC[160]R/2, LPDNGIELC[160]R, 2, 612.308612, 915.446176, #N/A, 1, y8
LPDNGIELC[160]R/2, LPDNGIELC[160]R, 2, 612.308612, 582.276257, #N/A, 1, y4
LPDNGIELC[160]R/2, LPDNGIELC[160]R, 2, 612.308612, 453.233664, #N/A, 1, y3
LPDNGIELC[160]R/2, LPDNGIELC[160]R, 2, 612.308612, 561.76658, #N/A, 2, y10
LPDNGIELC[160]R[166]/2, LPDNGIELC[160]R[166], 2, 617.312746, 1041.481388, #N/A, 1, y9
LPDNGIELC[160]R[166]/2, LPDNGIELC[160]R[166], 2, 617.312746, 925.454445, #N/A, 1, y8
LPDNGIELC[160]R[166]/2, LPDNGIELC[160]R[166], 2, 617.312746, 592.284526, #N/A, 1, y4
LPDNGIELC[160]R[166]/2, LPDNGIELC[160]R[166], 2, 617.312746, 463.241933, #N/A, 1, y3
LPDNGIELC[160]R[166]/2, LPDNGIELC[160]R[166], 2, 617.312746, 566.770714, #N/A, 2, y10
                    
```

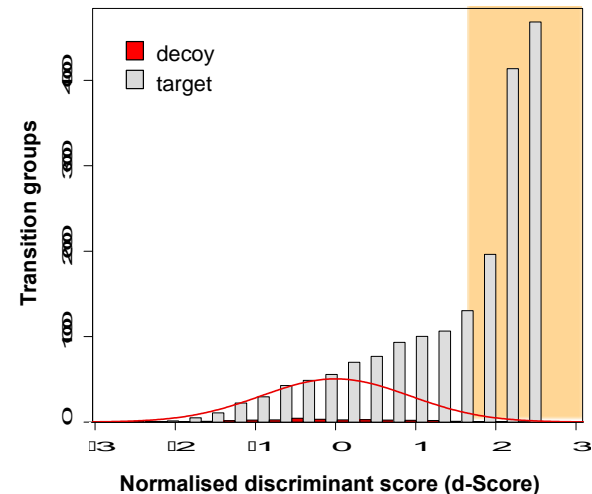
SSRCalc window	10	arbitrary units
Q1 mass window	0.7	Th
Q3 mass window	1.0	Th
Low mass threshold for transitions	300	Th
High mass threshold for transitions	1500	Th
Genome	Mycobacterium bovis ▾	
Consider isotopes up to	3	amu
Missed Cleavages	1	▾
Find UIS up to order*	2	

Validation of the Mtb Proteome Library in unfractionated whole cell lysates by SRM



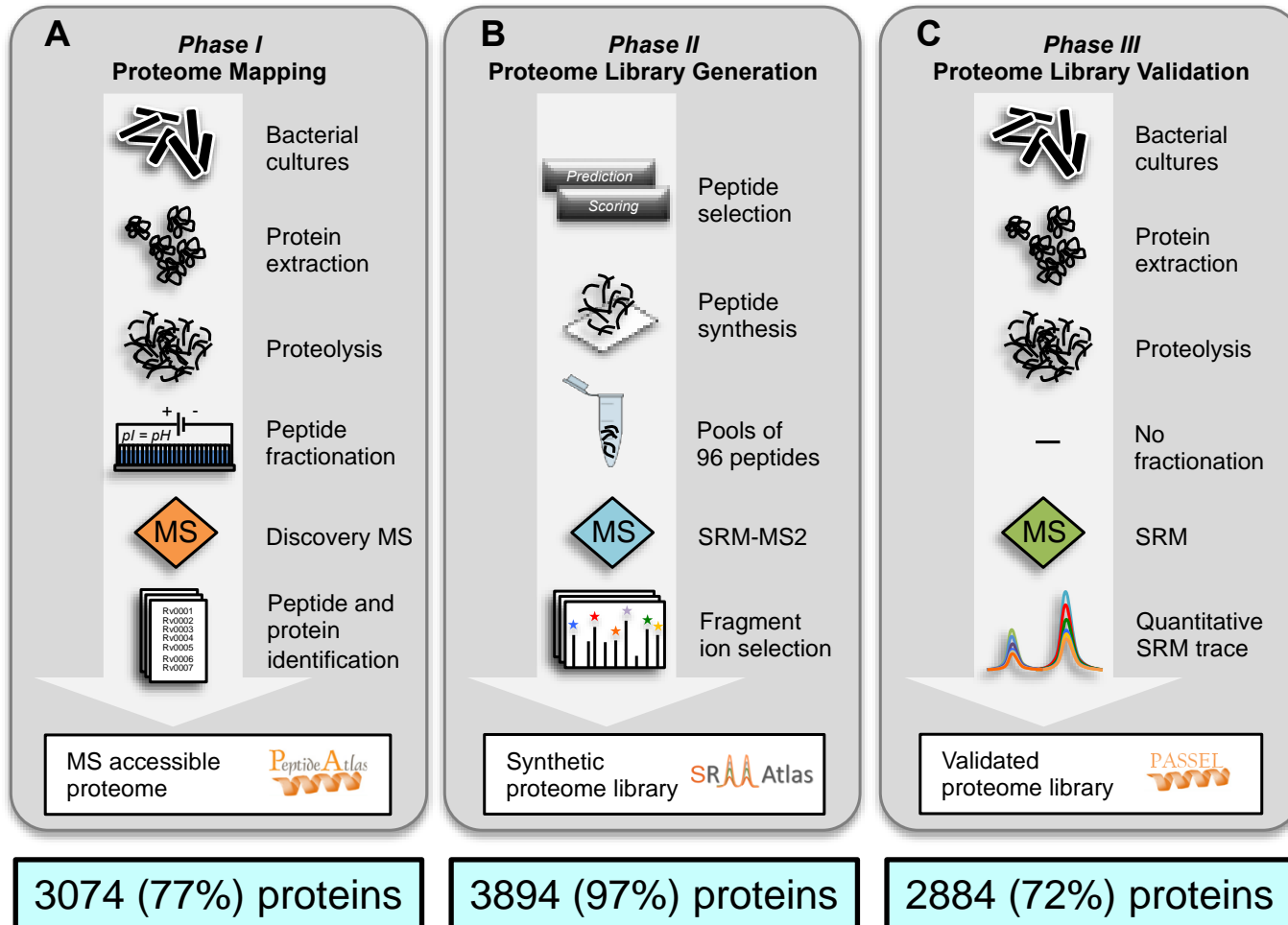
mProphet analysis

d-Score < 1.78
m-Score < 0.01



To validate all these SRM assays, over 200 scheduled SRM runs were needed.

The Mtb Proteome Library contains SRM assays for the entire proteome of Mtb



The Mtb Proteome Library is a publicly available resource of MS reference data, SRM assays and their validation

www.PeptideAtlas.org

www.PeptideAtlas.org/PASSEL

SRMATLAS HOME Search All Builds Current Build Queries SRMATLAS PTPAtlas Submission

PeptideAtlas Build: M. tuberculosis 2013-05
 Protein Name: Rv0001

C Rv0001
 Protein Name: Rv0001
 Gene Name: dnaA
 Description: Rv0001dnaA
 Identification Status: canonical
 Protein Group: Rv0001
 Protein/Prophet Probability: 1.00000
 Mult Hyp Testing Probability: 1.00000
 Distinct Peptides: 31
 Total Observations: 443
 Protein/Prophet-adjusted N Obs: 126
 Normalized PSMs per 100K: 12.718

C Sequence Motifs

Observed Peptides: [Peptide list]

Sequence Coverage: [Graph]

Unlikely (theoretical): [Graph]

Sequence Position: [Graph]

Observed peptide with single genome mapping
 Protein coverage by observed peptides
 Peptides unlikely to be observed

C Sequence
 Sequence Display Mode: [Typic]

Protein Coverage = 62.1% (75.9% of likely observable sequence)

SRMATLAS HOME Search All Builds Current Build Queries SRMATLAS PTPAtlas Submission

Query Transitions Transition Lists SRMATLAS Builds PASSEL Experiments PASSEL Data

Works best under Firefox.

Get SRM Experiment Transitions
 Form Resultset ChromaView Plot

Select left and right chromatograms from the Results tab.
 Drag the red area over the context panel to shift the focus panel. When the cursor is a crosshair, you may draw a new context window. Lock/unlock X and Y axes as desired.

YVSTEEFTNDFINSLR (1933.906 Daltons) +2, light
 Experiment: M. tuberculosis SRM atlas (Schubert et al.)
 Spectrum file: oligas_H110902_105_TRID30.mzXML
 Chromatogram ID: 296421
 mQuest: best pg RT=27.726 SN=66.713 log max apex intens=4.564

YVSTEEFTNDFINSLR (1933.906 Daltons) +3, light
 Experiment: M. tuberculosis SRM atlas (Schubert et al.)
 Spectrum file: oligas_H110902_107_TRID31.mzXML
 Chromatogram ID: 296463
 mQuest: best pg RT=27.617 SN=23.614 log max apex intens=3.981

Intensity vs. Time (min) chromatograms for +2 and +3 charge states.

Y-axis: Intensity
 X-axis: Time (min)

Auto Scale Y Axis

www.SRMAtlas.org

SRMATLAS HOME Search All Builds Current Build Queries SRMATLAS PTPAtlas Submission

Query Transitions Transition Lists SRMATLAS Builds PASSEL Experiments PASSEL Data

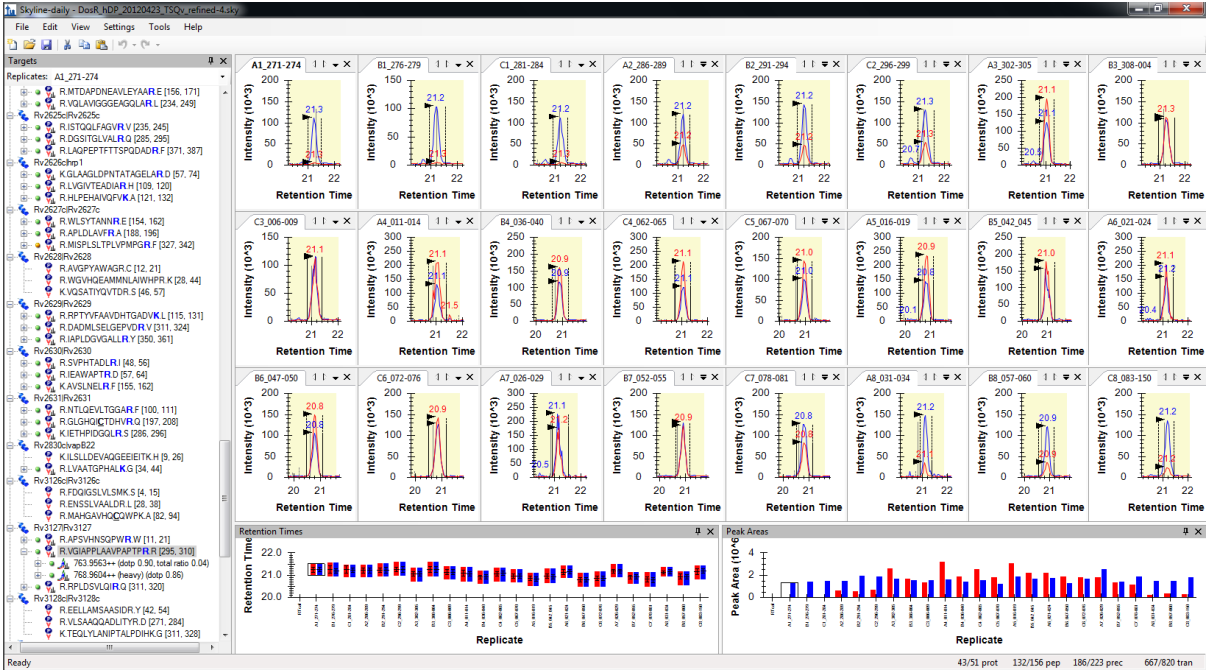
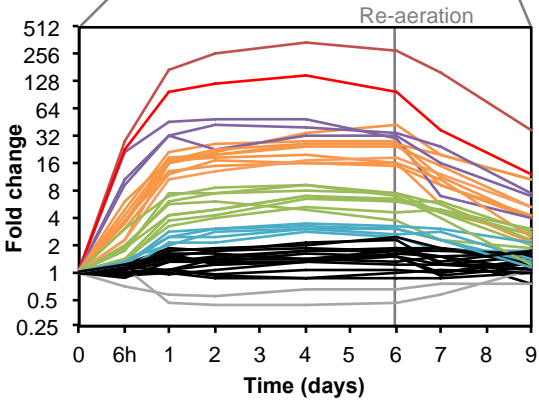
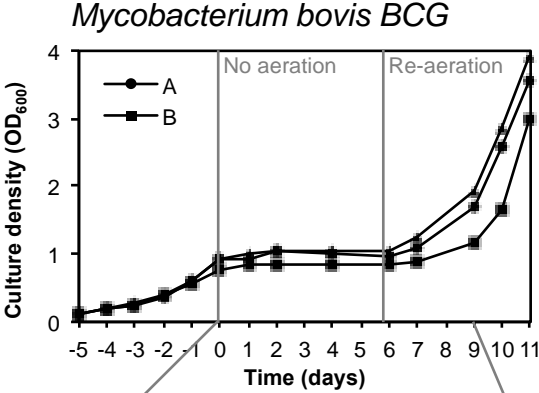
show column descriptions

Protein	Pvs AA	Sequence	For AA	Adj BS	Source	Q1_mz	Q1_std	Q2_mz	Q2_std	Ion	Rank	RI	SRR1	N_map	IonTrap
Rv0001	R	YTFDFVIGASNR	F	1.73	OTOF	748.87	2.504.2525	1.7771	33.9	1	1.7771	33.9	1	1	
Rv0001	R	YTFDFVIGASNR	F	1.73	OTOF	748.87	2.617.2365	1.96	2.8015	33.9	1	1.96	2.8015	33.9	1
Rv0001	R	YTFDFVIGASNR	F	1.73	OTOF	748.87	2.716.4049	1.17	3.5012	33.9	1	1.17	3.5012	33.9	1
Rv0001	R	YTFDFVIGASNR	F	1.73	OTOF	748.87	2.1268.6164	1.11	4.4579	33.9	1	1.11	4.4579	33.9	1
Rv0001	K	VDDGPPSDANLSAPLTPQOR	A	1.68	OTOF	1034.50	3.830.4734	1.17	1.7796	24.9	1	1.17	1.7796	24.9	1
Rv0001	K	VDDGPPSDANLSAPLTPQOR	A	1.68	OTOF	1034.50	2.528.2889	1.14	2.5860	24.9	1	1.14	2.5860	24.9	1
Rv0001	K	VDDGPPSDANLSAPLTPQOR	A	1.68	OTOF	1034.50	2.997.5425	1.19	3.2548	24.9	1	1.19	3.2548	24.9	1
Rv0001	K	VDDGPPSDANLSAPLTPQOR	A	1.68	OTOF	1034.50	2.910.5105	1.18	4.2217	24.9	1	1.18	4.2217	24.9	1
Rv0001	K	YVSTEEFTNDFINSLR	D	1.37	OTOF	967.96	2.749.4304	1.16	1.2722	37.4	1	1.16	1.2722	37.4	1
Rv0001	K	YVSTEEFTNDFINSLR	D	1.37	OTOF	967.96	2.1079.5480	1.19	2.1823	37.4	1	1.19	2.1823	37.4	1
Rv0001	K	YVSTEEFTNDFINSLR	D	1.37	OTOF	967.96	2.489.2780	1.14	3.1359	37.4	1	1.14	3.1359	37.4	1
Rv0001	K	YVSTEEFTNDFINSLR	D	1.37	OTOF	967.96	2.263.1390	1.12	4.1298	37.4	1	1.12	4.1298	37.4	1
Rv0001	R	FAHAAALAJAEAPAR	A	1.70	OTOF	493.94	3.219.11	1.12	1.100	29.3	1	1.12	1.100	29.3	1
Rv0001	R	FAHAAALAJAEAPAR	A	1.70	OTOF	493.94	3.614.35	1.16	2.81	29.3	1	1.16	2.81	29.3	1
Rv0001	R	FAHAAALAJAEAPAR	A	1.70	OTOF	493.94	3.1234.66	1.13	3.74	29.3	1	1.13	3.74	29.3	1
Rv0001	R	FAHAAALAJAEAPAR	A	1.70	OTOF	493.94	3.569.28	1.16	4.58	29.3	1	1.16	4.58	29.3	1
Rv0001	K	THLHAAGNYAQR	L	1.57	OTOF	363.69	4.374.22	1.13	1.70	21.4	1	1.13	1.70	21.4	1
Rv0001	K	THLHAAGNYAQR	L	1.57	OTOF	363.69	4.239.11	1.12	2.43	21.4	1	1.12	2.43	21.4	1
Rv0001	K	THLHAAGNYAQR	L	1.57	OTOF	363.69	4.352.20	1.13	3.36	21.4	1	1.13	3.36	21.4	1
Rv0001	K	THLHAAGNYAQR	L	1.57	OTOF	363.69	4.915.48	1.19	4.36	21.4	1	1.19	4.36	21.4	1

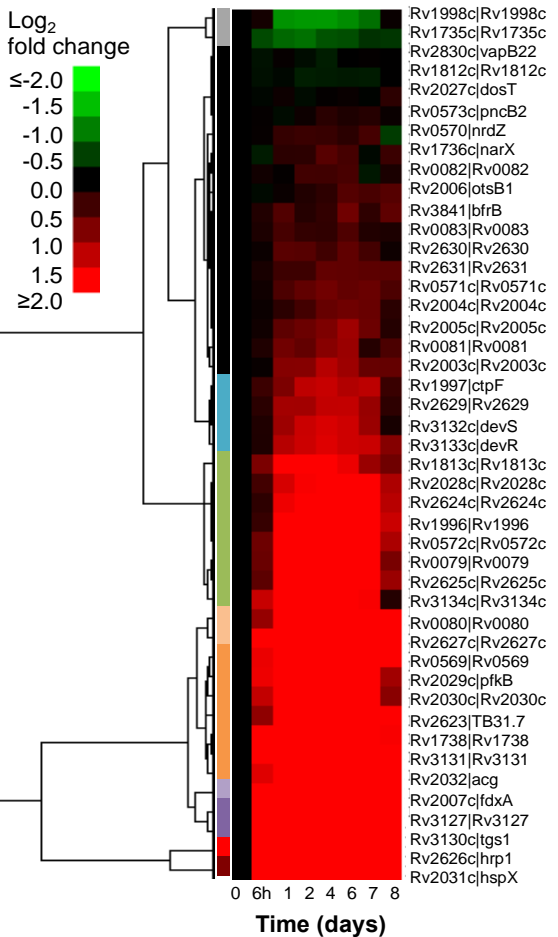
Application of the Mtb Proteome Library to study the dynamics of the DosR regulon of Mtb under hypoxia



DosR study summary	
Total number of proteins	53
Proteins with SRM assays	52
Proteins with validated assays in exponential and stationary growth phase	37
Proteins detected in DosR study	45



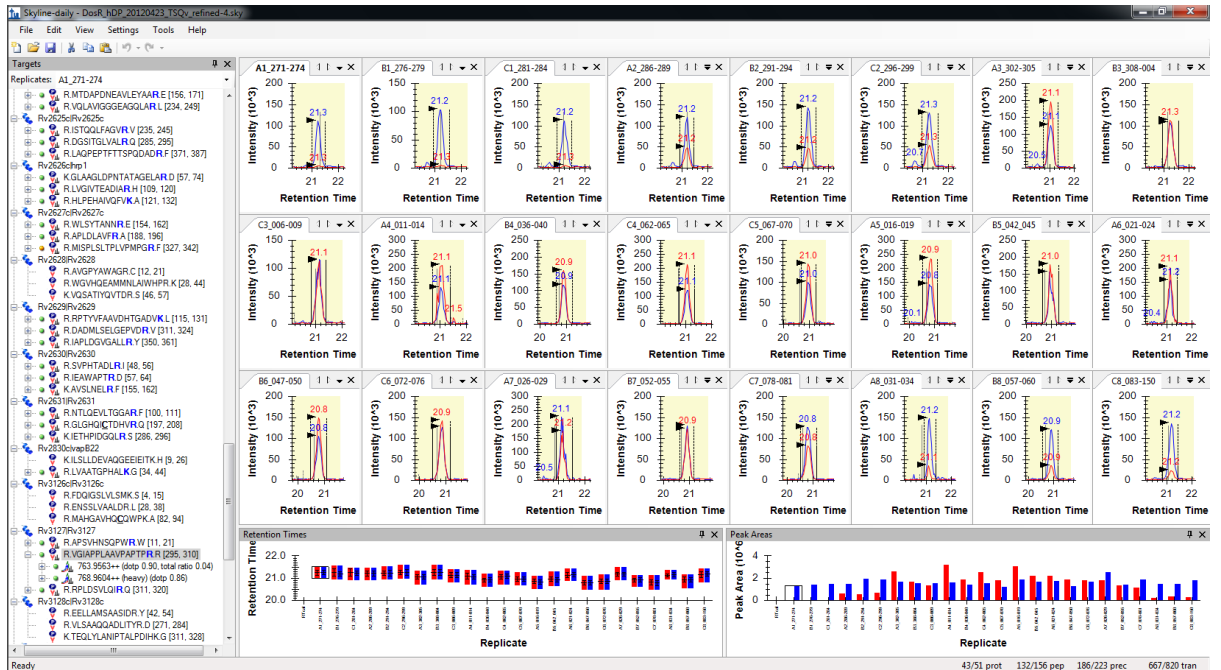
Application of the Mtb Proteome Library to study the dynamics of the DosR regulon of Mtb under hypoxia



- Rv1998c|Rv1998c
- Rv1735c|Rv1735c
- Rv2830c|vapB22
- Rv1812c|Rv1812c
- Rv2027c|dosT
- Rv0573c|pncB2
- Rv0570|nrdZ
- Rv1736c|narX
- Rv0082|Rv0082
- Rv2006|otsB1
- Rv3841|bfrB
- Rv0083|Rv0083
- Rv2630|Rv2630
- Rv2631|Rv2631
- Rv0571c|Rv0571c
- Rv2004c|Rv2004c
- Rv2005c|Rv2005c
- Rv0081|Rv0081
- Rv2003c|Rv2003c
- Rv1997|ctpF
- Rv2629|Rv2629
- Rv3132c|devS
- Rv3133c|devR
- Rv1813c|Rv1813c
- Rv2028c|Rv2028c
- Rv2624c|Rv2624c
- Rv1996|Rv1996
- Rv0572c|Rv0572c
- Rv0079|Rv0079
- Rv2625c|Rv2625c
- Rv3134c|Rv3134c
- Rv0080|Rv0080
- Rv2627c|Rv2627c
- Rv0569|Rv0569
- Rv2029c|pfbK
- Rv2030c|Rv2030c
- Rv2623|TB31.7
- Rv1738|Rv1738
- Rv3131|Rv3131
- Rv2032|acg
- Rv2007c|fdxA
- Rv3127|Rv3127
- Rv3130c|tgs1
- Rv2626c|hrp1
- Rv2031c|hspX

DosR study summary

Total number of proteins	53
Proteins with SRM assays	52
Proteins with validated assays in exponential and stationary growth phase	37
Proteins detected in DosR study	45



Statistical analysis of SRM data by SRMstats

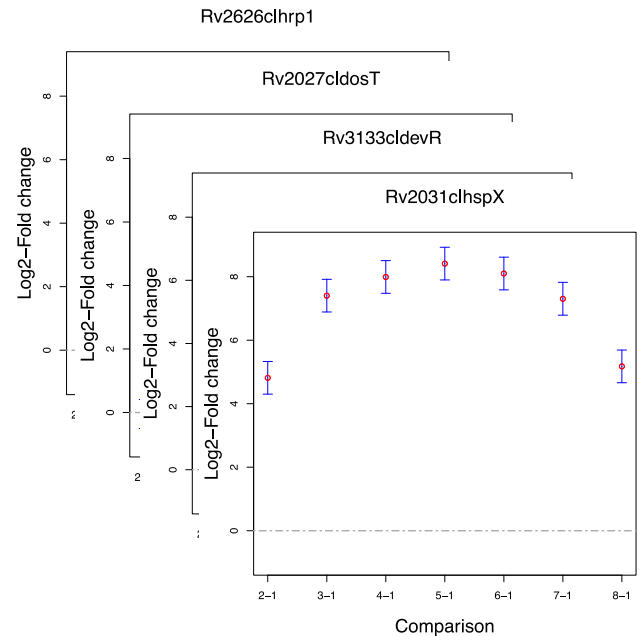
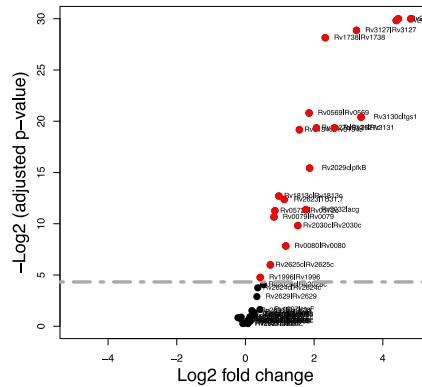
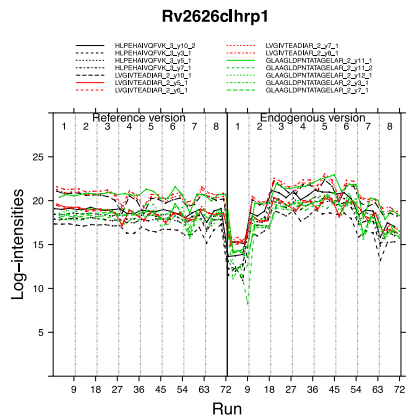


SRMstats

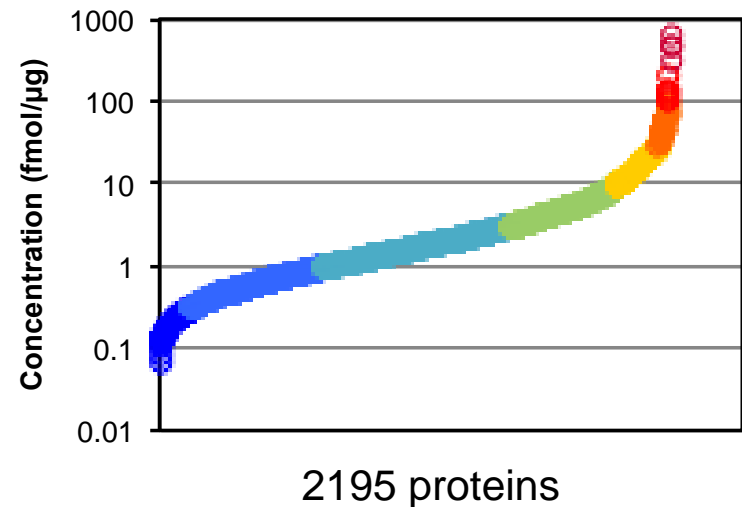
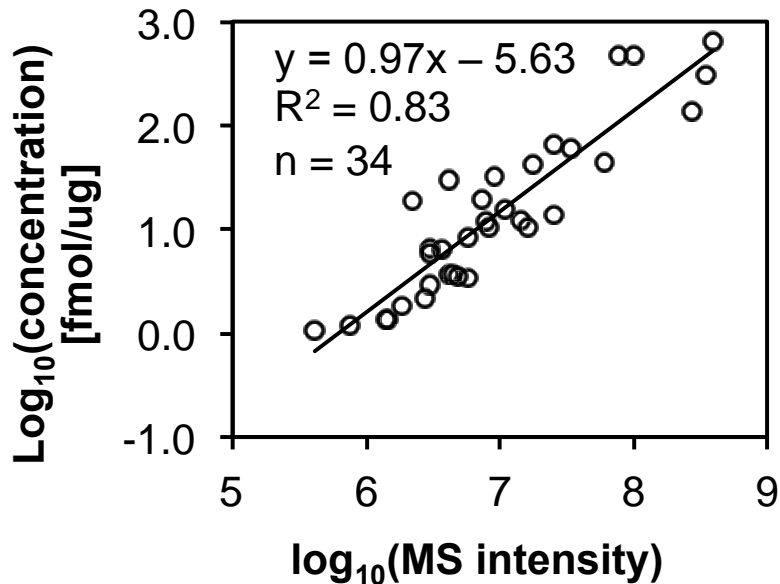
Protein significance analysis in SRM measurements

Goal

A statistical framework is proposed for protein quantification in SRM experiments based on a family of linear mixed-effects models. The framework is sensitive and flexible, and is applicable to a variety of experimental designs and to both label-based and label-free workflows.



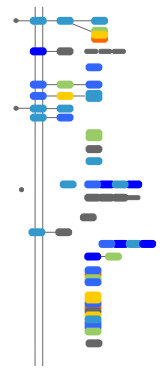
Absolute label-free quantification by SRM exploits the linear correlation of the sum of the top transitions of the top peptides per protein and the protein concentration



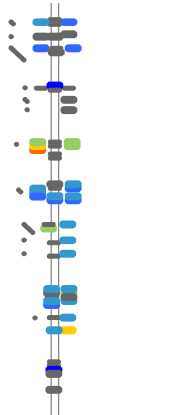
Linear correlation established using 34 anchor proteins quantified by AQUA peptides
MS intensity: sum of 2 most intense transitions of the 3 most intense peptides per protein

Proteome-wide absolute abundance estimates for Mtb

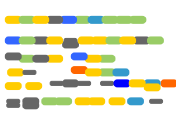
2-Component Systems



ABC Transporter



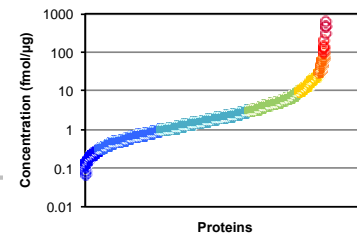
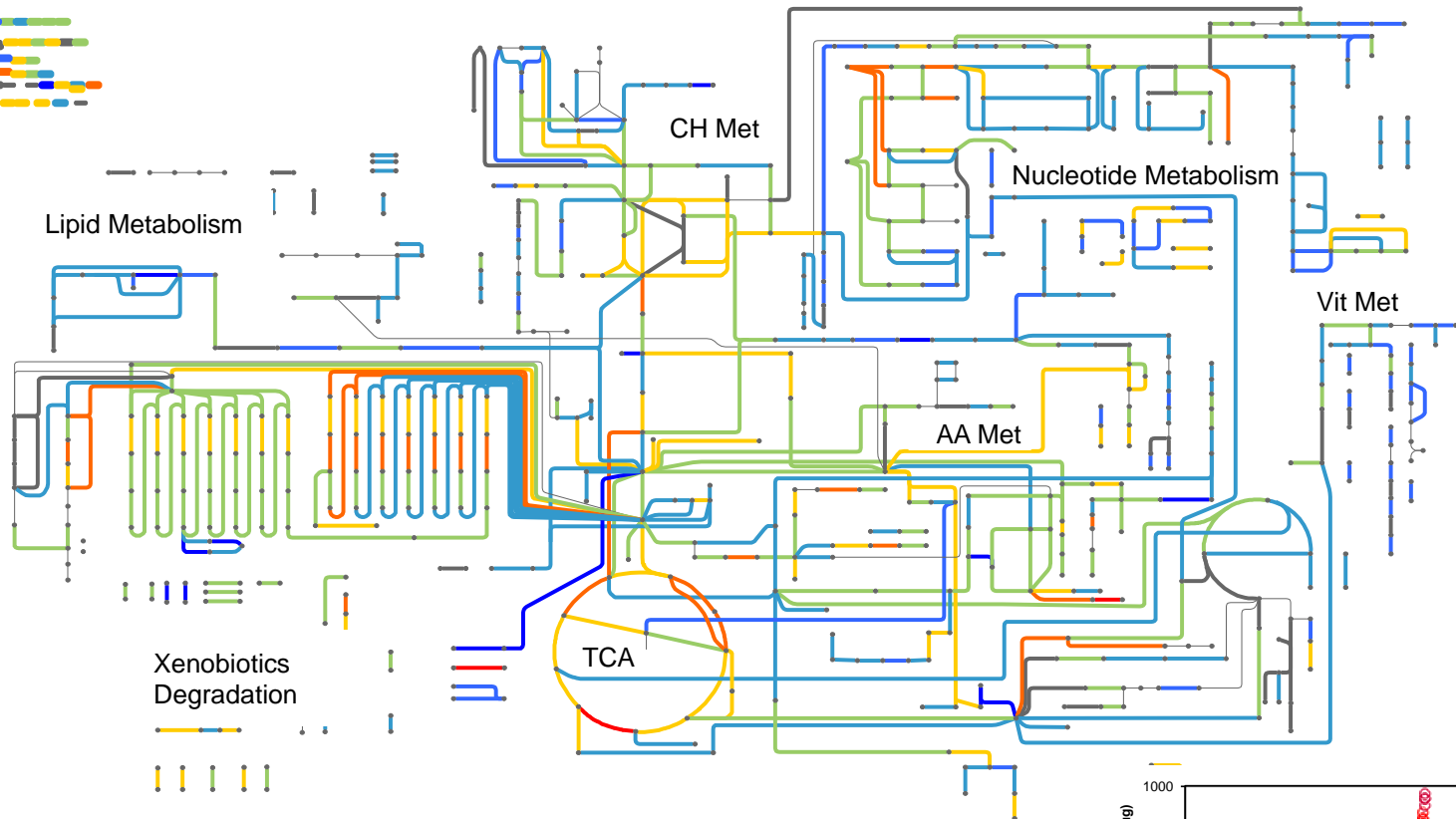
Ribosome



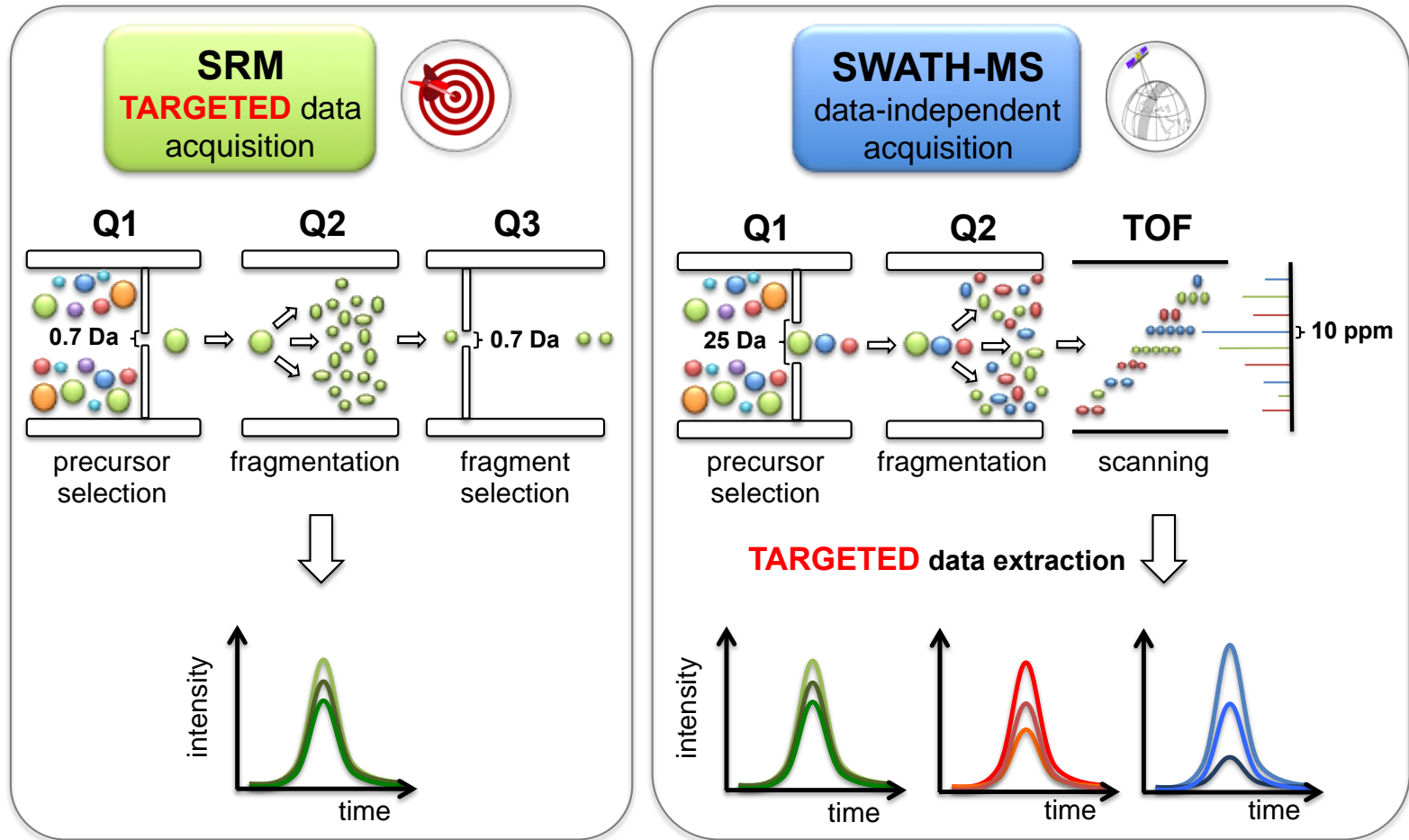
Bacterial Secretion



Xenobiotics Degradation



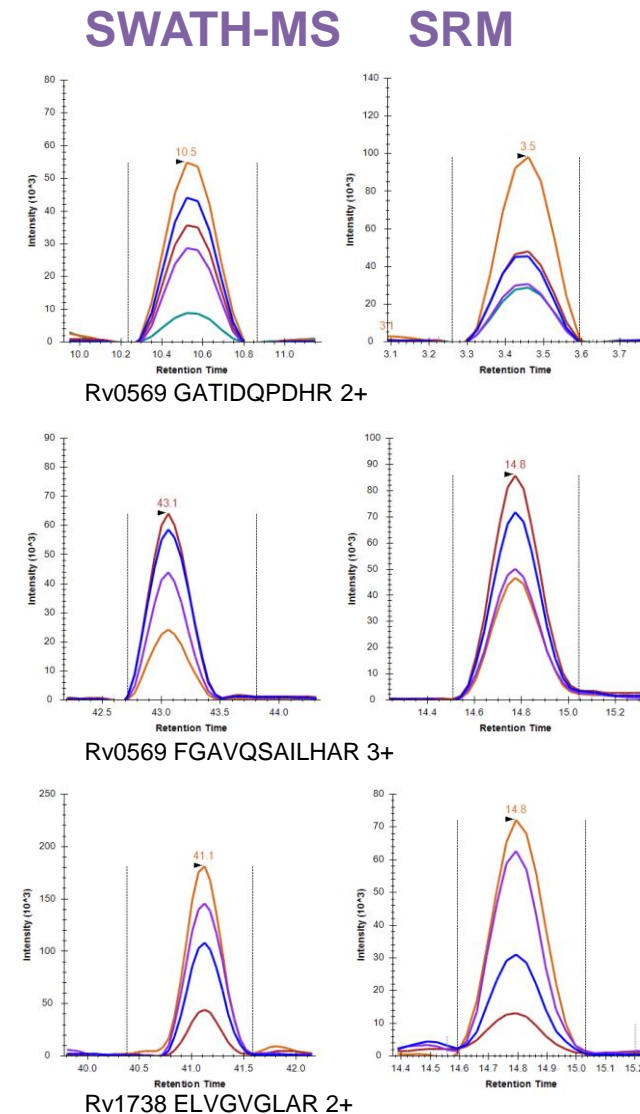
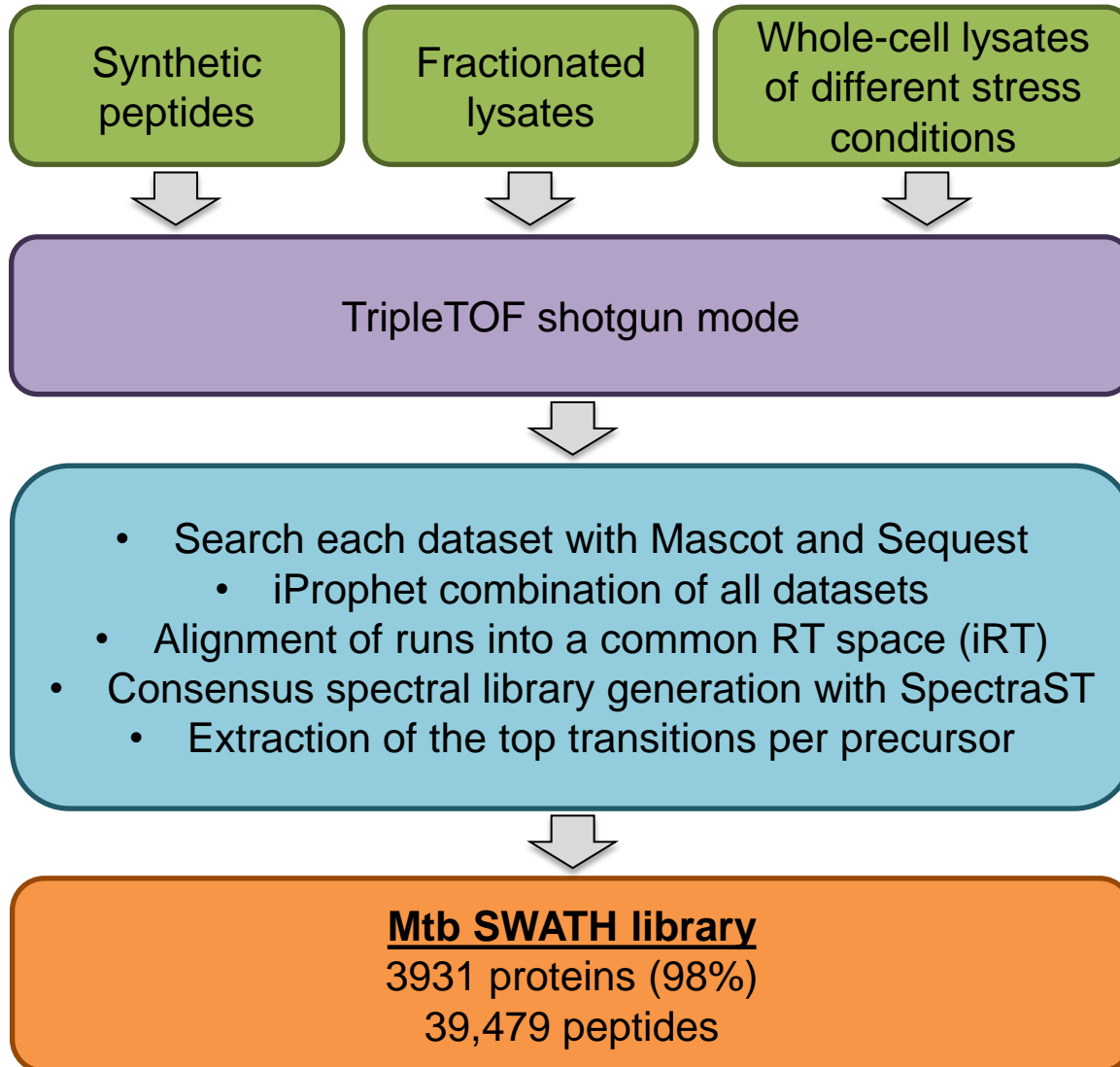
SWATH-MS: Data-independent acquisition with targeted data extraction



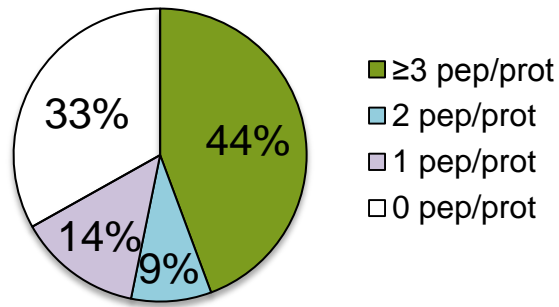
Gillet et al., MCP 2012, Targeted data extraction of the MS/MS spectra generated by data-independent acquisition: a new concept for consistent and accurate proteome analysis.

Figure by Christina Ludwig

Generation of the Mtb SWATH library



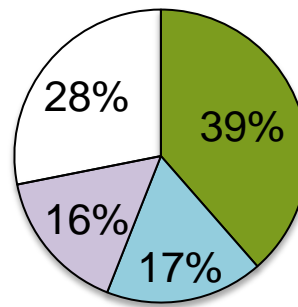
SWATH-MS allows reproducible, high proteome coverage measurements of Mtb in a single run



Proteome coverage by
SWATH-MS

single injection
2683 proteins (67%)

(openSWATH software)

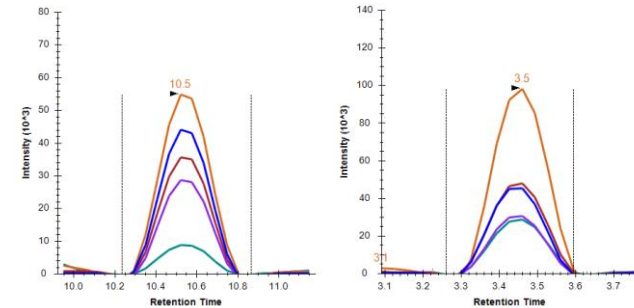


Proteome coverage by
SRM

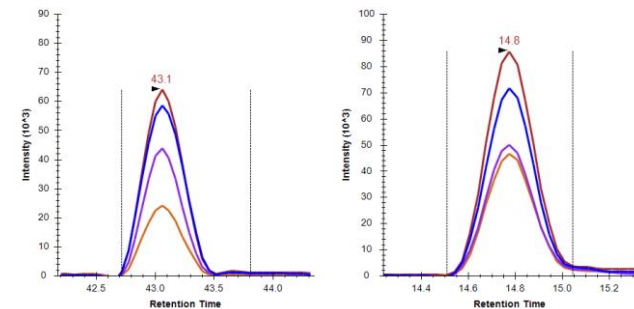
>200 injections
2884 proteins (72%)

(mProphet software)

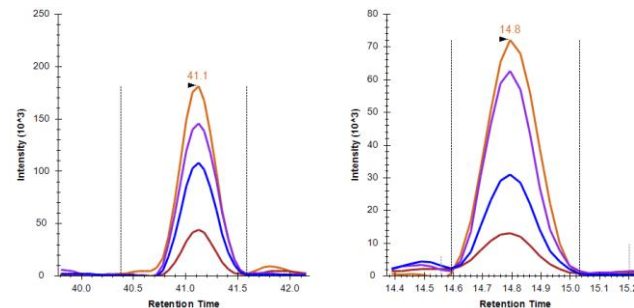
SWATH-MS SRM



Rv0569 GATIDQPDR 2+



Rv0569 FGAVQSAILHAR 3+



Rv1738 ELVGVGLAR 2+



Summary

- The Mtb Proteome Library is a public resource of SRM assays for the entire proteome of Mtb
 - SRM assays generated from crude synthetic peptides and validated in whole cell lysates
 - Data analysis with **Skyline** and supporting tools: iRT peptides, SRMCollider, mProphet, SRMstats
 - Data can be browsed on and downloaded from www.SRMAtlas.org, www.PeptideAtlas.org/passel
- Application of the Mtb Proteome Library to study the dynamics of the DosR regulon of Mtb under hypoxia
- Absolute label-free quantification by SRM exploits the linear correlation of the sum of the top transitions of the top peptides per protein and its absolute concentration.
- Expansion of the Mtb Proteome Library for use with SWATH-MS
- SWATH-MS allows reproducible, high proteome coverage measurements of Mtb in a single run

ETH Zurich

- Prof. Ruedi Aebersold
- Christina Ludwig
- Jeppe Mouritsen
- George Rosenberger
- Hannes Röst (Mon 5:45 pm and WP 595, Wed)
- Ludovic Gillet (TOD pm, Tue 5:30 pm)
- Ben Collins (WP 688, Wed)
- Alessio Maiolica, Mariette Matondo

University of Ghana

- Dr. Patrick K. Arthur

Max Planck Institute Berlin

- Prof. Stefan Kaufmann
- Dr. Martin Gengenbacher

ISB Seattle

- Prof. Rob Moritz
- Dave Campbell
- Zhi Sun
- Terry Farrah
- Samuel Bader (ABSciex, Mon 7 am)

University of Washington

- Brendan MacLean (TP 499, Tue)

Acknowledgements

