Protocol

Protocol for global proteome, virome, and metaproteome profiling of respiratory specimen (VTM) in COVID-19 patient by LC-MS/MS-based analysis



In this protocol, we describe global proteome profiling for the respiratory specimen of COVID-19 patients, patients suspected with COVID-19, and H1N1 patients. In this protocol, details for identifying host, viral, or bacterial proteome (Meta-proteome) are provided. Major steps of the protocol include virus inactivation, protein quantification and digestion, desalting of peptides, high-resolution mass spectrometry (HRMS)-based analysis, and downstream bioinformatics analysis.

Gaurav Tripathi, Nupur Sharma, Vasundhra Bindal, ..., Ekta Gupta, Jaswinder Singh Maras, Shiv Kumar Sarin

jassi2param@gmail.com (J.S.M.) shivsarin@gmail.com (S.K.S.)

Highlights

HRMS-based Proteome, Virome, and Metaproteome analysis of COVID 19 patients

Protocol for Sample preparation and Database curation for HRMS-based analysis

Downstream bioinformatics data analysis for Proteome, Virome, and Metaproteome dataset

Tripathi et al., STAR Protocols 3, 101045 March 18, 2022 © 2021 The Author(s). https://doi.org/10.1016/ j.xpro.2021.101045



Protocol



Protocol for global proteome, virome, and metaproteome profiling of respiratory specimen (VTM) in COVID-19 patient by LC-MS/MS-based analysis

Gaurav Tripathi,^{2,4} Nupur Sharma,² Vasundhra Bindal,² Manisha Yadav,² Babu Mathew,² Shvetank Sharma,² Ekta Gupta,³ Jaswinder Singh Maras,^{2,4,5,6,*} and Shiv Kumar Sarin^{1,*}

¹Department of Hepatology, Institute of Liver and Biliary Sciences, New Delhi, India ²Department of Molecular and Cellular Medicine, Institute of Liver and Biliary Sciences, New Delhi 110070, India ³Department of Virology, Institute of Liver and Biliary Sciences, New Delhi, India ⁴These authors contributed equally ⁵Technical contact ⁶Lead contact *Correspondence: jassi2param@gmail.com (J.S.M.), shivsarin@gmail.com (S.K.S.) https://doi.org/10.1016/j.xpro.2021.101045

SUMMARY

In this protocol, we describe global proteome profiling for the respiratory specimen of COVID-19 patients, patients suspected with COVID-19, and H1N1 patients. In this protocol, details for identifying host, viral, or bacterial proteome (Meta-proteome) are provided. Major steps of the protocol include virus inactivation, protein guantification and digestion, desalting of peptides, high-resolution mass spectrometry (HRMS)-based analysis, and downstream bioinformatics analysis. For complete details on the use and execution of this profile, please refer to Maras et al. (2021).

BEFORE YOU BEGIN

This protocol describes specific steps to execute proteome, metaproteome, and virome study in a COVID VTM sample (Maras et al., 2021). However, this protocol can be used for executing proteome, metaproteome, and virome from any sample. In the case of Plasma samples where some proteins are very high in concentration, e.g., albumin and immunoglobulins, you have to specifically deplete them first and then follow the same protocol {Chutipongtanate, 2017 #57}

Nasopharyngeal specimen (NP) collection (performed by a trained healthcare provider, only): (Pondaven-Letourmy et al., 2020)

Tilt head of patient backward around 70 degrees. Then Cautiously insert a mini tip swab through the nostril, parallel to the palate (not upwards) until you feel resistance or the distance is equivalent to that from the ear to the patient's nostril, indicating contact with the nasopharynx. Cautiously rub and roll the mini tip swab and leave the swab in place for several seconds to absorb secretions. Now gently remove the swab while rotating it. Finally, Place the mini tip swab into the transport tube (VTM media) provided.

Note: If a deviated septum or blockage creates difficulty obtaining the specimen from one nostril, use another nostril.

Oropharyngeal (OP) (throat) specimen collection

Insert the mini tip swab into the pharynx and tonsillar areas. Then roll swab over both tonsillar pillars and posterior oropharynx.







Note: Avoid touching the tongue, teeth, and gums.

Finally, Place the swab, tip first, into the viral transport media provided.

KEY RESOURCE TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER		
Chemicals, peptides, and recombinant proteins				
Ammonium bicarbonate	Sigma-Aldrich	A6141		
Sodium dodecyl sulfate	Sigma-Aldrich	L3771		
Dithiothreitol	Sigma-Aldrich	D0632		
lodoacetamide	Acros Organics	122270050		
Trypsin protease, MS grade	Thermo Fisher Scientific	90305		
Trifluoroacetic acid	Sigma-Aldrich	T6508		
Acetonitrile	JT Baker	14650359		
Formic acid	Sigma-Aldrich	F0507		
Urea	Sigma-Aldrich	U0631		
LC-MS grade water	Sigma-Aldrich	39253		
Bovine serum albumin	Thermo Fisher Scientific	23208		
Acetone	JT Baker	9006–03		
Acrylamide	Thermo Fisher Scientific	AM9022		
SDS	Thermo Fisher Scientific	15525017		
APS	Thermo Fisher Scientific	HC2005		
TEMED	Thermo Fisher Scientific	T9281		
Tris-Base	Thermo Fisher Scientific	17926		
HCL	Thermo Fisher Scientific	24308		
Glycine	Thermo Fisher Scientific	A37730IN		
Glacial acetic acid	Sigma-Aldrich	A0808		
Coomassie Brilliant Blue	Thermo Fisher Scientific	20278		
β-Mercaptoethanol	Sigma-Aldrich	M6250		
Bradford reagent	Thermo Fisher Scientific	22663		
Software and algorithms				
Proteome Discoverer 2.2 using Sequest HT	Thermo Fisher Scientific	OPTON-30945		
Xcalibur	Thermo Fisher Scientific	OPTON-30965		
Uniprot	(https://www.uniprot.org):			
Others				
Benchtop centrifuge	Eppendorf	5427R		
Sonicator	Helix Biosciences	HBSNII-92		
Vortex	Sigma-Aldrich	Z258423		
Incubator	Memmert	INB200		
Vacuum evaporator	Genevac	DNA-12060-C00		
Microplate reader	Thermo Fisher Scientific	VL0000D0		
Nano Reverse-phased capillary HPLC system	Thermo Fisher Scientific	UltiMate 3000 HPLC		
Q-exactive + orbitrap	Thermo Fisher Scientific	IQLAAEGAAPFADBMBCX		
Heated electrospray ionization (H-ESI) ion source	Thermo Fisher Scientific	H-ESI probe		
SDS PAGE Apparatus	Bio-Rad Laboratories	165-8000		
Viral Transfer media	RMBio	VTM-CHT		
Eppendorf tubes	Eppendorf	30125150		
ELISA plates	Thermo Fisher Scientific	44-2404-21		
Pierce C18 spin columns	Thermo Fisher Scientific	89870		
Water bath	LabcOM	76308-830		
Nano LC Column	Thermo Fisher Scientific	164570		
Calibration solution	Thermo Fisher Scientific	88323/88324		



MATERIALS AND EQUIPMENT

Tris's buffer		
Reagent	Final concentration	Volume
Tris's buffer	0.1 M Tris, pH 7.8	500 mL
Dissolve 6.05 g tris in LC-MS gra	de water and makeup to a total volume of 500 mL. Adjust pH t	o 7.8 with HCl.

Note: Tris buffer solutions can be stored at 25°C or at +4°C for two weeks.

Ammonium bicarbonate buffer		
Reagent	Final concentration	Volume
Ammonium bicarbonate buffer	50 mM ammonium bicarbonate in 0.1 M Tris buffer, pH 7.8	250 mL
Dissolve 0.97 g of ammonium bicard	ponate in tris buffer and make up a final volume of 250 mL.	

Note: Tris's buffer solutions can be stored at 25°C or at +4°C for two weeks.

Urea buffer		
Reagent	Final concentration	Volume
Urea buffer	8 M urea in 50 mM ammonium bicarbonate buffer	125 mL
Dissolve 110.06 g of ure	ea in 50 mM ammonium bicarbonate buffer and make up a final volume of 125 mL.	

Note: Urea solutions should always be freshly prepared and should be stored below 30°C.

Iodoacetamide solution			
Reagent	Final concentration	Volume	
lodoacetamide solution	50 mM iodoacetamide in urea buffer	10 mL	
Dissolve 92 mg of iodoacetamide in u	rea buffer and makeup to a total volume of 10 mL with 8 M u	ırea buffer.	

△ CRITICAL: Iodoacetamide acts as an acutely toxic, irritant, and health hazard, according to MSDS information. To prevent exposure, wear suitable protective gloves, eye protection, face masks, etc.

Note: Iodoacetamide is unstable and light-sensitive. Prepare solutions immediately before use and perform alkylation in the dark.

Dithiothreitol solution			
Reagent	Final concentration	Volume	
Sample buffer	1 M DTT in LC-MS grade water	10 mL	
Dissolve 1.55 gm of DTT in LC-N	1S grade water and makeup to a total volume of 10 mL.		

Note: Dithiothreitol solution is stored at -20° C and should use within 3 months.

LC /MS-MS – Solvent A		
Reagent	Final concentration	Volume
LC - Solvent A	0.1% (v/v) formic acid in LC-MS grade water	1,000 mL
Gently mix 1 mL formic aci	d with LC-MS grade water and makeup to a total volume of 1,000 mL w	vith LC-MS grade water.





Note: Always prepare fresh.

LC/MS-MS – Solver	nt B	
Reagent	Final concentration	Volume
LC - Solvent B	95% (v/v) acetonitrile, 0.1% (v/v) formic acid in LC-MS grade water	1,000 mL
Gently mix 950 mL o	f acetonitrile with 0.1 mL of formic acid and makeup to a total volume of 1,000 m	L with LC-MS grade
water.		

Note: Always prepare fresh.

Solution A for peptide purification (Wash buffer)		
Reagent	Final concentration	Volume
Wash buffer	50% (v/v) acetonitrile in LC-MS grade water	5 mL
In 25 mL of acetonitrile, a	dd 25 mL of LC-MS grade water to make the final volume 50 mL.	

Note: Always prepare fresh.

Solution B for peptide purification (Equilibration buffer)		
Reagent	Final concentration	Volume
Equilibration buffer	5% (v/v) acetonitrile in LC-MS grade water	50 mL
n 2.5 mL of acetonitrile, add 47.5 mL of LC-MS grade water to make the final volume 50 mL.		

Note: Always prepare fresh.

Solution C for peptide purification (Elution buffer)		
Reagent	Final concentration	Volume
Elution buffer	90% (v/v) acetonitrile in LC-MS grade water	5 mL
In 45 mL of acetonitrile, ac	ld 5 mL of LC-MS grade water to make the final volume 50 mL.	

Note: Always prepare fresh.

Tris's HCL buffer (PH-8.	8)	
Reagent	Final concentration	Volume
Tris's base	1.5 M Tris's base in deionized water.	100 mL
Dissolve 18.15 gm of Tris volume to 100 mL.	base in 80 mL of deionized water, now adjust the pH to 8.8 using 6 N $$	HCL and then adjust final

Note: Tris buffer solutions can be stored at 25°C or at +4°C for two weeks.

Tris's buffer (pH-6.8)		
Reagent	Final concentration	Volume
Tris's base	0.5 M Tris's base in deionized water.	100 mL
Dissolve 6 gm of Tris base in 60 r 100 mL.	nL of deionized water, now adjust the pH to 6.8 using 6 N HCL and	d then adjust final volume to



Note: Tris's buffer solutions can be stored at 25°C or at +4°C for two weeks.

10% (W/V) SDS		
Reagent	Final concentration	Volume
SDS	10% SDS in deionized water.	100 mL
Dissolve 10 gm of SDS in 9	0 mL of deionized water, then adjust the final volume to 100 mL.	

Note: Store 10% SDS stock solution at 20°C. Do not store in the fridge, as the SDS will reprecipitate.

10% (W/V) APS		
Reagent	Final concentration	Volume
APS	10% APS in deionized water.	1 mL
Dissolve 0.1 gm of SDS in 1 mL of deionized	water.	

Note: Always prepare fresh.

Tris-Glycine buffer (Running	buffer)	
Reagent	Final concentration	Volume
Tris's base	250 mM	
Glycine	1.92 M	
SDS	1%	
water	to 1 L	1 L
Dissolve 30.30 gm tris, 144.10 g	m Glycine, and SDS 10 gm in 900 mL of water and make fina	l volume 1 L.

Note: don't adjust it will be reached automatically pH 8.3.

Note: Tris buffer solutions can be stored at 25°C or at +4°C for two weeks.

1% Bromophenol Blue		
Reagent	Final concentration	Volume
Bromophenol blue	1% bromophenol in water.	1 mL
Dissolve 100 mg bromophenol Blu	e in 10 mL of water.	

Note: 1% Bromophenol Blue can be stored at 25°C or at +4°C for two years.

Staining solution		
Reagent	Final concentration	Volume
Coomassie blue R250	0.1% (w/v),	
methanol	20% (v/v)	
Acetic acid	10% (v/v)	

Dissolve 0.4 gm Coomassie blue in 200 mL of 40% (v/v) HPLC grade methanol in water with stirring as required. Filter the solution to remove any insoluble material. Add 200 mL of 20% (v/v) acetic acid in water.





Note: Always prepare fresh.

Destaining solution		
Reagent	Final concentration	Volume
methanol	50% (v/v)	
Acetic acid	10% (v/v)	
water	to 1 L	1 L

Add 500 mL of HPLC- grade methanol to 300 mL of HPLC-grade water, Add 100 mL of reagent grade acetic acid mix and adjust final volume to 1000 mL with water. The final concentrations will be 50% (v/v) methanol in water with 10% (v/v) acetic acid.

Note: Always prepare fresh.

STEP-BY-STEP METHOD DETAILS Heat inactivation of virus (Batejat et al., 2021)

© Timing: 30 min

Note: By the end of this step, the virus gets inactivated, and now all the remaining sample processing steps can be performed in the BSL-2 facility.

CAUTION: SARS-CoV-2 Samples should be handled in a Biosafety Level 3 (BSL-3) laboratory using BSL-3 practices till virus inactivation is performed and post inactivation, all other steps can be performed on BSL-2 facility.

- 1. Nasopharyngeal samples (NPS) and oropharyngeal samples together were collected into 1 mL of viral transport media.
- 2. Around 500 uL of the sample was incubated with a pre-warmed dry heating block maintained at 92°C for 15 min.

Homogenization of respiratory specimen (Dowling et al., 2020)

© Timing: 5–7 min per sample

Note: By the end of this step, all the cell's macromolecules like proteins, metabolites, lipids, etc., will be released in solution.

3. Respiratory specimens may contain host cells, mucus, bacterial and viral components. Thus, it is essential to homogenize the samples.

CAUTION: The high-frequency sound emitted by the Sonicator can damage hearing; therefore, place the sample in a noise isolating chamber and always close the door while operating. Do not grasp an activated horn or touch the tip of a vibrating probe. It can cause severe tissue damage and burns.

- 4. Take 500 μ L of respiratory specimen in a new microcentrifuge tube (MCT) and keep it on the ice during homogenization.
- 5. Place the sample in a noise isolating chamber and submerge the Sonicator probe into the sample.

Note: The probe should not touch the walls of MCT as it will break the tube and destroy the sample.



- 6. Close the door of the chamber after properly placing the sample tube.
- 7. Run the program at Power 20%, Run time 5 min (Cycle 10 s ON, 10 s OFF) and temperature 22°C.
- 8. Remove and wipe the probe with ethanol.

Note: Keep the sample on the ice during all the homogenization steps step 2-6.

III Pause Point: Respiratory specimen can be stored at -80° C for up to 4 weeks.

Organic solvent-mediated protein precipitation (Bradford, 1976; Simpson and Beynon, 2010)

© Timing: 90 min

Note: By the end of this step, all proteins will be precipitated.

- 9. 200 uL of VTM was placed in acetone compatible tube.
- 10. Add the 6-fold volume of cold acetone to VTM.
- 11. Vortex and incubate for 1-4 h at -80° C.
- 12. Centrifuge for 10 min at 18,000 g.
- 13. Carefully decant the supernatant without disturbing the protein pellet.
- 14. Allow pellet to air dry for 10–30 min in an uncapped tube. (Do not over dry otherwise, the pellet may not resuspend properly)
- 15. Resuspend the pellet in 50 uL of Ammonium Bicarbonate Buffer (ABC buffer).

Protein estimation using Bradford assay (Bradford, 1976)

© Timing: 30 min

Note: By the end of this step, the protein concentration of all samples can be quantified.

- 16. To prepare BSA standards and associated calibration curve, make a stock solution of 2 mg/mL of BSA in PBS.
- 17. Now label vials A to I, and add the stock or the master solution and Buffer as directed in the table below to achieve the desired results concentration of standard.

	Volume of diluent	Volume and source vial of BSA	Final concentration
Vial	(μL)	(μL)	(μg/μL)
A	0	300 of stock	2000
В	125	375 of stock	1500
с	325	325 of stock	1000
D	175	175 of vial B dilution	750
E	325	325 of vial C dilution	500
F	325	325 of vial E dilution	250
G	325	325 of vial F dilution	125
н	400	100 of vial G dilution	25
1	400	0	0

18. In a 96 well plate, 5 uL of pre-prepared BSA standards along with blank were added.

- 19. 5 uL of unknown samples was also loaded.
- 20. Now to each well, 245 uL of Bradford reagent was added.
- 21. Keep it on a shaker for 15-30 s.
- 22. Now incubate at 25°C for 10 min in the dark.





- 23. Take absorbance at 595 nm within 60 min post-incubation.
- 24. Plot a standard absorbance curve at 595 nm on the "Y" axis versus the protein concentration on the "X" axis.
- 25. Record the value "x" of unknown samples from the graph corresponding to absorbance reading and determine the protein concentration.

SDS PAGE mediated intra-sample variability detection (Nowakowski et al., 2014)

⁽¹⁾ Timing: 5 h

Note: By the end of this step, the intrasample variability among samples and in their associated group can be detected.

- 26. In-gel casting apparatus assemble glass plates and spacers.
- 27. Mix the component of resolving gel (12%).
- 28. Pour the resolving gel mixture between the gel plates till a level 2 cm below the top of the shorter plate.
- 29. Pace a layer of Distilled H2O over the top of the resolving gel to prevent meniscus formation in the resolving gel.
- 30. Incubate resolving gel for 30 min at 25°C
- 31. Drain the water from the top of the resolving gel and wick any remaining water away with a Kimwipe.
- 32. Mix stacking gel components and pour into gel plates (on top of resolving gel) to fill the gel plates.
- 33. Insert comb at the top of the spacers.
- 34. Remove the comb and assemble cast gel into the SDS apparatus.
- 35. Add 1× bromophenol blue to each sample.
- 36. Add freshly prepared $1 \times$ running buffer to both chambers of the apparatus
- 37. Now load DNA ladder in first well, 5 ug BSA in another well, and 30 ug equivalent unknown sample protein in other wells.
- 38. Run the gel at 100 V till the dye front migrates into the running gel (\sim 15 min) and then increase to 120 V until the dye front reaches the bottom of the gel (\sim 1.5 h.).
- 39. Now Remove the gel from the apparatus and by removing the spacers and glass plates.
- 40. Place the gel into a small tray.

Note: Never use a metal spatula to separate the glass plates.

- 41. Add \sim 25 mL staining solution and incubate for > 30 min with gentle shaking.
- 42. Add \sim 5 mL destaining solution and destain for \sim 1 min with gentle shaking, then Pour off and discard the destain solution.
- 43. Further, add \sim 30 mL of destain solution and destain with gentle shaking until the gel is visibly distained (> 2 h).
- 44. Pour off and discard the destain solution and rewash the gel with distilled water.
- 45. Visualize on GelDoc.

In-solution protein digestion (Gundry et al., 2009)

© Timing: 26 h

Note: By the end of this step, all the proteins will be digested, and tryptic peptides will be generated.

46. 50 ug equivalent proteins are diluted using the buffer till the final volume becomes 100 uL.



- 47. Add 20 uL of 10 mM Dithiothreitol (DTT) and incubate at 60°C for 1 h in a water bath to reduce proteins.
- 48. Post-reduction, add 15 uL of 10 mM Iodoacetamide and incubate for 30 min in the dark for alkylation of proteins,
- 49. Add 5 uL of modified sequencing grade trypsin (1 ug/5 uL) reconstituted in 100 uL ABC buffer.
- 50. Incubate at 37°C for 20–24 h in a water bath.
- 51. After incubation, add 5 uL of 0.1% (v/v) Formic acid to the reaction mixture to inhibit trypsin activity.

C18 column mediated desaltation of peptides (Dowling et al., 2020)

© Timing: 90 min

Note: By the end of this step, all the contaminants, e.g., salt, will be removed.

- 52. Please open a new Pierce C18 column, place it in an empty 1.5 mL Eppendorf tube,
- 53. Incubate it with 250 uL of Buffer C for 5 min (Column Washing).
- 54. Centrifuge for 2 min on a tabletop centrifuge and discard the flowthrough.
- 55. Add 200 uL of Buffer A in the column, then incubate for 2 min (Column Washing).
- 56. Centrifuge for 2 min on a tabletop centrifuge and discard the flowthrough.
- 57. Add 200 uL of Buffer B in the column,
- 58. Incubate for 2 min (Column equilibration).
- 59. Discard the flowthrough.
- 60. Now load the digested sample.
- 61. Incubate it for 5 min.
- 62. Centrifuge for 2 min on a tabletop centrifuge and then reload the flowthrough (peptide-binding).
- 63. Incubate again for 5 min and repeat step 49.
- 64. Repeat steps 49 and 50 two more times.
- 65. Centrifuge for 2 min on a tabletop centrifuge and discard the flowthrough.
- 66. Add 200 uL of Buffer B in the column, then incubate for 2 min (Washing).
- 67. Centrifuge for 2 min on a tabletop centrifuge and discard the flowthrough.
- 68. Add 30 uL of Buffer C and Centrifuge for 2 min on a tabletop centrifuge.
- 69. Repeat step 55.
- 70. Finally, add 40 uL Buffer C and Centrifuge for 2 min on a tabletop centrifuge.
- 71. Take the flowthrough and discard the used column.
- 72. Lyophilize the samples at 4°C and 60 millibar pressure till the sample is dry (do not over dry sample; otherwise, they will be tough to reconstitute).
- 73. Reconstitute in 40 uL of 0.1% (v/v) Formic acid.
- 74. Centrifuge at 15,000 g for and load the supernatant in an HPLC vial.

LC-MS/MS analysis (Bhat et al., 2020; Das et al., 2017, 2019; Dowling et al., 2020)

© Timing: 180 min per sample

Note: By the end of this step, the MS.raw files would be curated.

- 75. The peptides were ionized by nano-electrospray and subsequent tandem mass spectrometry (MS/MS) on a Q-ExactiveTM Plus (Thermo Fisher Scientific, San Jose, CA, United States).
- 76. The peptides were eluted by a 3–95% gradient of buffer B (aqueous 80% acetonitrile in 0.1% formic acid) with a continuous flow rate of 300 nL/min for about 60 min on a 25-cm analytical C18 column (C18, 3 mm, 100 A).



Protocol



Figure 1. Overview of data curation process at UNIPROT website

- 77. The peptides were analyzed using a mass spectrometer with the collision-induced dissociation mode with the electrospray voltage 2.3 kV.
- 78. Analysis on the orbitrap was performed with full scan MS spectra with a resolution of 70,000 from m/z 350 to 1800. Mascot algorithm (Mascot 2.4, Matrix Science).
- 79. Significant proteins were identified at (p < 0.05) and q values (p < 0.05). The threshold of false discovery rate was kept at 0.01.

Database curation for host proteome, metaproteome, and virome (Maras et al., 2021)

Note: By the end of this step, the database would be curated on which MS.raw files would be mapped.

- 80. The raw files generated post-LC-MS/MS analysis were analyzed using Proteome discoverer 2.2 for protein identification and quantification.
- 81. One of the most crucial parts is database curation, against which the mass spectrometric files need to be searched. So, in our case, we created three databases: a Virome (COVID) database, another for humans, and one for metaproteome.
- 82. To curate the database, visit https://www.uniprot.org/proteomes/.
- 83. Find proteome for your organism of interest by searching for them from proteome ID, e.g., UP000005640 Homo sapiens) or taxonomic name or ID.
- Now click on the download button> Select all protein entries > Fasta (canonical, compressed, and isoform). (Figure 1)

Note: the databases in Uniprot are of three categories reviewed (Uniprot/Swissport), Unreviewed (Uniprot/TREMBL), or both (UniprotKB) options. Always prefer the reviewed databases because reviewed databases are carefully annotated with minimal redundancy and can integrate into other databases efficiently.

Protocol



Thermo Proteome Discoverer 2.2.0.388		
le View Administration Tools Window	v Help	
💱 🞲 🗳 🔒 🎒		a ex 🛛 🗠 📄 👗 🐗 👼 🗸
Start Page X		
C .		
Drotoom	Discoverer 2.2	
	le Discoverer 2.2	
-		
Start	Recent Studies	Recent Results
New Study/Analysis	funcal analysis	E 20210513 ENV REPEAT
Open Study		
Open Result	HELA	B al 08-1
open neoditin	E HEK	hby poornima
	SORTED CELL PROTEOME HAMID SIR	hby blast iitoo
	🖬 zabardasti	E NUPUR6
	Charu 1	Hamed Proteomics
	H CHARU	a ABRASSI
	New Study	BRASSI-(1)
	+ jeeto sir final	hbv_blast_jitoore
	BRASSICA	ei jeeto 110221
	ieetu sir analysis pro	SALIVA PROTEOMICS DATABASE SEARCH RESULTS

Figure 2. Pictographic representation of step 87 associated to proteome discoverer 2.2 mediated data analysis section

85. We used the following databases

Application	Organism	Source
Host Proteome	Homo sapiens	Uniprot homo sapiens (Human) database (UP000005640; reviewed)
Virome	COVID-19	https://covid-19.uniprot.org/uniprote
Metaproteome	Amalgamation	bacterial/fungal sequence (UniprotSwP_20170609, with sequences 467231 and MG_BG_UPSP with sequences2019194)

Proteome discoverer 2.2 mediated data analysis (Maras et al., 2021)

Note: By the end of this step, the excel data files would be curated.

Note: The steps for data curation from raw files for Proteome, Metaproteome, and virome are the same, and the same raw files can be used for data curation analysis of all three.

The only difference among the three is that each need to be run individually against their respective database, e.g., Host proteome should be mapped against Human Uniprot database)

- 86. Open Proteome discoverer 2.2.
- 87. Click New Study. (Figure 2)
- 88. Name your study > Decide the root directory > Add your raw files > Press ok. (Figure 3)
- 89. Click on New Analysis. (Figure 4)
- 90. Click on Input files > Select all files and drag it below the processing workflow window > Click on the icon of the processing step. (Figure 5)
- 91. Click on Open common > Select your processing workflow and press enter. (Figure 6)
- 92. Click on Spectrum files and > go to Search settings > Select the Protein database, e.g., for COVID virome, select the already curated COVID database> Select the type of digestion we performed, e.g., tryptic peptides. (Figure 7)

Note: This is the place where you will select the database for proteome, metaproteome, and virome or any other organism associated.

93. Now, select each step of the Processing workflow.



Figure 3. Pictographic representation of step 88 associated to proteome discoverer 2.2 mediated data analysis section

Note: Setting for each parameter group is given in the image below. (Figure 8)

94. Now click on consensus workflow > then Open common > and select your workflow. (Figure 9) Step 10-Now select each step of the processing workflow.

Note: Setting for each parameter group is given in the image below. (Figure 10)

- 95. Click on grouping and quantification> Then select Sample type. (Figure 11)
- 96. Click on Run. (Figure 12)
- 97. Wait for 24-72 h. for data analysis to happen, and finally, you will get a file like below. (Figure 13)

Note: You have to repeat Proteome discoverer 2.2 mediated data analysis five times.

- a. For proteome run all raw files of COVID positive, COVID negative, and H1N1 samples together.
- b. For virome, similarly run all raw files of COVID positive, COVID negative, and H1N1 samples together.
- c. For the metaproteome, run three times for COVID positive, once for COVID negative, and then finally once for H1N1.)

File View Administration Tools Window Help
💱 🕼 😂 🛃 🎒 👚 🌱 🖾 🖂 📖 🐼 🔊 👘 🖼 🖽 📾
Start Page × Study: Covid * ×
Mark Add Files Add Fractions Kernove Files Open Containing Folder New Analysis Open Analysis Template Study Definition Input Files Samples Analysis Results
Study Summary
Study Name: Covid Study Directory: El.covid lipids analysis/Covid Last Changed: 9/1/2021 2:5:18 PM Creation Direct: 9/1/2021 2:5:18 PM

Figure 4. Pictographic representation of step 89 associated to proteome discoverer 2.2 mediated data analysis section

STAR Protocols Protocol



Protocol



🗌 Ås Bøtelt 🥥 Paul
🗌 Ås Banh 🥥 Pue

Figure 5. Pictographic representation of step 90 associated to Proteome discoverer 2.2 mediated data analysis section

Note: From here, the further downstream data processing is the same for Host proteome and virome and different for metaproteome.

Downstream data Processing for Metaproteome analysis (Maras et al., 2021)

Note: By the end of this step final metaproteome excel file will be curated.

- 98. By the end of step 13, you will have three files for the metaproteome, one for each group.
- 99. Open the excel file of one group. (Figure 14)
- 100. Copy all the tryptic peptides
- 101. Visit the Unipept website https://unipept.ugent.be/. (Figure 15)
- 102. Enter the list of tryptic peptides here > Click search (Figure 16).

File Ver Administration Tock Widder Hele Ver Administration Tock Wid	Thermo Proteome Discoverer 2.2.0.388	
Image: Image	File View Administration Tools Window Help	
Image: Part Part Image: Part Part Part Part Part Part Part Part		
Add Fale: Add Fale: Add Fale: Concourds Add Fale: Concourds A Canaditation Start Detailed: Concourds A Canaditation Add Fale: Add Fale: Concourds A Canaditation Add Fale: Add Fale: Concourds A Canaditation Add Fale: Concourds A Canaditation Concourds A Canaditation Detailed: Concourds A Canaditation Concourds A Canaditation Startum File: Concourds A Canaditation Concourds A Canaditation Manad Falabasia Concourds A Canaditation Concourds A Canaditation Startum File: Concourds	Start Page × Study: Covid ×	
Study Detailoring Ingust Falls Samples Analysiss Results Note State Detailoring Samples Analysiss Results Samples Samples State Data Samples Samples Samples State Data Samples <td< td=""><td>🙀 Add Files 🚜 Add Fractions 💥 Remove Files 🧔 Open Containing Folder 🛞 New Analysis 🍏</td><td>Open Analysis Template</td></td<>	🙀 Add Files 🚜 Add Fractions 💥 Remove Files 🧔 Open Containing Folder 🛞 New Analysis 🍏	Open Analysis Template
Deta book Sector Sec	Study Definition Input Files Samples Analysis Results Workflows Grouping & Quantification Workflow Nodes	n 👔 Open 📸 Open Common 🚠 Save 🌉 Save Common 🏂 Auto Layout 💢 Clear
	Data Input	Workflow: PWF QE Precursor Quan and LFQ SequentIHT MASCOT Percelator ILBS covid
B) Section Files CC Section Files CC Section Files Cector Monor Assume Titer Section Files Note Feature Files Section Files	🖗 Spectrum Files	Description: Processing workflow for once more based quartification, using Sequent/Twith Percelator validation. Specify the FASTA database labels used and any additional modifications.
Secture Retrieval Image: Secture Sector Peature Detector Machagement Risk Reporter Tiler Secture Transmitter Machagement Risk Secture Transmitter Massamed Massamed <t< td=""><td>翻 Spectrum Files RC</td><td></td></t<>	翻 Spectrum Files RC	
Peature Detection Routent Selector Peature Detector Monar Aquant Detector Monar Aquant Detector Monar Aquant Detector Recurssing Note Peak Filter Sectrum Filter Sectr	Spectrum Retrieval	
Peter Detection & Questification Minors Feature Detector Reporte Eards Questifier Secture Thire Sectur	Spectrum Selector	Workflow Tree
k Monsa Facue Detector Reporter loss Quantifier Reporter loss Quantifier Reporter loss Quantifier Reporter loss Quantifier Non-Facuers Filter Sectum Grouper Sectum Grouper Sectum Grouper Sectum Filter Sectum	Feature Detection & Quantification	· · · · · · · · · · · · · · · · · · ·
All Reporter Josa Quantifier Sectrum Processing Noise Resk Filter Noise Resk Filter Sectrum Konnelizer Sectrum	💱 Minora Feature Detector	Spectrum Files
Sectum Processing Noise Propriet Filter Sectum Couper Sectum Consultarie Sectum Control Couper Miss Amands Miss Amands Miss Amands Sectum Control Couper <tr< td=""><td>ul Reporter Ions Quantifier</td><td>RC RC 0</td></tr<>	ul Reporter Ions Quantifier	RC RC 0
 Note Pasteline State Sochraghent Filter Sochram Filter Sochram	Spectrum Processing	
WonFagnett Filter Sectrum Konsaltzer Sectrum Konsaltzer Sect	😡 Noise Peak Filter	
Partial kixed Spectrum Grouper Spectrum Morralizer Spectrum Morralizer Spectrum Rise Spectrum R	😡 Non-Fragment Filter	
Section Grouper	😡 Parallel Xtract	Spectrum 1 Partonic 6
	Spectrum Grouper	Selector
Top Peaks Filter Secture Triller Secture Confidence Filter Secture Confidence Filter Secture Database Search MS Amanda 2.0 MS Amanda 2.0 Mascot MS Amanda 2.0 Mascot MS Amanda 2.0 MS A	Spectrum Normalizer	
Secture Filter Secture Confidence Filter Secture Confidence Filter Secture Confidence Filter Secture Database Search MsAmada 2.0 Secture Library Search Secture Confidence Filter Secture Confidence Filter Mascet 4 MsAmada MsAmada MsAmada MsAmada MsAmada MsAmada MsAmada MsAmada MsAmada MsAmada MsAmada MsAmada MsAmada MsAmada MsAmada	😡 Top N Peaks Filter	
Start Event Filter Sequest H1 2 Masscot Secture Ontabase Search MS Amanda 2.0 Ms Amada 2.0 Ms Amada 2.0 Ms Amada 2.0 Pht-Preview Sequest H1 Sequest H2 MS Amada 2.0 Ms Amada 2.0 <	Spectrum Filters	
Bectum Confidence Filter Sectum Confidence Filter Percolator 3 Fixed Value B Fixed Value B Fixed Value B Fixed Value B Sectum Confidence Filter B Fixed Value B B B Fixed Value B B	Scan Event Filter	Mascol 4
Sequence Database Search MsAmada 2.0 Masce: Proview Sequence Search Masce: Proview Sequence Search Secretaria Search Secretaria Search Secretaria Search Secretaria Search MsAmada 2.0	Spectrum Confidence Filter	
Sequence Database Search d M SAmanda 2.0 M Marcata P MI-Preview Sepectral Library Search M Sepectral Sepectral Library Search M Sepectral	ur) Spectrum Properties Filter	
Al MS Amanda MS Amanda 2.0 MS Amanda 2.0 MS Amanda 2.0 MS Amanda 2.0 MS Amanda 2.0 MS Amanda 2.0 Percolator 3 MS Amanda 2.0 Percolator 3 MS Amanda 2.0 Percolator 3 MS Amanda 2.0 Percolator 3 MS Amanda 2.0 Percolator 1 MS Amanda 2.0 Percolator 1 Percolator 1 Percola	Sequence Database Search	
In Maximula 200 Image: Constraint of a constrai	MS Amanda	Percolator 3 Percolator 8
Partage P	Macat	
MSPeriod Spectral Library Search MSecond	PMI-Skypeir	
SequestHT Spectral Library Search MSPepSearch	PMI-Preview	
© Spectral Library Search ♀ MSPepSearch	Sequest HT	
K MSeptench	Spectral Library Search	
	MSPepSearch	
P PMValidation	PSM Validation	
E Food valuePSN validator	Eixed Value PSM Validator	
Percolator	Percolator	· · · · · · · · · · · · · · · · · · ·
Target Decoy PSM Validator	Target Decoy PSM Validator	¢ >

Figure 6. Pictographic representation of step 91 associated to proteome discoverer 2.2 mediated data analysis section



Protocol



Figure 7. Pictographic representation of step 92 associated to proteome discoverer 2.2 mediated data analysis section

- 103. You will get a sunburst plot from Unipept (Figure 17).
- 104. Now download the data from Unipept by clicking on the download tab. (Figure 18).
- 105. You will get a file like this (Figure 19).
- 106. Repeat these steps for the other two groups.

*					
Spectrum 1 Selector 1	💓 Sequ	iest HT	Mascot	4	Percolator 3
Show Advanced Parameters					
✓ 1. General Settings		1	Show Advanced Parameters		Show Advanced Parameters
Precursor Selection Use MS1 Precursor			1. Input Data		
2. Spectrum Properties Filter			Protein Database		✓ I. Input Data
Lower RT Limit 0	Show Advanced Parameters		Enzyme Name		Maximum Delta Cn 0.05
Opper KT Umt 0	V 1. Input Data		Maximum Missed Cleavage Sites	2	Maximum Paula 0
First Scan U	Protein Database	CoV2_NCBI_sequences	Instrument		Maximum Hank U
Last Scan U	Enzyme Name	Trypsin (Full)	Taxonomy		2. Decoy Database Search
I grove Specified Sci	Max. Missed Cleavage Sites	2	V 2. Tolerances		Treast EDD (Strict) 0.01
Highest Charge Stati 0	Mn. Peptide Length	6	Precursor Mass Tolerance	10 ppm	Target FDR (Strict) 0.01
Min. Practure Mass 350 Da	Max. Peptide Length	144	Fragment Mass Tolerance	0.02 Da	Target FDR (Relaxed) 0.05
Max. Precursor Mass 5000 Da	2. Tolerances		Use Average Precursor Mass	False	Validation based on avValue
Total Intensity Three 0	Precursor Mass Tolerance	10 ppm	✓ 3. Modification Groups		validation based on q-value
Minimum Peak Court 1	Fragment Mass Tolerance	0.02 ba	From Quan Method		
3. Scan Event Filters	Use Average Precursor Mass	False	4 Dynamic Modifications		
Mass Analyzer (Not specified)	V 3 Spectrum Matching	rase	Show Al Modifications	False	Eived Value
MS Order Is Not MS1	Use Neutral Loss a loos	True	1. Dynamic Modification		rixed value
Activation Type (Not specified)	Use Neutral Loss b loos	True	2. Dynamic Modification		DSM Validator
Min. Collision Energy 0	Use Neutral Loss y lons	True	3. Dynamic Modification		
Max. Collision Energ 1000	Use Ranking lons	True	4. Dynamic Modification		
Scan Type Is Full	Weight of a lons	0	5 Dynamic Modification		terrent sector s
Polarity Mode (Not specified)	Weight of b lons	1	6. Dynamic Modification		
✓ 4. Peak Filters	Weight of c lons	0	7. Dynamic Modification		
S/N Threshold (FT< 1.5	Weight of x lons	0	8 Dynamic Modification		Chan Advand Demosters
5. Replacements for Unrecognized Properties 1	Weight of y lons	1	9 Dunamic Modification		Show Advanced Parameters
Unrecognized Charg Automatic	Weight of z lons	0	× 5 Static Modifications		
Unrecognized Mass ITMS	 4. Dynamic Modifications 		1 Oute Modification		✓ 1. Input Data
Unrecognized MS O MS2	Max. Equal Modifications Per Peptide	3	2 Optio Modification		
Unrecognized Activa CID	1. Dynamic Modification	Oxidation / +15.995 Da	2 Outo Modification		Maximum Delta Cn 0.05
Unrecognized Polart +	2. Dynamic Modification	None	A Dutic Medication		No. Data A
Unrecognized MS R 60000	3. Dynamic Modification	None	 Code Modeledon Date Madiantes 		Maximum Rank 0
Unrecognized MSn 30000	4. Dynamic Modification	None	5. Static Modification		
1	5. Dynamic Modification	None	6. Static Modification		

Figure 8. Pictographic representation of step 93 associated to proteome discoverer 2.2 mediated data analysis section

Protocol





Figure 9. Pictographic representation of step 94 associated to proteome discoverer 2.2 mediated data analysis section

- 107. Now, you will have this excel file for all groups.
- 108. Now, compile all three files and master peptide files to get the abundance of each bacterium (Figure 20).
- 109. During this process, you can identify each disease group-specific metaproteome and compile the relative abundance of metaproteome.
- 110. Now use this file for statistical analysis

Note: the files generated at step 13 of proteome and virome can also be used for statistical analysis.



Figure 10. Pictographic representation of step 94 associated to proteome discoverer 2.2 mediated data analysis section



Protocol



Figure 11. Pictographic representation of step 95 associated to proteome discoverer 2.2 mediated data analysis section

EXPECTED OUTCOMES

The importance of Global Proteomics in SARS-CoV-2 biology can help us understand the virus-host protein interaction and provide the index of its variability depending on the cell type used or sample type studied. Together, this can help us understand the mechanism of virus infection and provide us with potential biomarkers. On the contrary SARS-CoV-2 infection often precedes with bacterial co-infection and is linked with longer duration and more severe infection (van den Brand et al., 2014) thus metaproteome studies can give us insights into how the COVID-19 infection changes the microbiome of the human body and is this change in microflora can be associated to the severity of the disease. This protocol describes how to execute global proteomics to identify host proteome, virus-linked proteome (Virome), and bacteria-linked proteome (metaproteome) in the respiratory specimen of the study group. This protocol describes an approach that combines the Sample collection, virus inactivation, protein isolation, alkylation, digestion, desalination, and peptide sequencing using a Q-exactive + mass spectrometer bioinformatics approaches for data analysis like database curation, using proteome discoverer 2.2 and metaproteome analysis.

The expected outcome of these studies would be a list of a few hundred to several thousands of proteins and their expression pattern. Now via the proteomics approach, you can understand the change in host proteome pattern, virome can reveal the insights of virus proteins in the host, and finally, the metaproteome study can reveal the change in it and intensity of change. For example, in our COVID study (Maras et al., 2021), we identified 6 significantly differentially expressed COVID-19 viral proteins; around 1256 host proteins were identified. We also found an increase in respiratory pathogens (e.g., Burkholderiales, Klebsiella pneumonia) and decreased lactobacillus salivary (FDR<0.05) in the COVID-19 specimen. Finally, an increase in the basal level of WARS (tryptophan-tRNA ligase) and MX1 (MX dynamin-like GTPase 1) was correlated with SARS-CoV-2 infection outcome. Further, these were validated in a separate cohort of 200 patients.



			_	ð	\times
					• 4 Þ
		r			
nalysis		As Batch	🎲 Run	Save	×
Consensus Ste	р 💫				×
Workflow: C	WF_Comprehensive	e_Enhanced			
A F	nnotation_LFQ_and _11042020	_Precursor_Quan_IL	.BS.pdCo	onsensus	W
Result File: a	LO8-1.pdResult				
Child Steps	(1)			L	Add

Figure 12. Pictographic representation of step 96 associated to proteome discoverer 2.2 mediated data analysis section

QUANTIFICATION AND STATISTICAL ANALYSIS

Systems bioinformatics

() Timing: days to weeks

For further downstream analysis of proteome, virome, and Metaproteome data is performed using the following software's-

Thermo Proteome Discoverer 2.2.0.388												-	o ×
File View Administration Tools Window Help													
🕼 🕼 ڬ 🔒 🎒 👫 📖 💎 🐼 💷 🗔 🗖	Lalu L		-		🗟 🐻 🖪								
Start Page X Study: Covid X Administration X hby poornima X Hamed Proteomics X													- d b
	•			- N Observer									
Protein Groups Peptide Groups Pams Marma Spectrum into Input Piles	a special	ized traces Consi	ensus reatures Ret	suit Statistics		0	land the second second			100 x1 0		-	
Checked Protein FD + Master Accession Description	Exp. q-valu [Coverage [%]	Sequence Coverage	# Peptides # P	'SMs # Unique Peptides #	Protein Groups # AA	s MW [kDa] calc.	pl Score N + [4	Score S • •	# Peptdes	Biological Process	+	Cellular Compc
1 0 High V P50/09 Actin, cytoplasmic 1 OS=Homo sapiens 0X=9506 GN=AC	0.000	87%		35	827 14	1 3/5	41.7 5.4	8 8200	659.01	29 34			
P02/68 Serum albumin US+Homo sapiens UX+9606 GN+ALB PE	0.000	75%		56	96/ 56	1 603	69.3 6.4	8 8982	581.20	42 56			
3 4 P4332/ Fatty acid synthese US=Homo septens UX=3606 UN=FAS	0.000	56%		96	564 35	1 251	2/3.3 6.4	4 51/1	535.97	P3 3P			
 High P 10803 Bu KLA Heat shock protein, mitochondrial US+Homo saper 	0.000	/9%		43	4/4 43	1 57.	61.0 5.0	/ 623/	432.60	33 42			
Private knase PKM US=Homo sapens UX=9606 UN=PH Private knase PKM US=Homo sapens UX=9606 UN=PH	0.000	/3%		41	408 41	1 53	57.3 7.8	4 4609	4///43	35 41		_	
High V P0/43/ Tubulan beta chain US+Homo sapiens UX+9505 GN=108	6 0.000	81%		28	503 5	1 44	49.6 4.2	3 5452	462.23	2/ 28			
7 C High V PU8238 Heat shock protein HSP 90-beta US+Homo sapiens UX+9	6 0.000	65%		44	526 27	1 72	83.2 5.0	3 5831	446.//	35 41		_	
Reginary Polyson Reat shock protein HSP Sol-apina OS-Homo sapiens OX-	0.000	61%		**	404 30	1 73	04.0 0.0	2 5001	440.00	91 90		_	
9 9 P683/1 Tubulin beta-45 chain US+Homo sapiens UX+9606 GH+1	0.000	81%		23	409 1	1 44	43.0 4.0	3 5143	420.03	2/ 23			
11 c Right V PO4406 Glyceradenyde/Sphotphate denydrogenase OS+Romo s	a 0.000	63%		52	420 61	1 33	004 45	4 2042	415.10	25 52			
12 II III III III IIII V P14625 Endoplasmin OS-Homo septens OX-Sole Git-InSP3081	T 0.000	00%		23	433 51	1 00.	40.0 51	- 3343	900.92	30 02			
12 W QBQCS Tobula april-10 chair 05 mono sepiens 0X 8500 GN	0.000	12%		31	413 2	1 44	43.3 5.1	0 4430	304.24	20 31			
13 - High V Q13885 Tobula beta-2A chain OS=Homo saplens OX=9606 GN=1	0.000	12%		21	412 1	1 44	43.3 4.0	9 4002	3/0.45	24 27			
14 W USBVAT Tubula beta-26 chain US#homo sapiens OX#8506 GN#1	t 0.000	81%		20	910 1	1 44	43.3 4.2	a 4020	3/5.29	23 28			
15 - Hon V Q/1036 Tubulin apris-14 chain US=Homo sapiens UX=9606 GN=	0.000	12%		30	3/1 1	1 45	50.1 5.0	6 4058	367.32	24 30			
16 - High V P13639 Elongation factor 2 US+Romo sapiens UX+96U6 GN+EEP	2 0.000	68%		00	339 49	1 85	90.3 6.8	3 3249	349.12	40 49			
17 4 P04350 Tubulin beta-4A chain US+Homo sapiens UX+3505 GN+1	0.000	73%		20	365 3	1 44	49.6 4.8	8 38/6	3.95.28	23 25		للللب	
18 - High V P68104 Elongation factor 1-alpha 1 US=Homo sapiens UX=9606 0	0.000	65%		25	33/ 25	1 46.	50.1 9.0	3105	334.51	19 24			
19 P High V P05/33 Alpha-enclase US=Homo saprens GX=9505 GN=EN01 Pr	0.000	76%		339	344 35	1 43	47.1 7.3	9 4055	334.14	30 39			
20 P68365 Tubulin alpha-4A chain OS+Homo sapiens OX+9506 GN+	0.000	56%		25	321 4	1 44	49.9 5.0	6 3293	325.46	21 25			
21 -P High V 043707 Alpha-actinin-4 OS+Homo sapiens OX+9606 GN+ACTN4	P 0.000	73%		56	383 41	1 91	104.8 5.4	4 3560	317.85	42 56			
22 😔 📄 High 🧹 Q14697 Neutral alpha-glucosidase AB OS+Homo sapiens OX+960	6 0.000	60%		38	317 38	1 94	106.8 6.1	4 2679	310.23	28 38			
23 Image: 23 Image: 23 Image: 23 Image: 23 Image: 23 Image: 24 Image: 24 Image: 23 Image: 24 Ima	0.000	66%		38	330 14	1 64	70.9 5.5	2 3208	304.12	27 37			
24 ··· High V P68032 Actin, alpha cardiac muscle 1 OS+Homo sapiens OX+9600	5 0.000	41%		21	431 2	1 37	42.0 5.3	9 4580	303.89	16 20			
25 ·□ High	G 0.000	80%		22	276 22	1 28	30.8 5.5	2 3311	302.34	19 22			
26 🗢 📃 High 🧹 075874 Isocitrate dehydrogenase [NADP] cytoplasmic OS+Homo e	н 0.000	76%		0 29	308 27	1 41-	46.6 7.0	1 3267	276.81	24 28			
27 ↔ Thigh V P35579 Myosin-9 OS=Homo sepiens OX=9606 GN=MYH9 PE=1 S	0.000	39%		64	270 52	1 196	226.4 5.6	0 3013	276.37	46 63			
28 😑 🔲 High 🧹 P55072 Transitional endoplasmic reticulum ATPase OS=Homo sap	÷ 0.000	55%) 45 :	295 45	1 80	89.3 5.2	6 2784	271.31	34 45			
29 🗢 💽 High 🧹 P05787 Keratin, type II cytoskeletal 8 OS=Homo sapiens OX=9606	0.000	80%		41	293 34	1 483	53.7 5.5	9 3030	257.31	32 41			
30 😐 🔲 High 🧹 K7EM38 Actin. cytoplasmic 2 (Fragment) OS=Homo sapiens OX=96	X 0.000	62%) 9	180 1	1 13	14.5 5.2	7	253.14	9			
31 🗢 🔲 High 🧹 Q13509 Tubulin beta-3 chain OS+Homo sapiens OX+9606 GN+TU	E 0.000	60%		22	257 6	1 450	50.4 4.5	3 3300	249.36	19 22			
32 😐 🔳 High 🧹 P11021 Endoplasmic reticulum chaperone BiP OS=Homo sapiens	C 0.000	57%		37	274 35	1 65	72.3 5.1	6 2770	247.14	29 37			
33 🗢 🚺 High 🧹 Q00610 Clathrin heavy chain 1 OS+Homo sapiens OX+9606 GN+C	0.000	46%		55	305 55	1 167	191.5 5.6	9 2893	244.67	40 55			
34 🗢 🔲 High 🧹 Q9NY65 Tubulin alpha-8 chain OS+Homo sapiens OX+9606 GN+T	0.000	37%		17	223 1	1 44	50.1 5.0	6 2186	241.73	13 17			
35 ··· C High ✓ P04075 Fructose-bisphosphate aldolase A OS=Homo sapiens OX+	< 0.000	84%		26	228 23	1 36-	39.4 8.0	9 2773	241.23	20 25			
36 ⊕ 🚺 High 🗸 Q06830 Peroxiredoxin-1 OS=Homo sapiens OX=9606 GN=PRDX1	0.000	92%		23	264 19	1 195	22.1 8.1	3 2427	239.10	18 23			
37 ·□ High ✓ P12277 Creatine kinase B-type OS=Home sapiens OX=9606 GN=1	0.000	75%		24	266 24	1 38	42.6 5.5	9 3037	228.03	17 23			
38 ··· ■ High	0.000	74%		27	202 27	1 41	44.6 8.1	0 1908	224.19	22 27			





Tryptic peptides									
	ם לא מי פי י								
Fi	ile Home Insert Page Layout Formulas Da								
	~								
1	peptide								
2	FFSYGDCMFLA								
3	AGEIIDGAELAR								
4	ALACVDAVAVFDEDTPER								
5	ALGSQGGAVLGPEAVR								
6	ANPVQDAQGISIESR								
7	APSEDIFVGGAGTPLR								
8	ASADVMAR								
9	CGDHSDTDAGMDTVFVNATR								
10	DDLAGVGAPYYR								
11	DEVNDLEAAGTSVIQVDEPALR								
12	DGILDFVATDHAPHAPEEK								
13	DSGVEVVIALK								
14	DTVGAASAK								
15	DVANQVTVHEVVGDVK								
16	EAGFVDVEVK								
47	EEL DOVODDDEV								

Figure 14. Pictographic representation of step 99 associated to downstream data processing for metaproteome analysis

- a. Uniprot (https://www.uniprot.org): A comprehensive database of protein sequences. (The usage is explained in detail in the bioinformatics section, especially database curation).
- b. Enricher (https://maayanlab.cloud/Enrichr/): Program for the functional enrichment analysis of proteins.

(Open the site > Copy the list of Gene symbols and paste in the Query section > Press Submit > Open the Reactome database > Download)

c. PANTHER (http://www.pantherdb.org): Program for the functional enrichment analysis of proteins.

(Open the site > Copy the list of Gene symbols and paste in Query section > Select ID type > Select organism > Select statistical overrepresentation test > Choose Reactome annotation > Submit > Select reference genome and run.

- d. Metaboanalyst 4.0 (https://www.metaboanalyst.ca/): A web server used for all the statistical analysis associated with the proteome, virome, and Metaproteome data.
 (Open site > Click on start > Click on statistical analysis > select data type, file and format > Submit and select preferred normalization > Perform desired statistics)
- e. Cytoscape (https://cytoscape.org/): a global cross-correlation map between the virome, metaproteome, and the pathways linked to the proteins and metabolites using Cytoscape.
- f. Unipept (https://unipept.ugent.be/): Database that's designed for metaproteomic data analysis with a focus on interactive data visualizations.

(The usage is explained in detail in the bioinformatics section, especially database curation)

LIMITATIONS

Although mass spectrometry-based proteomics and viromics are non-replaceable techniques in terms of how much they are essential in giving the snapshot of Protein expression of whole-cell,

Protocol



Welcome Unipept 4.3 Unipept is an open source web application developed at Ghent University that is designed for metaproteomics data analysis with a focus on interactive datavisualizations. Unipept is powered by an index Unipept 4.3 has just been released and contains these major containing all UniProt entries, a tweaked version of the NCBI taxonomy additions: and a custom lowest common ancestor algorithm. This combination Load and process multiple samples at once with the enables a blazingly fast biodiversity analysis of large and complex metaproteomics analysis tool. metaproteome samples. This functionality is also available via an API · Compare multiple samples with each other using the new and a set of command line tools. Next to these core functions, Unipept heatmap. Functional analysis has been expanded with support for also has a tool for selecting unique peptides for targeted proteomics InterPro annotations and for comparing genomes based on peptide similarity. UniProt updated to version 2020.01 Tryptic Peptide Analysis Metaproteome Analysis Unique

Figure 15. Pictographic representation of step 101 associated to downstream data processing for metaproteome analysis

unlike genome, the proteome of cells is highly dynamic, and thus, multiple samples and time points need to be considered if study demands temporal profiling (Betzen et al., 2015). The starting sample should be fresh as degradation in the proteins would result in lower identification. The extensive protein and peptide cleanup procedure used in the protocol tends to lose some proteins and peptides, resulting in lower identification.

TROUBLESHOOTING

Problem 1

Interference of chemicals such as polymers in LC-MS/MS analysis (steps 83-87).

Potential solution

In such a scenario, the prepared peptides samples need to be reprecipitated. Add $6 \times$ acetone to your peptide samples and incubate them at -20° C for 1 h > centrifuge at maximum g possible and discard the supernatant. Dissolve the pellet again and reanalyze in mass-spectrometer.

If still the problem persists, reprepare your samples, and this time before using any plastic ware, soak them in any organic solvent (e.g., ethanol or acetone) for 24–72 h to remove any contaminant polymer.

Problem 2

The trap column of Nano-LC gets blocked (steps 83–87).

Potential solution

In such a scenario, first, you disconnect the column from source > then reverse the direction of flow in the column by reversing the direction of column > then start the isocratic flow and wait till water





Unipept		
	Metaproteomics Analysis Peptidome Analysis API CLI Metagenomics	
	METAPROTEOMICS ANALYSIS	CREATE SAMPLE DATA PRIDE LOCAL DATA
	Selected datasets Please select one or more datasets from the right hand panel to continue the analysis.	Peptide list
	Search settings Equate I and L Filter duplicate peptides Advanced missed cleavage handling SEARCH Start OVER	Name this dataset e.g. Sample B5 Store dataset in browser's local storage + ADD TO SELECTED DATASETS

Figure 16. Pictographic representation of step 102 associated to downstream data processing for metaproteome analysis

starts to appear or flow from the open end of the column once it happens. Rearrange the column and start.

Problem 3

The air bubble got trapped in the nano-LC pump (steps 83–87).



Figure 17. Pictographic representation of step 102 associated to downstream data processing for metaproteome analysis

Protocol



METAPROTEOMICS ANALYSIS	+	COVID
 COVID 309 peptides 	9/2/2021	Peptide list FFSYGDCMFLA AGEIIDGAELAR ALACVDAVAVFDEDTPER ALGSQGGAVLGPEAVR ANPVQDAQGISIESR Search settings □ Equate I and L ✓ Filter duplicate peptides □ Advanced missed cleavage handling
COMPARE SAMPLES	3	We managed to match 308 of your 308 peptides. Unfortunately, 0 peptides couldn't be found.

Figure 18. Pictographic representation of step 104 associated to downstream data processing for metaproteome analysis

Potential solution

In such a scenario, first, you disconnect the column from source > then join the column end with waste outlet> then start purge > the flush finally performs a leak test if passed. Rearrange the column and start.

Problem 4

No Peaks/Very Small Peaks (steps 83-87).

	⊟ 5× ở× ∓		COVID_mpa - Excel	A (
F	ile Home Insert Page Layout For	mulas Data Review View Help	Q Tell me what you want to do			
1	A	В	C D E F	G		
1	peptide	Ica	superking kingdom subkingdom superphylum	phylum		
2	FFSYGDCMFLA	Acidobacterium capsulatum ATCC 51196	Bacteria	Acidobacteria		
З	AGEIIDGAELAR	Streptomyces	Bacteria	Actinobacteria		
4	ALACVDAVAVFDEDTPER	Streptomyces	Bacteria	Actinobacteria		
5	ALGSQGGAVLGPEAVR	Mycolicibacterium	Bacteria	Actinobacteria		
6	ANPVQDAQGISIESR	Streptomyces	Bacteria	Actinobacteria		
7	APSEDIFVGGAGTPLR	Streptomyces	Bacteria	Actinobacteria		
8	ASADVMAR	Corynebacteriales	Bacteria	Actinobacteria		
9	CGDHSDTDAGMDTVFVNATR	Mycobacteriaceae	Bacteria	Actinobacteria		
10	DDLAGVGAPYYR	Mycobacterium	Bacteria	Actinobacteria		
11	DEVNDLEAAGTSVIQVDEPALR	Streptomycetaceae	Bacteria	Actinobacteria		
12	DGILDFVATDHAPHAPEEK	Rubrobacter xylanophilus	Bacteria	Actinobacteria		
13	DSGVEVVIALK	Actinomycetia	Bacteria	Actinobacteria		
14	DTVGAASAK	Thermobifida fusca YX	Bacteria	Actinobacteria		
15	DVANQVTVHEVVGDVK	Actinomycetia	Bacteria	Actinobacteria		
16	EAGFVDVEVK	Saccharopolyspora erythraea NRRL 2338	Bacteria	Actinobacteria		
17	FEI DOVEDDDEK	Common handward all an	Destants.	A set of a local set of a		

Figure 19. Pictographic representation of step 105 associated to downstream data processing for metaproteome analysis



Protocol

⊟ 5× ♂× ∓				all togethe	er metaprote	tins bile fina	01092021 -	Excel			A	Gaurav Tri	pathi 🎴		- 6	×
File Home Insert Page Layout Formula	as Data	Review	r View	Help	Qт	ell me what	you want	to do							Z	Share
$ \begin{array}{c c} & & & \\ & & $		♦ ✓ ✓ Alignment	환 Wrap Te 태 Merge 8	xt k Center ~ r	General \$ ~ 1	% > %	- C - Fo - Fo	onditional F rmatting ~ S	format as Table ~ Sty tyles	Cell In /les ~	sert Delete	Format	∑ AutoSur ↓ Fill ~ ✓ Clear ~	m X A Sort 8 Filtery Editing	Find & Select ~	^
CR12 * : × ✓ fr =LOG10(AX1	2)															v
CM	CN	со	CP	cq	CR	CS	СТ	cu	CV	CW	CX	CY	CZ	DA	DB	DC 🔺
1 species	Abundanc	Abundanc	Abundanc	Abundanc	Abundanc	Abundanc	Abundand	Abundanc	Abundanc	Abundanc	Abundanc	Abundand	Abundanc	Abundanc	Abundanc	Abund
2 Acidobacterium capsulatum ATCC 51196	CAGB+	CAGB+ ~	CAGB+i *	CAGB+(*	CAGB+	CAGB+(~	CAGB+(*	CAGB+(*	CAGBH -	CAGB+	CAGB 👻	CAGB 💌	CAGB 👻	CAGB 💌	CAGB 💌	CAGB
3 Streptomyces	3.985056	3.950268	4.448969	4.19153	4.578209	4.441559	4.598167	4.184353	6.249744	4.26313	4.545819	4.327645	4.041082	4.050932	4.151964	4.4355
4 Streptomyces	4.64913	4.698538	6.417867	4.383369	4.528956	4.204618	4.539119	4.2041	4.346952	4.650422	3.970124	4.422697	4.703109	4.023395	4.612617	3.8718
5 Mycolicibacterium	4.986855	4.658235	6.04066	4.074063	3.686446	6.086496	4.249126	6.004014	5.787249	4.628735	4.560202	3.483277	4.444202	6.972999	4.491865	6.220:
6 Streptomyces	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
7 Streptomyces	4.030544	4.246576	4.57909	4.696856	3.882883	3.574533	3.982993	4.669799	4.622663	4.475023	4.60735	5.625842	4.164092	4.192561	3.98645	4.4565
8 Corynebacteriales	3.700823	4.591095	6.207588	5.490694	4.67042	4.462236	4.135257	4.497986	4.512902	4.21164	4.365295	4.199541	4.621495	4.373854	4.515191	4.0092
9 Mycobacteriaceae	3.985056	3.950268	4.448969	4.19153	4.578209	4.441559	4.598167	4.184353	6.249744	4.26313	4.545819	4.327645	4.041082	4.050932	4.151964	4.4355
10 Mycobacterium	4.64913	4.698538	6.417867	4.383369	4.528956	4.204618	4.539119	4.2041	4.346952	4.650422	3.970124	4.422697	4.703109	4.023395	4.612617	3.8718
11 Streptomycetaceae	4.986855	4.658235	6.04066	4.074063	3.686446	6.086496	4.249126	6.004014	5.787249	4.628735	4.560202	3.483277	4.444202	6.972999	4.491865	6.220:
12 Rubrobacter xylanophilus	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
13 Actinomycetia	4.030544	4.246576	4.57909	4.696856	3.882883	3.574533	3.982993	4.669799	4.622663	4.475023	4.60735	5.625842	4.164092	4.192561	3.98645	4.4565
14 Thermobifida fusca YX	3.700823	4.591095	6.207588	5.490694	4.67042	4.462236	4.135257	4.497986	4.512902	4.21164	4.365295	4.199541	4.621495	4.373854	4.515191	4.009:
15 Actinomycetia	4.535377	3.604453	4.291042	6.305179	4.679005	6.15751	4.603604	5.038701	4.195364	4.123753	4.704925	4.498028	5.465421	3.619459	4.129436	3.673:
16 Saccharopolyspora erythraea NRRL 2338	5.787774	6.056748	7.387746	8.076922	6.187671	7.771453	6.198297	6.660543	8.623757	8.076085	6.978339	7.786113	7.988383	7.208519	6.926465	7.31
17 Corynebacteriales	4.498857	4.39969	6.208059	4.391948	4.454063	5.691522	4.521783	4.845506	5.673777	3.826419	4.115631	5.495551	6.017553	4.624968	4.523843	5.130
18 Actinomycetia	7.281888	6.917465	3.960728	7.473476	3.86017	4.110373	6.729083	6.849299	7.51256	7.111949	6.609995	7.165759	3.852846	4.274494	7.109594	4.234/
19 Mycobacterium tuberculosis complex	4.48225	4.221428	4.012584	3.970307	4.320598	3.910714	4.41736	5.203909	4.536621	3.591205	4.674772	4.15319	4.323898	6.260772	3.490904	4.5360
20 Tropheryma whipplei	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
21 Mycobacteriaceae	4.631716	4.631057	4.345048	3.95807	4.614321	4.22511	3.722386	4.535484	3.869263	5.493717	4.673505	4.611447	4.501957	4.464785	3.462749	4.4546
22 Bifidobacterium	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
23 Mycobacterium tuberculosis complex	4.299363	6.357258	4.62546	4.692415	6.243259	6.502643	6.482995	6.319068	4.22065	4.215612	5.541137	4.354929	7.159359	7.337209	4.28351	7.091
24 Streptomyces	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
35 Streatemuses arisesus	Anochec	4 100150	4 05 300	5 605037	A 610336	E 40111	2.046430	C 220220	4.136044	E ADECA	A 641763	e sennee	4 410307	3 600305	4.040046	4.1001
cago_mpa_18082021 GS HC	rinai all n	netaproteii	she	214	Ð			: 4								•
Ready												=	m =		-	- + 100%

Figure 20. Pictographic representation of step 108 associated to downstream data processing for metaproteome analysis

Potential solution

In such a scenario, check that the lamp is on and the cables are well connected. Next, ensure that the flow is normal and the automatic sampler is working properly, not blocked or jammed.

Problem 5

No or low protein detection (steps 60-82).

Potential solution

In such a scenario take your flowthrough and repurify peptide from it again. Never discard the flowthrough.

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Dr. Jaswinder Singh Maras (jassi2param@gmail.com).

Materials availability

The study did not generate any materials.

Data and code availability

No data is provided, and the raw data for the manuscript is available on request to the lead contact Dr. Jaswinder Singh Maras (jassi2param@gmail.com).

ACKNOWLEDGMENTS

We want to acknowledge the Science and Engineering Research Board Department of Science and Technology Government of INDIA for providing funds to build the manuscript. The work was supported by project DST (DST-SERB) (EMR/2016/004829).

Protocol



AUTHOR CONTRIBUTIONS

J.S.M. conceptualized the work. The manuscript was written by G.T. and J.S.M., with help from M.Y., N.S., B.M., V.B., S.S., and S.K.S. The manuscript was read and approved by all authors.

DECLARATION OF INTERESTS

The authors declare no competing interests.

REFERENCES

Batejat, C., Grassin, Q., Manuguerra, J.C., and Leclercq, I. (2021). Heat inactivation of the severe acute respiratory syndrome coronavirus 2. J. Biosaf Biosecur. 3, 1–3.

Betzen, C., Alhamdani, M.S., Lueong, S., Schroder, C., Stang, A., and Hoheisel, J.D. (2015). Clinical proteomics: promises, challenges and limitations of affinity arrays. Proteomics Clin. Appl. 9, 342–347.

Bhat, A., Das, S., Yadav, G., Chaudhary, S., Vyas, A., Islam, M., Gupta, A.C., Bajpai, M., Maiwall, R., Maras, J.S., et al. (2020). Hyperoxidized albumin modulates platelets and promotes inflammation through CD36 receptor in severe alcoholic hepatitis. Hepatol. Commun. 4, 50–65.

Bradford, M.M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal Biochem. *72*, 248–254.

Das, S., Hussain, M.S., Maras, J.S., Kumar, J., Shasthry, S.M., Nayak, S., Arora, V., Vijayaraghavan, R., Sharma, S., Maiwall, R., et al. (2019). Modification patterns of urinary albumin correlates with serum albumin and outcome in severe alcoholic hepatitis. J. Clin. Gastroenterol. 53, e243–e252.

Das, S., Maras, J.S., Hussain, M.S., Sharma, S., David, P., Sukriti, S., Shasthry, S.M., Maiwall, R., Trehanpati, N., Singh, T.P., et al. (2017). Hyperoxidized albumin modulates neutrophils to induce oxidative stress and inflammation in severe alcoholic hepatitis. Hepatology 65, 631–646.

Dowling, P., Gargan, S., Zweyer, M., Henry, M., Meleady, P., Swandulla, D., and Ohlendieck, K. (2020). Protocol for the bottom-up proteomic analysis of mouse spleen. STAR Protoc. 1, 100196.

Gundry, R.L., White, M.Y., Murray, C.I., Kane, L.A., Fu, Q., Stanley, B.A., and Van Eyk, J.E. (2009). Preparation of proteins and peptides for mass spectrometry analysis in a bottom-up proteomics workflow. Curr. Protoc. Mol. Biol. *Chapter 10*, Unit10 25.

Maras, J.S., Sharma, S., Bhat, A., Rooge, S., Aggrawal, R., Gupta, E., and Sarin, S.K. (2021). Multi-omics analysis of respiratory specimen characterizes baseline molecular determinants associated with SARS-CoV-2 outcome. iScience 24, 102823.

Nowakowski, A.B., Wobig, W.J., and Petering, D.H. (2014). Native SDS-PAGE: high resolution electrophoretic separation of proteins with retention of native properties including bound metal ions. Metallomics *6*, 1068–1078.

Pondaven-Letourmy, S., Alvin, F., Boumghit, Y., and Simon, F. (2020). How to perform a nasopharyngeal swab in adults and children in the COVID-19 era. Eur. Ann. Otorhinolaryngol. Head Neck Dis. 137, 325–327.

Simpson, D.M., and Beynon, R.J. (2010). Acetone precipitation of proteins and the modification of peptides. J. Proteome Res. *9*, 444–450.

van den Brand, J.M., Haagmans, B.L., van Riel, D., Osterhaus, A.D., and Kuiken, T. (2014). The pathology and pathogenesis of experimental severe acute respiratory syndrome and influenza in animal models. J. Comp. Pathol. *151*, 83–112.