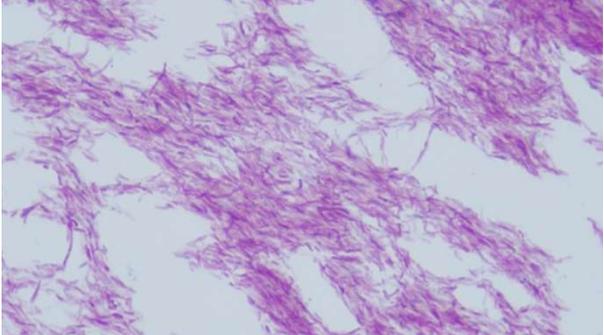


SIGNIFICANCE

This is the first study that compares the saliva and the sputum from active TB patients and their contacts. Our findings strongly suggest that TB patients show not only an activation of processes that are related to complement activation and modulation of inflammation, but also an imbalance in carbohydrate and lipid metabolism. In addition, those individuals who do not get infected after direct exposure to the pathogen display a typical proteomic signature in the sputum which is a reflection of the secretion from the nasal and oral mucosa, the first immunological barriers that *M. tuberculosis* encounters in the host. Thus, this result indicates the importance of the processes related to the innate immune response in fighting the initial events of the infection.



Eliciting Mucosal Immunity in TB

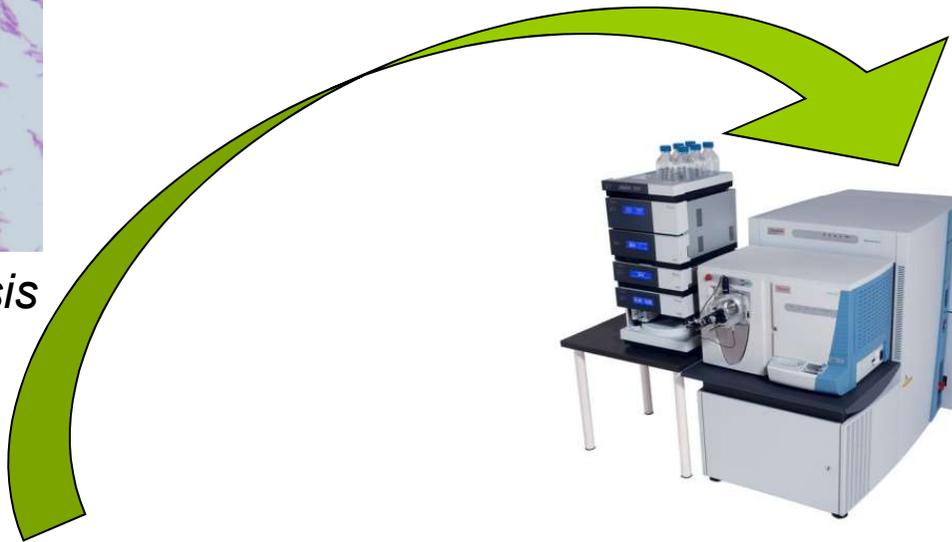


Mycobacterium tuberculosis



TUBERCULOSIS PATIENTS AND CONTACTS SAMPLE COLLECTION

**-SPUTUM
-SALIVA**



Shotgun Quantitative Proteomics

Specific Protein Signature of

Tuberculosis patients:

High levels of markers of inflammation and acute-phase response vs. latent-infected contacts in both fluids

Uninfected contacts:

Increased levels of proteins related to innate immune response vs. latent-infected contacts in sputum

HIGHLIGHTS

- Proteomic analysis of saliva and sputum in tuberculosis patients and contacts.
- Both fluids differ in terms of protein composition.
- Active TB patients show markers of inflammation and complement activation.
- Active TB patients present a decrease in enzymes related to sugar metabolism.
- Markers of innate immune response are higher in the sputum of uninfected contacts.

1 **High-Resolution Quantitative Proteomics applied to the study of the**
2 **specific protein signature in the sputum and saliva of active**
3 **tuberculosis patients and their infected and uninfected contacts**

4 Jesús Mateos^{1*}, Olivia Estévez², África González-Fernández², Luis Anibarro^{2,3,4},
5 Ángeles Pallarés³, Rajko Reljic⁵, José M. Gallardo¹, Isabel Medina¹ and Mónica
6 Carrera^{1^}.

7

8 ¹*Spanish National Research Council (CSIC), Vigo, Pontevedra, Spain*

9 ²*Biomedical Research Centre (CINBIO), Galician Singular Center of Research, Galicia*
10 *Sur Health Research Institute (IIS-GS), University of Vigo, Vigo, Pontevedra, Spain.*

11 ³*Tuberculosis Unit, Infectious Diseases, Internal Medicine service, Complejo*
12 *Hospitalario Universitario de Pontevedra, Galicia Sur Health Research Institute (IIS-*
13 *GS, Pontevedra, Spain.*

14 ⁴*Mycobacterial Infections Study Group (GEIM) of the Spanish Society of Infectious*
15 *Diseases and Clinical Microbiology (SEIMC), Madrid, Spain.*

16 ⁵*St. George's, University of London, UK.*

17

18 *CORRESPONDING AUTHOR: Jesús Mateos, PhD

19 Spanish National Research Council (CSIC), 36208, Vigo, Pontevedra, Spain.

20 Phone: +34 986231930

21 Fax: +34 986 292762

22 e-mail: jmateos@iim.csic.es

23 ^ CO-CORRESPONDING AUTHOR: Mónica Carrera, PhD

24 Spanish National Research Council (CSIC), 36208, Vigo, Pontevedra, Spain.

25 Phone: +34 986231930

26 Fax: +34 986 292762

27 E-mail: mcarrera@iim.csic.es

28

29

30 **ABSTRACT (154words)**

31 Our goal was to establish panels of protein biomarkers that are characteristic of patients
32 with microbiologically confirmed pulmonary tuberculosis (TB) and their contacts,
33 including latent TB-infected (LTBI) and uninfected patients. Since the first pathogen-
34 host contact occurs in the oral and nasal passages the saliva and sputum were chosen as
35 the biological fluids to be studied. Quantitative shotgun proteomics was performed
36 using a LTQ-Orbitrap-Elite platform. For active TB patients, both fluids exhibited a
37 specific accumulation of proteins that were related to complement activation,
38 inflammation and modulation of immune response. In the saliva of TB patients, a
39 decrease of in proteins related to glucose and lipid metabolism was detected. In contrast,
40 the sputum of uninfected contacts presented a specific proteomic signature that was
41 composed of proteins involved in the perception of bitter taste, defense against
42 pathogens and innate immune response, suggesting that those are key events during the
43 initial entry of the pathogen in the host.

44

45

46

47

48

49

50

51

52

53 **SIGNIFICANCE**

54

55 This is the first study to compare the saliva and sputum from active TB patients and
56 their contacts. Our findings strongly suggest that TB patients show not only an
57 activation of processes that are related to complement activation and modulation of
58 inflammation but also an imbalance in carbohydrate and lipid metabolism. In addition,
59 those individuals who do not get infected after direct exposure to the pathogen display a
60 typical proteomic signature in the sputum, which is a reflection of the secretion from the
61 nasal and oral mucosa, the first immunological barriers that *M. tuberculosis* encounters
62 in the host. Thus, this result indicates the importance of the processes related to the
63 innate immune response in fighting the initial events of the infection.

64

65

66 **HIGHLIGHTS**

67

- 68 -Proteomic analysis of saliva and sputum in tuberculosis patients and contacts.
- 69 -Both fluids differ in terms of protein composition.
- 70 -Active TB patients show markers of inflammation and complement activation.
- 71 -Active TB patients present a decrease in enzymes related to sugar metabolism.
- 72 -Markers of innate immune response are higher in the sputum of uninfected contacts.

73

74

75 1. Introduction

76

77 Tuberculosis (TB) is a major health problem globally (1). TB is caused by the
78 bacterium *Mycobacterium tuberculosis* (MTB). Roughly one third of the world's
79 population carries MTB in a dormant form (2). TB is responsible for the death of more
80 than 1.8 million people each year, making it one of the leading causes of mortality and
81 the most common cause of death by a single infectious agent (3). Since evidence of the
82 infection has been found in human remains dating back 5000 years, TB is considered
83 one of the oldest diseases in mankind (4). Pulmonary TB is air transmitted from an
84 active TB patient to a healthy individual through coughing or sneezing. Infection occurs
85 through nasal/oral inhalation of aerosol droplets carrying MTB (5). The smaller droplets
86 are able to reach the lower lung, and, after recruitment of macrophages and dendritic
87 cells, they form the granuloma, which is a host-defensive structure that is characteristic
88 of latent infections (6). Granulomas provide a fibrotic physical barrier between the
89 infected, necrotic area and the healthy neighboring tissue (7). LTBI patients cannot
90 infect a healthy individual; however, eventually, activation of the pathogen occurs in
91 approximately 10% of the cases due to recurrent infections, immunosuppression or a
92 weakened health state of the host (8).

93 Remarkably, many individuals in close contact with an active TB patient do not
94 become infected. Increasing evidence suggests that genetic heterogeneity of the host
95 affects immune response to intracellular pathogens, such as mycobacteria (9-11). Innate
96 immune response is the first line of defense and plays a key role in the quality, strength
97 and efficacy of the adaptive immune response (12). However, it remains unknown
98 whether or not those healthy household contacts present a specific proteomic signature
99 that could be linked to this particular state of special innate protection. Vaccines offer a

100 safe and cost-effective method to protect large populations against infectious diseases
101 or, alternatively, mitigate their clinical course (13). However, many poverty-related and
102 neglected infectious diseases such as TB continue to escape attempts to develop
103 effective vaccines against them (14). BCG (Bacille Calmette-Guerin), the vaccine
104 currently in use against TB, was developed eighty years ago and is widely used for
105 prevention, with an efficiency of more than 80% in children under 4 years (15).
106 Unfortunately, BCG efficiency in adolescents and adults is variable, ranging from 0 to
107 80% (16). The World Health Organization Global Strategy for the period 2015-2035
108 draws special attention to the urgent need to develop a new efficient vaccine against TB
109 (3, 17). Eliciting Mucosal Immunity in Tuberculosis (EMI-TB) consortium ([www.emi-](http://www.emi-tb.org)
110 [tb.org](http://www.emi-tb.org)) is a Horizon-2020-funded action that focuses on selecting candidates for
111 developing a new, nasal-administered vaccine against TB. Our laboratory (CSIC) is
112 integrated in the working-package 3, and our task is the proteomic analysis of biological
113 samples.

114 Proteomics provides a unique tool to analyze cellular and organism activity at
115 the protein level. Thus, proteomic profiling allows the elucidation of the links between
116 broad cellular pathways and individual molecules that were previously impossible to
117 predict using only traditional biochemical analysis. Biological fluids from human
118 subjects are a promising source for analyzing biological markers of health and disease
119 (18). Biological fluids contain biomolecules (including lipids, peptides, amino acids,
120 cytokines, proteases, enzymes, and antibodies) that present different physicochemical
121 properties. Since analytical proteomics has experienced extensive progress in the last
122 decade because of the emergence of mass-spectrometry-based techniques (MALDI-
123 TOF/TOF, LC-MS/MS), its combination with classical techniques for protein separation
124 (e.g., iso-electric focusing, chromatography, and two-dimensional gel electrophoresis)
125 facilitate the identification and characterization of thousands of proteins in a single

126 experiment. Proteomics is expected to be the tool of choice for the search for diagnostic
127 or therapeutic biomarkers and for the identification and characterization of the proteins
128 encoded by the genome (19, 20). One of the main objectives of proteomics is the
129 identification of markers of disease by comparing the protein status in normal and
130 pathological conditions.

131 Our specific aim for this study, as part of the Horizon2020 EMI-TB project, was to test
132 two different biological fluids, sputum and saliva to establish the proteomic signature
133 that is specific to active TB patients and their contacts, including LTBI and uninfected
134 patients. We aim to determine indicators of the innate immune response that prevents
135 the development of the disease and improve our understanding of the processes that are
136 activated during MTB latency and active TB disease.

137

138 **2. Material and methods**

139

140 **2.1. Patient selection and database management**

141 Active culture-positive pulmonary TB patients and their close contacts were
142 prospectively enrolled in patients attending the TB Unit in Pontevedra, Galicia, Spain
143 (Table 1). Collection of samples initiated on September 2015, within the framework of
144 the EMI-TB project (Project ID: 643558; Eliciting Mucosal Immunity to Tuberculosis;
145 Ongoing project H2020-EU.3.1: SOCIETAL CHALLENGES; “Health, demographic
146 change and well-being”). Exclusion criteria were ages less than 18 y, coinfection with
147 the human immunodeficiency virus (HIV), any other immunosuppressive medical
148 condition or concomitant use of immunosuppressive drugs. Patients with previous TB
149 infection or LTBI were also excluded for the study. All patients accepted to be included
150 in the study and signed informed consent. The database was elaborated using codes,
151 keeping the anonymity of the person and including all the clinical fields considered
152 relevant for the study. The informed consent was prepared, and all patients and contacts
153 received a detailed explanation of the project and confidentiality. The total number of
154 patients included in the study was 99, including 26 active TB patients and 73 contacts
155 (Table 1). The contacts were screened using the tuberculin skin test (TST) and/or the
156 Quantiferon-TB-Gold test (QFT) and again after 8-12 weeks after the last possible
157 exposure to the index case if the first test was negative. Chest radiography was
158 performed to exclude active TB in patients with a positive TST-QFT result. The patients
159 were classified as follows: active TB, LTBI or uninfected, following Spanish national
160 guidelines (21). A “Contact Score” was assigned to the contacts following this criteria:
161 acid-fast bacilli (AFB) microscopy index: from 0 to 4; cavitory X-rays: No=0, Yes=1;
162 household contact: No=0, Yes=1; hours of exposure/day: 0-3 hours: 1, 4-7 hours: 2, 8-

163 11 hours: 3, \geq 12 hours: 4; type of exposure: Outdoors= 0.25, Different room: 1, As a
164 bar: 2, As a class: 2.5, As an office: 3, As a room or car: 4; weeks in contact with Index
165 Case: < 12 weeks: 0, \geq 12 weeks: 1; sleeps in the same room: No=0, Yes=1; first-degree
166 family relationship: No=0, Yes=1.

167

168 **2.2. Collection and *M. tuberculosis* decontamination of saliva and sputum samples**

169 Samples of saliva and sputum were collected from the selected patients in the
170 Tuberculosis Unit of the “*Complejo Hospitalario Universitario de Pontevedra*”
171 Hospital Facility following the ethical committee authorization from the Galician Ethics
172 Committee for Clinical Research (CEIC, Ref 2014/492). Saliva was constituted only for
173 a transparent liquid produced in the mouth without clots and was collected and
174 processed following standard procedures in the hospital facilities (22). Not induced
175 sputum samples (2-4 mL) were collected in sterile 50 mL polypropylene tubes and
176 stored at 4°C until processing. After addition of four volumes of 1% β -mercaptoethanol
177 in Phosphate Buffered Saline (PBS), the samples were shaken at 37°C for 15 min and
178 centrifuged at 300 rpm for another 15 min to collect the supernatant fraction.

179 After collection, saliva and sputum samples were transferred to the
180 Microbiology Unit of the same hospital, where they were processed for the inactivation
181 of *Mycobacterium tuberculosis*. The mechanical disruption method was used to ensure
182 the liquefaction of samples without damaging the proteins and RNA. For this purpose,
183 zirconium-silica beads were used to homogenize the samples in a BeadBeater (Biospec,
184 Bartlesville, OK 74005, USA). Once decontaminated, aliquots of the samples were
185 seeded and cultured. The elimination of MTB from the treated samples was confirmed
186 via microbial culture on Coletsos culture medium. This assay was performed in the
187 Laboratory of Microbiology of the “*Complejo Hospitalario de Pontevedra*”. After

188 confirmation of the total elimination of the pathogen, frozen samples were transferred to
189 University of Vigo and stored at -80°C.

190 **2.3. Protein preparation**

191 Total protein from individual samples was determined by measuring the
192 absorbance at 280 nm using the NanoDrop™ instrument (Thermo Fisher Scientific, San
193 Jose, CA, USA) and was precipitated by adding six volumes of cold acetone and
194 overnight incubation at -20°C. After centrifugation, the dried protein pellet was
195 resuspended in 0.1 M triethylammonium bicarbonate (TEAB) buffer solution. The
196 protein concentration for TMT labeling was determined using the Bicinchoninic Acid
197 (BCA) assay (Sigma-Aldrich, St. Louis, MO, USA). Next, the protein integrity of all
198 individual samples was confirmed via 10% acrylamide SDS-PAGE of 1 µg aliquots and
199 silver staining of the resulting gels. The samples in which protein degradation was
200 detected were discarded. The densitometry of the entire lane was used, if necessary, to
201 correct protein quantification.

202

203 **2.4. Shotgun proteomics study design**

204 A schematic of the study design is presented in Figure 1A. Samples were
205 selected for the quantitative proteomics study according to their protein concentration
206 and integrity. Three complete Tandem Mass Tag (TMT) 10plex labeling procedures
207 (Thermo Fisher Scientific) were performed for each biological fluid using individual
208 samples. Within each TMT experiment, three active TB (channels 126, 127N and
209 127C), three infected LTBI (channels 128N, 128C and 129N) and three uninfected
210 patients (channels 129C, 130N and 130C) were included, plus a standard sample
211 resulting from mixing equal amounts of proteins for the nine samples included (channel
212 131). A summary of the samples used for the study is presented in Table 2.

213

214 **2.5. TMT 10plex labeling**

215 For the labeling, 100 µg of each individual sample was resuspended in a final
216 volume of 100 µL of 0.1 M TEAB buffer solution, reduced/alkylated and digested with
217 trypsin for 16 h at 37°C. Labeling with TMT 10plex reagents (Thermo Fisher Scientific)
218 was performed following the manufacturer's instructions. Briefly, LC-grade acetonitrile
219 was used to dissolve the reagents (41 µL per 0.8 mg of reagent). After labeling for 1
220 hour and quenching with 8 µL of 5% hydroxylamine for 15 min, both steps at room
221 temperature, all the channels were mixed in a single tube, aliquoted and dried in a
222 speed-vac.

223

224 **2.6. Peptide fractionation by High-pH Reversed Phase**

225 Aliquots of 100 µg of the total labeled protein were reconstituted in 300 µL of
226 trifluoroacetic acid (TFA), and the peptides were fractionated using the High-pH
227 Reversed Phase fractionation kit (Thermo Fisher Scientific) following the
228 manufacturer's instructions. The peptide concentration in the resulting fractions (10 plus
229 the Washing and the Flow-through) was determined via colorimetric analysis using the
230 Quantitative Colorimetric Peptide Assay (Thermo Fisher Scientific).

231

232 **2.7. LC-MS/MS analysis and Orbitrap-Elite settings**

233 One microgram of the samples of each fraction, as determined using the
234 colorimetric assay, were injected and analyzed in the Proteomics Facility of the
235 University of Vigo (CACTI, Vigo, Spain) via LC-MS/MS using a Proxeon EASY-nLC
236 II liquid chromatography system (Thermo Fisher Scientific) coupled to a LTQ-Orbitrap
237 Elite mass spectrometer (Thermo Fisher Scientific). The peptides were separated on an
238 RP column (EASY-Spray column, 50 cm × 75 µm ID, PepMap C18, 2 µm particles,

239 100 Å pore size, Thermo Fisher Scientific) with a 10 mm precolumn (Accucore XL
240 C18, Thermo Fisher Scientific) using 0.1% formic acid (mobile phase A) and 98% ACN
241 with 0.1% formic acid (mobile phase B). A 240 min linear gradient from 5% to 35% B
242 was applied at a flow-rate of 300 nL per min. Ionization was performed in a nanosource
243 using a spray voltage of 1.95 kV and a capillary temperature of 275°C. The peptides
244 were analyzed in positive mode (1 μ scan; 400–1600 amu), followed by 10 data-
245 dependent HCD MS/MS scans (1 μ scans), using a normalized collision energy of 38%
246 and an isolation width of 1.5 amu. Dynamic exclusion was enabled with a repeat count
247 of 1, a repeat duration of 30 s, a duration of the exclusion of 80 s, and a relative
248 exclusion width of 10 ppm. Unassigned charged ions were excluded from the analysis.

249 **2.8. *Mass spectrometry data processing***

250 Raw data were loaded in the Xcalibur software (Thermo Fisher Scientific) for
251 inspection of the chromatography profile and confirmation of the labeling of the
252 peptides. Next, protein identification and quantification was performed using the
253 Proteome Discoverer 2.1 software (Thermo Fisher Scientific). Peak lists were generated
254 with a precursor signal-to noise ratio of 1.5, and default settings were used to search the
255 latest UniProtKB Release using the SEQUEST algorithm. The enzyme specificity was
256 set to trypsin and one missed cleavage was tolerated. TMT-labeling and
257 carbamidomethylation of cysteine were set as fixed modifications, whereas oxidation of
258 methionine and N-terminal acetylation were set as variable modifications. The precursor
259 ion mass tolerance was set to 7 ppm, and the product ion mass tolerance was set to 0.8
260 Da. A decoy database search was performed to determine the peptide false discovery
261 rate (FDR) with the Target Decoy PSM Validator module. Quantification was
262 performed using the Quantification Module, and normalization was performed against
263 total peptide amount. A 1% peptide FDR threshold was applied.

264 Samples were categorized by the patient type (active TB, LTBI, uninfected and
265 Internal Standard). Quantification jobs were alternatively launched using a) the Patient
266 Type option for the global analysis and b) the individual ratio option for the
267 nonparametric statistical analysis.

268

269

270 **2.9. Selection of quantified proteins**

271 Several filters were sequentially applied to the global quantification results to
272 obtain the final list of quantified proteins: A) proteins quantified with at least two
273 unique peptides, B) proteins quantified in the three TMTs for each biological fluid, C)
274 p-value ≤ 0.0001 after performing the Kruskal-Wallis statistical test for all the different
275 ratios. A summary of the filtering process is presented in Figure 1B.

276

277 **2.10. Statistical analysis of the modulated proteins**

278 The normalized ratios for the all the proteins quantified in the three TMTs with
279 at least two unique peptides were extracted from the Proteome Discoverer 2.1 software
280 using the “export to Excel” option and were used for the Kruskal-Wallis statistical
281 analysis using the R software. Briefly, the normalized protein ratios (27 for each
282 comparison) were imported into R commander console and represented in box-
283 diagrams. For each protein, data were used to analyze the differences between the three
284 studied ratios (27 ratios uninfected/LTBI, 27 ratios uninfected/TB and 27 ratios
285 LTBI/TB) by applying a Kruskal-Wallis test. Differences in the modulation were
286 considered as significant when p-value ≤ 0.0001 .

287 The final list (UniProtKB accession number) of modulated proteins was
288 analyzed using the String 10.1 software for the determination of pathways and
289 biological processes modulated (p-value ≤ 0.001) in each group of patients.

290

291

292 .

293 **3. Results**

294

295 **3.1. *Sputum and saliva present differences in terms of protein composition and*** 296 ***quantification***

297 Proteomic datasets are deposited at the MassIVE repository
298 (www.massive.ucsd.edu). Raw and processed files (EMI_TB_PROTEOMICS_CSIC;
299 #MSV000081574) are public and freely accessible. A total of 1218 and 847 proteins
300 were identified with at least one unique peptide in the sputum and saliva samples,
301 respectively. Among them, 766 and 562 were quantified with at least two unique
302 peptides. A total of 755 proteins were identified in both biological fluids
303 (Supplementary Figure 1A). The sputum appeared to be more complex than the saliva,
304 presenting a high number of specific protein isoforms, 548, whereas only 164 proteins
305 were detected exclusively in the saliva. Additionally, the sputum and saliva appeared to
306 differ in protein composition, because we found no correlation between similar TMT
307 ratios when comparing both fluids (Supplementary Figure 1B, C and D).

308

309 **3.2. *Specific protein signature of active TB patients and contacts in sputum***

310 For statistical analysis only those proteins were selected that were quantified in
311 the three TMTs with at least two unique peptides (Figure 1B). A nonparametric
312 Kruskal-Wallis test (supplementary information) was applied to detect statistically
313 significant differences in the individual ratios (27 uninfected/LTBI, 27 TB/LTBI and 27
314 uninfected/TB). The results are represented as volcano-plots (Figure 2 and 6). Those
315 proteins presenting a minimal 1.5-fold change and a p-value ≤ 0.0001 were considered
316 as modulated.

317 Thus, we decided to investigate if the sputum presents a characteristic proteomic
318 signature for each group of patients. The volcano plots suggested that there is an
319 accumulation of a small subset of proteins in active TB *versus* LTBI patients (Figure
320 2A, right). Examples of these proteins are haptoglobin, alpha-1-acid glycoproteins 1 and
321 2 and fibrinogen. The list of significant proteins that are abundant in active TB patients
322 was analyzed using the String 10.1 software, which indicated a strong interaction
323 network (Figure 3A). Statistical pathway analysis showed ($p\text{-value} \leq 0.001$) that most of
324 the proteins are related to platelet degranulation, inflammation and acute phase response
325 (Supplementary information 2). On the other hand, String 10.1 analysis of the proteins
326 that were decreased in TB patients versus LTBI patients (Figure 2A, left) showed a
327 weak interaction network (Figure 3B) between proteins related to endopeptidase activity
328 and taste perception.

329 Uninfected patients exhibited an accumulation of a set of proteins, compared
330 with LTBI patients (Figure 2B, right), including several basic proline-rich proteins;
331 cystatins S, D and N; carbonic anhydrase 6; and secreted-frizzle related protein 1. String
332 10.1 analysis (Figure 3C) indicated that these proteins play a role in bitter-taste
333 perception and endopeptidase activity (Supplementary information).

334 Finally, to select proteins that are specifically increased in LTBI patients, we
335 crosschecked the list of proteins that were accumulated in LTBI versus active TB
336 patients (Figure 2A, left) with the list of proteins that were accumulated in LTBI versus
337 uninfected contacts (Figure 2B, left). Only five proteins, mammaglobin-B, retinal
338 dehydrogenase 1, ectopic-p-granules protein 5 homolog and BPI-fold-containing family
339 A and B members, are accumulated in LTBI versus both uninfected and active TB
340 patients (Figure 3D). In this case, due to the low number of proteins, String 10.1
341 pathway analysis was not possible.

342 The ratios for all statistically significant proteins are represented as box-plots
343 (supplementary info and Figures 4 and 5). Proteins such as alpha-1-acid glycoproteins 1
344 and 2, haptoglobin, fibrinogen alpha and beta and protein S100P were increased in
345 active TB versus both LTBI and uninfected contacts (Figure 4). This result indicates
346 that the increase in these proteins is characteristic of active TB patients, which
347 differentiates active TB patients from uninfected and LTBI contacts. Similarly, proteins
348 that were detected to be increased in uninfected contacts versus LTBI patients, such as
349 carbonic anhydrase 6, mucin-7, cystatin S, secreted frizzle-related protein 1, basic
350 salivary protein 2 and cysteine-rich secretory protein 3 (Figure 5), were also increased
351 in uninfected versus active TB patients.

352

353 **3.3. *Specific protein signature of active TB patients and contacts in saliva***

354 A similar statistical analysis was followed to select the specific proteomic
355 signature in saliva samples. The volcano plots (Figure 6A, right) suggest an
356 accumulation of a subset of proteins in the saliva of active TB patients versus LTBI
357 patients, such as haptoglobin, fibrinogens and alpha-1-acid glycoprotein 1 and 2. String
358 analysis demonstrated a strong interaction network (Figure 7A) between proteins
359 mainly related to complement activation and acute-phase response (Supplementary
360 information 3).

361 In this case, the proteins that were decreased in the saliva of active TB versus
362 LTBI patients (Figure 6A, left) were demonstrated to share a very strong interaction
363 network (Figure 7B) and were related to carbohydrate metabolism and GTPase-
364 mediated signal transduction (Supplementary info 3).

365 Unlike the sputum, the saliva samples of uninfected and LTBI contacts did not
366 exhibit significant differences (Figure 6B). Only mucin-like protein 1 appeared to be
367 slightly accumulated in uninfected contacts (Figure 6B, right). On the other hand,

368 cathepsin G and cystatin-F (Figure 6A, left) were detected to be slightly increased in the
369 saliva of LTBI patients versus both uninfected and active TB patients (Figure 7C).

370 Box-plots representing the three ratios were prepared for all the significant
371 proteins in the saliva (Proteomic data repository and Figures 8 and 9). As observed in
372 the sputum samples, several proteins including haptoglobin, alpha-1-acid glycoprotein 1
373 and 2, immunoglobulin-gamma 4 chain, fibrinogens and dermcidin (Figure 8) were
374 specifically increased in active TB patients versus both uninfected and LTBI contacts.
375 Another set of proteins, which is represented by glutathione synthetase,
376 lactoylglutathione lyase, protein disulfide isomerase, triose-phosphate isomerase,
377 tropomyosin alpha 4 and ras GTPase-activating-like protein (Figure 9), was specifically
378 decreased in active TB patients versus both uninfected and LTBI contacts.

379

380 **3. Discussion**

381 The sputum and saliva have been, classically, a source of information for the
382 study of Mycobacteria (23) and have been used, more recently, for detection of the
383 pathogen *M. tuberculosis* (24). Separately, these two fluids have been used for
384 biomarker discovery in different diseases (25, 26); however, only few reports compare
385 both fluids (27, 28). The present study is, to our knowledge, the first to focus
386 specifically on biomarker discovery in TB through combining both types of samples.

387

388 ***Differences between saliva and sputum in terms of protein composition***

389 Analysis of the modulation of the identified proteins indicated strong differences
390 between the fluids. Global quantification results (Supplementary figure 1 A) suggested
391 that the sputum is, as expected, a more complex biological fluid than saliva. Dispersion
392 diagrams (Supplementary figure 1B, C and D) show that no correlation exists when

393 comparing quantification ratios from the saliva versus the sputum samples for all the
394 possible ratios (uninfected/LTBI, uninfected/TB and LTBI/TB), which indicates that the
395 saliva and sputum differ in protein composition and proportion.

396

397

398 *Specific proteomic signature of active TB patients*

399 Statistical analysis of the quantification ratios for the two biological fluids
400 demonstrated not only similarities but also some important differences. Proteins related
401 to cell degranulation, inflammation, acute-phase response and defense against bacteria
402 appeared to be accumulated in the sputum and saliva of the active TB patients (Figures
403 2A and 6A) compared to that in the uninfected and LTBI contacts (Figures 4 and 8).
404 This is in concordance with previous studies on the serum (29) and saliva (30) using
405 antibody-based techniques and on the serum (31) using MS-based technology.

406 Additionally, the active TB patients presented, specifically in the saliva, a very
407 significant decrease in the proteins related to sugar metabolism and GTPase-related
408 signal transduction. (Figure 6A, left). Specifically, changes in proteins related to
409 carbohydrate metabolism such as those represented in Figure 9 concur with previous
410 studies. Shin *et al.*, detected an imbalance in carbohydrate and lipid metabolism using
411 NMR-based metabolomics in the tissue of mice that were infected with MTB (32).
412 More recently, using similar approaches, Zhou *et al.* reported an increase in anaerobic
413 glycolysis rate in the sera of TB patients (33), which was characterized by elevated
414 levels of lactate and pyruvate. Our findings may represent a reflection of the disease-
415 associated metabolome adaptations of both the microbe and host, as has been previously
416 described using metabolomics (34, 35). Since the adaptive immune response systems
417 rely on various microbiota interactions to promote immune cell maturation and function

418 (36), a microbial imbalance due to active TB infection may result in a weakened
419 immune system and a loss of ability to fight disease.

420

421 *Specific proteomic signature in uninfected contacts*

422 In contrast, the uninfected contacts presented a proteomic specific signature only
423 in the sputum (Figure 2B, right) and not in the saliva (Figure 6B, right), where no
424 significant differences were found between the uninfected and LTBI contacts. In the
425 sputum, an increase was detected for a subset of proteins that play a role in
426 endopeptidase activity and bitter-taste perception in the uninfected contacts versus both
427 LTBI and active TB patients (Figure 5).

428 Bitter-taste perception has been revealed in recent years as a key regulator of the
429 innate immune system in the respiratory tract (37-39). The central core components of
430 this machinery are the type-2 receptors (T2Rs) that belong to the G-protein-coupled
431 receptor family (GPCRs). Humans are known to have at least 25 different T2R subtypes
432 that are expressed in several cell types in the airways (40). Activation of T2Rs by
433 secreted bacterial substances promotes, via calcium-dependent signaling, the secretion
434 of antimicrobial peptides in the respiratory tract during acute bacterial infection (41).
435 On the other hand, many regulators of the anti-inflammatory process during infection
436 and allergy exhibit endo-peptidase activity, especially serine and cysteine protease
437 inhibitors (42). In recent years, new therapeutic approaches have been based on the use
438 of serine protease inhibitors (43, 44).

439 Since the protein content of the sputum is a reflection of the proteins secreted by
440 the oral and nasal mucosa, we postulate that these proteins could be conferring a special
441 innate immune status that protects these individuals against MTB infection, which helps
442 to eliminate the pathogen before it reaches the lungs. Some of the proteins that were
443 detected to be increased in the sputum of the uninfected contacts have been previously

444 identified in the oral epithelia (45), and they play a role in the defense against other
445 pathogenic bacteria. Cystatin-S and carbonic anhydrase 6 are important players in the
446 bitter-taste perception machinery (46). This process has emerged in recent years as a
447 key regulator of the innate immune response. In a CAH6 null mouse model presents
448 alterations in the lower respiratory tract reduced defense capacity and the renewal of the
449 lining epithelium (46, 47). Mucin-7, a glycosylated member of the mucin family (48),
450 binds selectively to *Staphylococcus aureus* in the respiratory tract (49). Mucins are
451 produced mainly by the goblet cells of the respiratory tract and are considered as highly
452 abundant proteins in the oral and nasal fluids.

453

454 ***Specific proteomic signature of LTBI patients***

455 The discovery of new biomarkers for latent TB infection is of paramount
456 importance to accomplish the goals of the WHO "End TB" strategy. This asymptomatic
457 condition could be prolonged for many years or even the lifetime, and it still lacks a
458 sensitive and specific method for diagnosis. The saliva and sputum are, a priori, ideal
459 fluids for this search due to the ease and lack of invasiveness in sample collection. A
460 *bona fide* latent TB infection biomarker should be specific to LTBI patients and should
461 differentiate them from both uninfected and active TB patients.

462 In the present study, we have focused on proteins that were specifically
463 increased in LTBI patients, five in the sputum (Figure 3D) and only two in the saliva
464 (Figure 7C). In the sputum, ectopic p-granules 5 homolog (EPG5), retinal
465 dehydrogenase 1 and mammaglobin B are accumulated in LTBI patients versus
466 uninfected and TB patients (Figure 10). EPG5 has been revealed to be part of the
467 autolysosomal formation machinery (50). Mutations in its coding gene are the causal
468 agent of Vici syndrome, a recessive genetic condition that is characterized by

469 immunodeficiency. Retinal dehydrogenase 1 is involved in detoxification of lipid
470 aldehydes (51), and mammaglobin-B, a member of the uteroglobin family, is a steroid-
471 binding protein (52). In the saliva, we found that cathepsin-G and cystatin-F were
472 specifically accumulated in LTBI patients. Cathepsin-G is an antibacterial protein with
473 activity against Gram-negative bacteria (53), and cystatin-F has been recently identified
474 as a key factor in the differentiation process from monocytes to macrophages (54, 55).

475

476 *Considerations on the biological variability of the proteomic results*

477 All the proteins selected present significant ($p\text{-value} \leq 0.0001$) differences in the
478 quantification ratios TB/LTBI, uninfected/LTBI and uninfected/TB. However,
479 individual box-plot diagrams show, in some cases, large variability in the quantification
480 ratios, which result eventually in "outliers". Regarding this observation, it is important
481 to consider two possibilities: a) the presence of "undetectable" LTBI patients in the
482 uninfected contact group and b) the presence of nondiagnosed active TB patients in the
483 LTBI patient group. Although the criteria followed for the diagnosis and classification
484 of the patients in the EMI-TB cohort has followed the highest clinical standards, we
485 cannot totally discard the possibility of having enrolled eventual false uninfected
486 contacts or undiagnosed active TB patients.

487

488 **4. Conclusion**

489 In summary, our work represents a step forward in studying the mechanisms
490 triggered in the host by MTB infection using shotgun proteomics. Our findings that
491 indicate accumulation of acute-phase response and inflammation players in active TB
492 patients agree with previous targeted, antibody-based studies. The observation that
493 several enzymes involved in sugar metabolism were decreased in these patients is in

494 concordance with recent metabolomic studies, which indicate an imbalance in the
495 glycolytic rate of the host during active TB. We also found a small set of proteins that
496 were specifically increased in latent TB patients. In this case, further studies and larger
497 patient cohorts are now required to decipher the biological and mechanistic significance
498 of this finding.

499 Finally, proteins that were identified to be related to the innate immune response were
500 significantly overrepresented in the sputum of uninfected individuals who have been in
501 close contact with an active TB patient; this result suggests that nasal and oral mucosa
502 play a critical role in the initial entry of the pathogen. This opens a new window of
503 opportunity for modulating their expression with the use of specific adjuvants, thus
504 enhancing the innate immune response as the first barrier against infection.
505 Furthermore, it supports the importance of nasal vaccination to fight TB. Additionally,
506 since the identified proteins play biological roles in nonclassical immune processes such
507 as bitter-taste perception and endopeptidase activity, our results indicate that other
508 biological and not necessarily immunity-related processes might play a role in the MTB
509 infection or even in the reactivation of the disease. In our opinion, the special resistance
510 status of a specific individual to MTB infection could be determined for the summed
511 action of a combination of several “markers of resistance” rather than the action of only
512 one “master protective player”. To gain more biological and clinical significance, our
513 findings must be further validated in a functional model of MTB infection.

514

515 **ACKNOWLEDGMENTS**

516 We are very grateful to the patients who participated in the present study for
517 their altruistic donation of samples and to the specialized personnel of “*Complejo*
518 *Hospitalario Universitario de Pontevedra*” (SERGAS) for sample collection and

519 processing. Additionally, we would like to thank Manuel Marcos (Proteomics Facility,
520 University of Vigo, CACTI, Vigo, Spain) for shotgun data acquisition and Alexandre
521 Alonso-Fernández (CSIC) for help with statistics. CSIC (M.C.), CINBIO (A.G.F.) and
522 SGUL (R.R.) groups are part of the Eliciting Mucosal Immunity in Tuberculosis (EMI-
523 TB) consortium. EMI-TB is a European Union Horizon 2020 funded action (Grant No
524 643558) that is focused on selecting and developing a novel vaccine candidate for
525 tuberculosis (TB). We would like to thank all its members, especially, Elina Garet and
526 Silvia Lorenzo-Abalde, for their collaboration and for helpful discussion and
527 suggestions.

528

529 **REFERENCES**

- 530 1. Scully T. Tuberculosis. *Nature*. 2013;502(7470):S1.
531 2. Sandgren A, Vonk Noordegraaf-Schouten M, van Kessel F, Stuurman A, Oordt-Speets A,
532 van der Werf MJ. Initiation and completion rates for latent tuberculosis infection treatment: a
533 systematic review. *BMC Infect Dis*. 2016;16:204.
534 3. Zumla A, George A, Sharma V, Herbert RH, Baroness Masham of I, Oxley A, et al. The
535 WHO 2014 global tuberculosis report--further to go. *Lancet Glob Health*. 2015;3(1):e10-2.
536 4. Zink AR, Sola C, Reischl U, Grabner W, Rastogi N, Wolf H, et al. Characterization of
537 *Mycobacterium tuberculosis* complex DNAs from Egyptian mummies by spoligotyping. *J Clin*
538 *Microbiol*. 2003;41(1):359-67.
539 5. Fennelly KP, Martyny JW, Fulton KE, Orme IM, Cave DM, Heifets LB. Cough-generated
540 aerosols of *Mycobacterium tuberculosis*: a new method to study infectiousness. *Am J Respir*
541 *Crit Care Med*. 2004;169(5):604-9.
542 6. Via LE, Lin PL, Ray SM, Carrillo J, Allen SS, Eum SY, et al. Tuberculous granulomas are
543 hypoxic in guinea pigs, rabbits, and nonhuman primates. *Infect Immun*. 2008;76(6):2333-40.
544 7. Adams DO. The granulomatous inflammatory response. A review. *Am J Pathol*.
545 1976;84(1):164-92.
546 8. Ahmad S. Pathogenesis, immunology, and diagnosis of latent *Mycobacterium*
547 *tuberculosis* infection. *Clin Dev Immunol*. 2011;2011:814943.
548 9. Kuipers K, van Selm S, van Opzeeland F, Langereis JD, Verhagen LM, Diavatopoulos DA,
549 et al. Genetic background impacts vaccine-induced reduction of pneumococcal colonization.
550 *Vaccine*. 2017;35(39):5235-41.
551 10. Conceicao EL, Nascimento-Sampaio FS, Schwingel PA, Oliveira ES, Rocha MS, Vieira I, et
552 al. Revisiting the Heterogeneous IFN-gamma Response of Bacille of Calmette-Guerin (BCG)-
553 Revaccinated Healthy Volunteers in a Randomized Controlled Trial: Effect of the Body Mass
554 Index and of the IFNG+874 A/T Polymorphism. *PLoS One*. 2016;11(7):e0160149.
555 11. Cousineau SV, Alizon S. Parasite evolution in response to sex-based host heterogeneity
556 in resistance and tolerance. *J Evol Biol*. 2014;27(12):2753-66.
557 12. Poland GA, Ovsyannikova IG, Kennedy RB. Personalized vaccinology: A review. *Vaccine*.
558 2017.
559 13. Delany I, Rappuoli R, De Gregorio E. Vaccines for the 21st century. *EMBO Mol Med*.
560 2014;6(6):708-20.
561 14. Bethony JM, Cole RN, Guo X, Kamhawi S, Lightowers MW, Loukas A, et al. Vaccines to
562 combat the neglected tropical diseases. *Immunol Rev*. 2011;239(1):237-70.
563 15. Hawn TR, Day TA, Scriba TJ, Hatherill M, Hanekom WA, Evans TG, et al. Tuberculosis
564 vaccines and prevention of infection. *Microbiol Mol Biol Rev*. 2014;78(4):650-71.
565 16. Roy A, Eisenhut M, Harris RJ, Rodrigues LC, Sridhar S, Habermann S, et al. Effect of BCG
566 vaccination against *Mycobacterium tuberculosis* infection in children: systematic review and
567 meta-analysis. *BMJ*. 2014;349:g4643.
568 17. Uplekar M, Raviglione M. WHO's End TB Strategy: From stopping to ending the global
569 TB epidemic. *Indian J Tuberc*. 2015;62(4):196-9.
570 18. Csoz E, Kallo G, Markus B, Deak E, Csutak A, Tozser J. Quantitative body fluid
571 proteomics in medicine - A focus on minimal invasiveness. *J Proteomics*. 2017;153:30-43.
572 19. Legrain P, Aebersold R, Archakov A, Bairoch A, Bala K, Beretta L, et al. The human
573 proteome project: current state and future direction. *Mol Cell Proteomics*. 2011;10(7):M111
574 009993.
575 20. Rai AJ, Gelfand CA, Haywood BC, Warunek DJ, Yi J, Schuchard MD, et al. HUPO Plasma
576 Proteome Project specimen collection and handling: towards the standardization of
577 parameters for plasma proteome samples. *Proteomics*. 2005;5(13):3262-77.

- 578 21. González-Martín J, García-García JM, Anibarro L, Vidal R, Esteban J, Blanquer R, et al.
579 [Consensus document on the diagnosis, treatment and prevention of tuberculosis]. *Enferm*
580 *Infecc Microbiol Clin*. 2010;28(5):297.e1-20.
- 581 22. Ruhl S. The scientific exploration of saliva in the post-proteomic era: from database
582 back to basic function. *Expert Rev Proteomics*. 2012;9(1):85-96.
- 583 23. Yeager H, Jr., Lacy J, Smith LR, LeMaistre CA. Quantitative studies of mycobacterial
584 populations in sputum and saliva. *Am Rev Respir Dis*. 1967;95(6):998-1004.
- 585 24. Holani AG, Ganvir SM, Shah NN, Bansode SC, Shende I, Jawade R, et al. Demonstration
586 of mycobacterium tuberculosis in sputum and saliva smears of tuberculosis patients using ziehl
587 neelsen and flurochrome staining- a comparative study. *J Clin Diagn Res*. 2014;8(7):ZC42-5.
- 588 25. Amado FM, Ferreira RP, Vitorino R. One decade of salivary proteomics: current
589 approaches and outstanding challenges. *Clin Biochem*. 2013;46(6):506-17.
- 590 26. Titz B, Sewer A, Schneider T, Elamin A, Martin F, Dijon S, et al. Alterations in the
591 sputum proteome and transcriptome in smokers and early-stage COPD subjects. *J Proteomics*.
592 2015;128:306-20.
- 593 27. Wilffert D, Donzelli R, Asselman A, Hermans J, Govorukhina N, Ten Hacken NH, et al.
594 Quantitative antibody-free LC-MS/MS analysis of sTRAIL in sputum and saliva at the sub-ng/mL
595 level. *J Chromatogr B Analyt Technol Biomed Life Sci*. 2016;1032:205-10.
- 596 28. Gaber F, Daham K, Higashi A, Higashi N, Gulich A, Delin I, et al. Increased levels of
597 cysteinyl-leukotrienes in saliva, induced sputum, urine and blood from patients with aspirin-
598 intolerant asthma. *Thorax*. 2008;63(12):1076-82.
- 599 29. De Groote MA, Sterling DG, Hraha T, Russell T, Green LS, Wall K, et al. Discovery and
600 Validation of a Six-Marker Serum Protein Signature for the Diagnosis of Active Pulmonary
601 Tuberculosis. *J Clin Microbiol*. 2017.
- 602 30. Phalane KG, Kriel M, Loxton AG, Menezes A, Stanley K, van der Spuy GD, et al.
603 Differential expression of host biomarkers in saliva and serum samples from individuals with
604 suspected pulmonary tuberculosis. *Mediators Inflamm*. 2013;2013:981984.
- 605 31. Achkar JM, Cortes L, Croteau P, Yanofsky C, Mentinova M, Rajotte I, et al. Host Protein
606 Biomarkers Identify Active Tuberculosis in HIV Uninfected and Co-infected Individuals.
607 *EBioMedicine*. 2015;2(9):1160-8.
- 608 32. Shin JH, Yang JY, Jeon BY, Yoon YJ, Cho SN, Kang YH, et al. (1)H NMR-based
609 metabolomic profiling in mice infected with Mycobacterium tuberculosis. *J Proteome Res*.
610 2011;10(5):2238-47.
- 611 33. Zhou A, Ni J, Xu Z, Wang Y, Lu S, Sha W, et al. Application of (1)h NMR spectroscopy-
612 based metabolomics to sera of tuberculosis patients. *J Proteome Res*. 2013;12(10):4642-9.
- 613 34. Preez ID, Luies L, Loots DT. Metabolomics biomarkers for tuberculosis diagnostics:
614 current status and future objectives. *Biomark Med*. 2017;11(2):179-94.
- 615 35. Luies L, du Preez I, Loots DT. The role of metabolomics in tuberculosis treatment
616 research. *Biomark Med*. 2017;11(11):1017-29.
- 617 36. Luies L, Reenen MV, Ronacher K, Walzl G, Loots DT. Predicting tuberculosis treatment
618 outcome using metabolomics. *Biomark Med*. 2017;11(12):1057-67.
- 619 37. Workman AD, Palmer JN, Adappa ND, Cohen NA. The Role of Bitter and Sweet Taste
620 Receptors in Upper Airway Immunity. *Curr Allergy Asthma Rep*. 2015;15(12):72.
- 621 38. Lee RJ, Cohen NA. Role of the bitter taste receptor T2R38 in upper respiratory infection
622 and chronic rhinosinusitis. *Curr Opin Allergy Clin Immunol*. 2015;15(1):14-20.
- 623 39. Freund JR, Mansfield CJ, Doghramji LJ, Adappa ND, Palmer JN, Kennedy DW, et al.
624 Activation of airway epithelial bitter taste receptors by Pseudomonas aeruginosa quinolones
625 modulates calcium, cyclic-AMP, and nitric oxide signaling. *J Biol Chem*. 2018;293(25):9824-40.
- 626 40. Chandrashekar J, Hoon MA, Ryba NJ, Zuker CS. The receptors and cells for mammalian
627 taste. *Nature*. 2006;444(7117):288-94.

- 628 41. Lee RJ, Kofonow JM, Rosen PL, Siebert AP, Chen B, Doghramji L, et al. Bitter and sweet
629 taste receptors regulate human upper respiratory innate immunity. *J Clin Invest.*
630 2014;124(3):1393-405.
- 631 42. Shigetomi H, Onogi A, Kajiwara H, Yoshida S, Furukawa N, Haruta S, et al. Anti-
632 inflammatory actions of serine protease inhibitors containing the Kunitz domain. *Inflamm Res.*
633 2010;59(9):679-87.
- 634 43. Bahgat MM, Blazejewska P, Schughart K. Inhibition of lung serine proteases in mice: a
635 potentially new approach to control influenza infection. *Virology.* 2011;8:27.
- 636 44. Saw S, Kale SL, Arora N. Serine protease inhibitor attenuates ovalbumin induced
637 inflammation in mouse model of allergic airway disease. *PLoS One.* 2012;7(7):e41107.
- 638 45. Gibbins HL, Proctor GB, Yakubov GE, Wilson S, Carpenter GH. Concentration of salivary
639 protective proteins within the bound oral mucosal pellicle. *Oral Dis.* 2014;20(7):707-13.
- 640 46. Patrikainen M, Pan P, Kuleskaya N, Voikar V, Parkkila S. The role of carbonic
641 anhydrase VI in bitter taste perception: evidence from the Car6(-)/(-) mouse model. *J Biomed*
642 *Sci.* 2014;21:82.
- 643 47. Tolvanen ME, Ortutay C, Barker HR, Aspatwar A, Patrikainen M, Parkkila S. Analysis of
644 evolution of carbonic anhydrases IV and XV reveals a rich history of gene duplications and a
645 new group of isozymes. *Bioorg Med Chem.* 2013;21(6):1503-10.
- 646 48. Xu D, Pavlidis P, Taskent RO, Alachiotis N, Flanagan C, DeGiorgio M, et al. Archaic
647 Hominin Introgression in Africa Contributes to Functional Salivary MUC7 Genetic Variation.
648 *Mol Biol Evol.* 2017;34(10):2704-15.
- 649 49. Heo SM, Choi KS, Kazim LA, Reddy MS, Haase EM, Scannapieco FA, et al. Host defense
650 proteins derived from human saliva bind to *Staphylococcus aureus*. *Infect Immun.*
651 2013;81(4):1364-73.
- 652 50. Cullup T, Kho AL, Dionisi-Vici C, Brandmeier B, Smith F, Urry Z, et al. Recessive
653 mutations in EPG5 cause Vici syndrome, a multisystem disorder with defective autophagy. *Nat*
654 *Genet.* 2013;45(1):83-7.
- 655 51. Xiao T, Shoeb M, Siddiqui MS, Zhang M, Ramana KV, Srivastava SK, et al. Molecular
656 cloning and oxidative modification of human lens ALDH1A1: implication in impaired
657 detoxification of lipid aldehydes. *J Toxicol Environ Health A.* 2009;72(9):577-84.
- 658 52. Lehrer RI, Xu G, Abduragimov A, Dinh NN, Qu XD, Martin D, et al. Lipophilin, a novel
659 heterodimeric protein of human tears. *FEBS Lett.* 1998;432(3):163-7.
- 660 53. Wasiluk KR, Skubitz KM, Gray BH. Comparison of granule proteins from human
661 polymorphonuclear leukocytes which are bactericidal toward *Pseudomonas aeruginosa*. *Infect*
662 *Immun.* 1991;59(11):4193-200.
- 663 54. Kos J, Nanut MP, Prunk M, Sabotic J, Dautovic E, Jewett A. Cystatin F as a regulator of
664 immune cell cytotoxicity. *Cancer Immunol Immunother.* 2018;67(12):1931-8.
- 665 55. Dautovic E, Perisic Nanut M, Softic A, Kos J. The transcription factor C/EBP alpha
666 controls the role of cystatin F during the differentiation of monocytes to macrophages. *Eur J*
667 *Cell Biol.* 2018;97(7):463-73.

668

669

Table 1. Demographic summary of the patients included in the EMI-TB cohort.

Patient Type	Gender	Age (mean \pmSD)	Contact score (mean \pmSD)
Active TB (n=26)	Female (15.4%), male (84,6%)	41.3 \pm 13.9	n.a.
LTBI (n=29)	Female (41.3%), male (58.7%)	47.4 \pm 14.7	10.9 \pm 2.6
Uninfected (n=44)	Female (52,3%), male (47,7%)	40.0 \pm 15.2	9.6 \pm 2.3

Table 2. Individual samples used for the shotgun proteomic study.

<i>TMT 1 Saliva</i>	<i>SAMPLE</i>	PO-20	PO-28	PO-38	PO-24	PO-27	PO-36	PO-21	PO-26	PO-44	Standard
	<i>Group</i>	TB	TB	TB	LTI	LTI	LTI	nonLTI	nonLTI	nonLTI	mix
	<i>TMT-Label</i>	126	127N	127C	128N	128C	129N	129C	130N	130C	131
<i>TMT 2 Saliva</i>	<i>SAMPLE</i>	PO-51	PO-61	PO-63	PO-30	PO-42	PO-46	PO-22	PO-52	PO-62	Standard
	<i>Group</i>	TB	TB	TB	LTI	LTI	LTI	nonLTI	nonLTI	nonLTI	mix
	<i>TMT-Label</i>	126	127N	127C	128N	128C	129N	129C	130N	130C	131
<i>TMT 3 Saliva</i>	<i>SAMPLE</i>	PO-53	PO-18	PO-64	PO-37	PO-41	PO-54	PO-23	PO-34	PO-65	Standard
	<i>Group</i>	TB	TB	TB	LTI	LTI	LTI	nonLTI	nonLTI	nonLTI	mix
	<i>TMT-Label</i>	126	127N	127C	128N	128C	129N	129C	130N	130C	131
<i>TMT 1 Sputum</i>	<i>SAMPLE</i>	PO-20	PO-28	PO-38	PO-24	PO-27	PO-36	PO-21	PO-23	PO-52	Standard
	<i>Group</i>	TB	TB	TB	LTI	LTI	LTI	nonLTI	nonLTI	nonLTI	mix
	<i>TMT-Label</i>	126	127N	127C	128N	128C	129N	129C	130N	130C	131
<i>TMT 2 Sputum</i>	<i>SAMPLE</i>	PO-16	PO-40	PO-51	PO-30	PO-41	PO-54	PO-34	PO-57	PO-59	Standard
	<i>Group</i>	TB	TB	TB	LTI	LTI	LTI	nonLTI	nonLTI	nonLTI	mix
	<i>TMT-Label</i>	126	127N	127C	128N	128C	129N	129C	130N	130C	131
<i>TMT 3 Sputum</i>	<i>SAMPLE</i>	PO-32	PO-53	PO-63	PO-31	PO-42	PO-55	PO-62	PO-65	PO-68	Standard
	<i>Group</i>	TB	TB	TB	LTI	LTI	LTI	nonLTI	nonLTI	nonLTI	mix
	<i>TMT-Label</i>	126	127N	127C	128N	128C	129N	129C	130N	130C	131

FIGURE LEGENDS

Fig 1: Schematic design and workflow of the proteomic shotgun analysis.

Fig 2: Volcano-plot representations of the statistical analysis in sputum. Analysis of the active TB/LTBI (A) and uninfected/LTBI (B) ratios for all the 469 proteins quantified in the three TMTs with at least two unique peptides.

Fig 3: String 10.1 interaction pathway analysis of the specific protein signature of active TB patients (A and B), uninfected contacts (C) and LTBI patients (D) in sputum samples.

Fig 4: Box-plot representations of the quantification ratios of selected proteins accumulated in the sputum of active TB patients.

Fig 5: Box-plot representations of the quantification ratios of selected proteins accumulated in the sputum of uninfected contacts.

Fig 6: Volcano-plot representations of the statistical analysis in saliva. Analysis of the active TB/LTBI (A) and uninfected/LTBI (B) ratios for all the 379 proteins quantified in the three TMTs with at least two unique peptides.

Fig 7: String 10.1 interaction pathway analysis of the specific protein signature of active TB (A and B) and LTBI (C) patients in saliva samples.

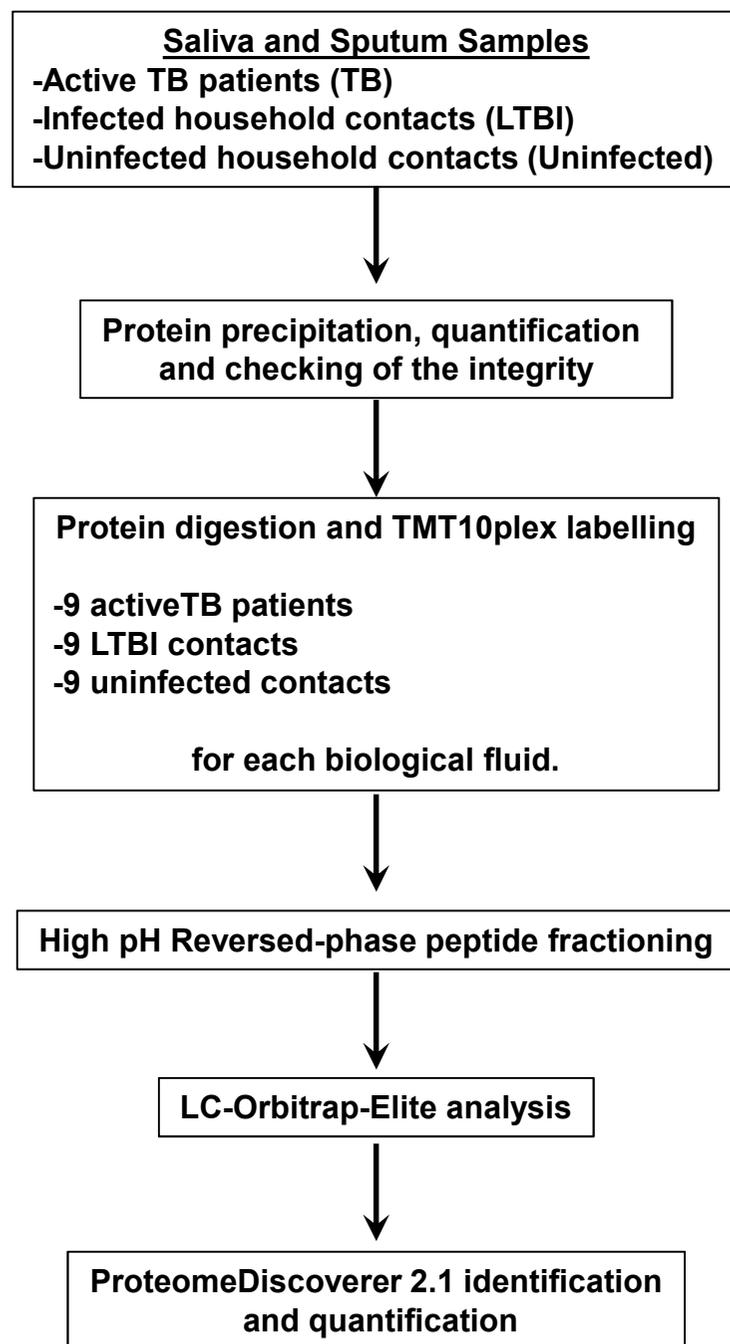
Fig 8: Box-plot representations of the quantification ratios of selected proteins accumulated in the saliva of active TB patients.

Fig 9: Box-plot representations of the quantification ratios of selected proteins decreased in the saliva of active TB patients.

Fig 10: Box-plot representations of the quantification ratios of selected proteins accumulated in the sputum and saliva of LTBI patients.

Figure 1
[Click here to download Figure: new Figure 1.pdf](#)

A



B

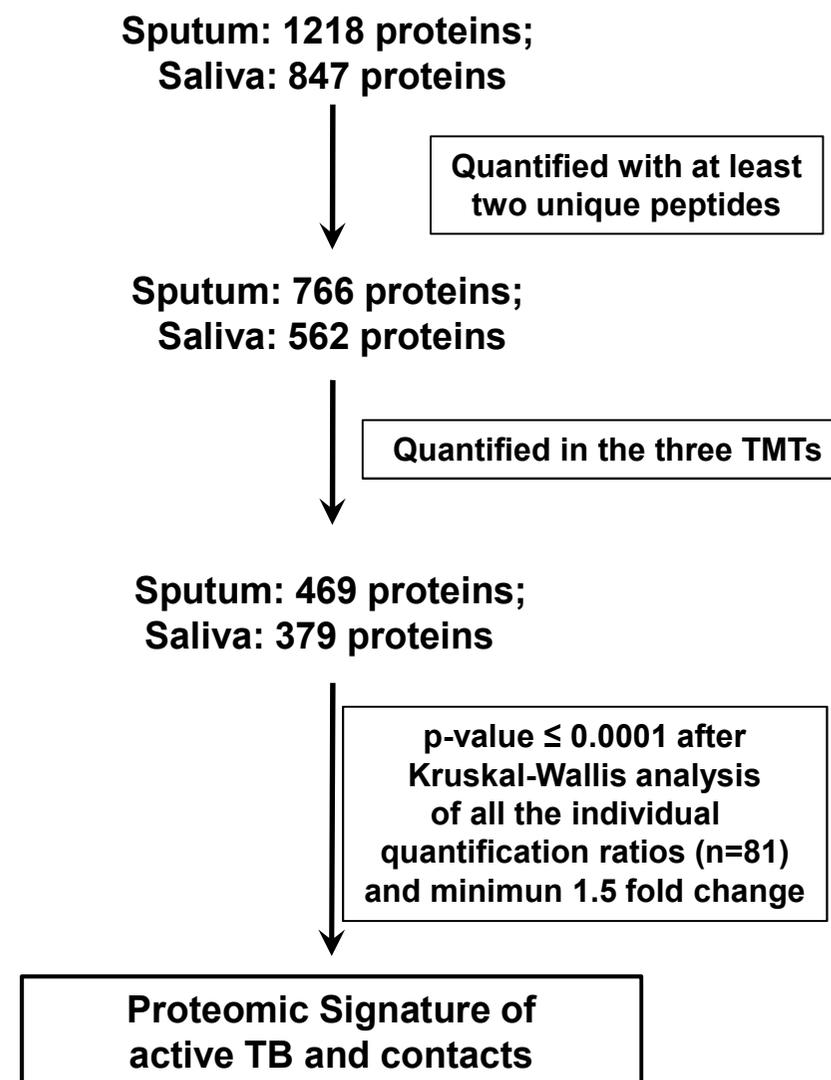
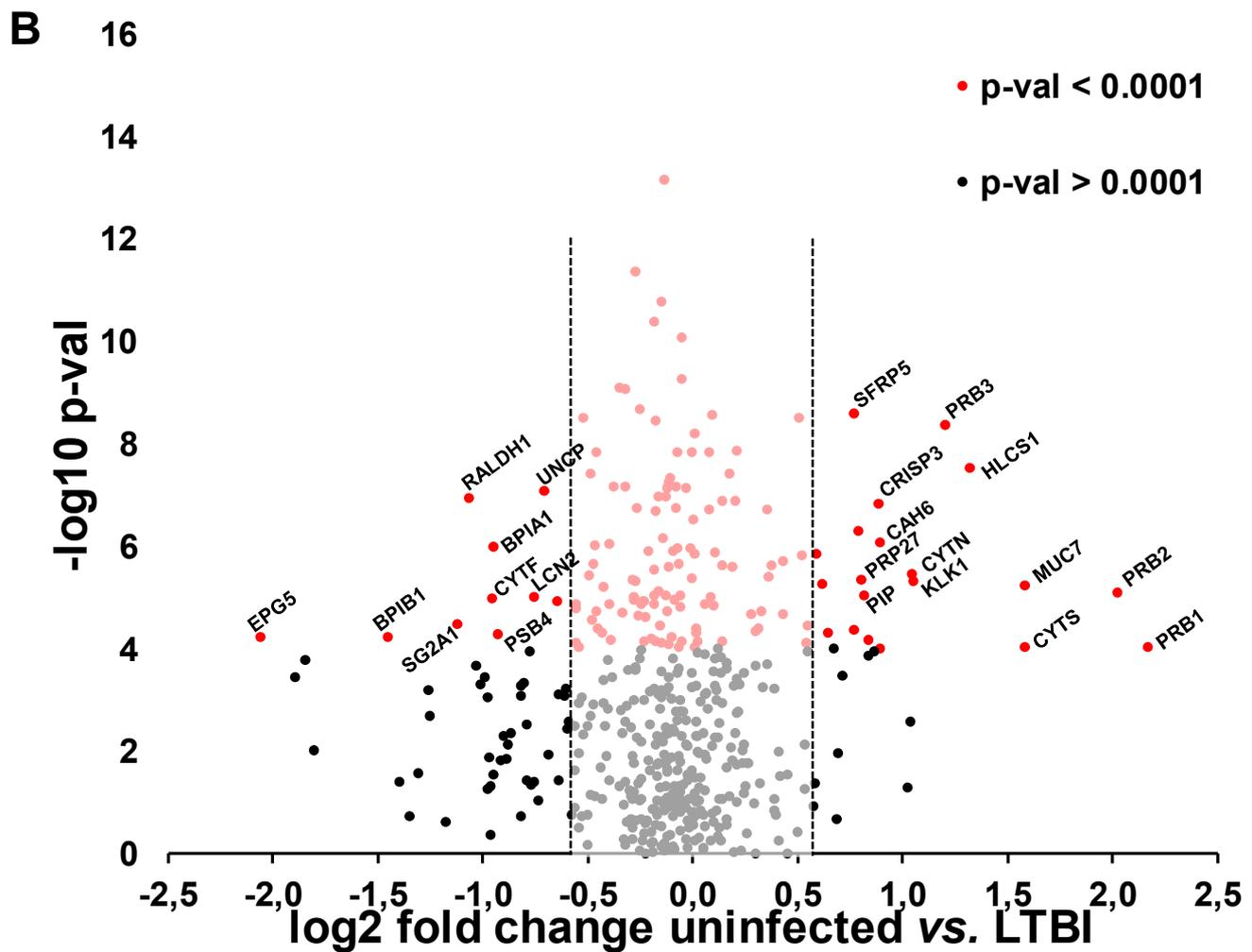
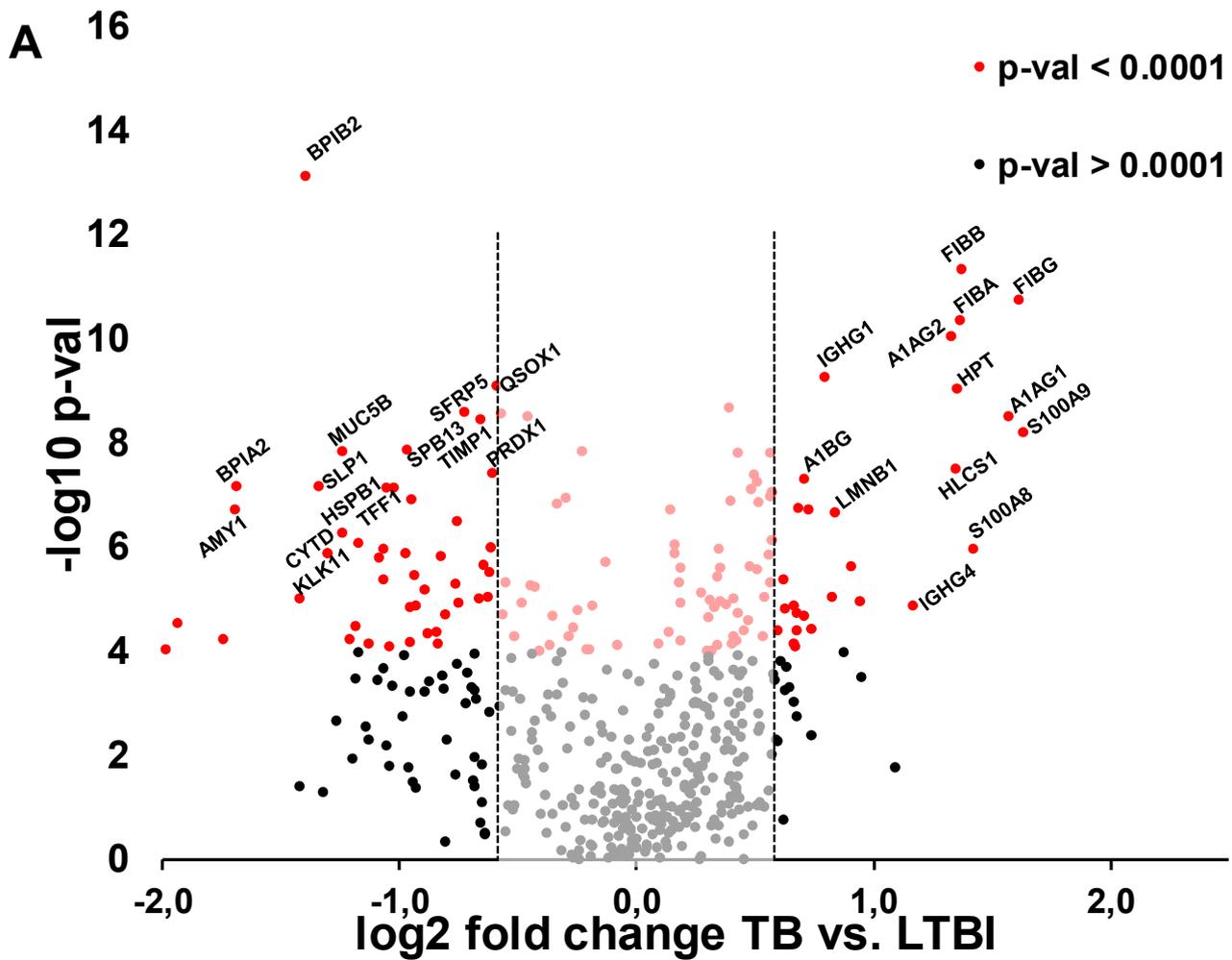
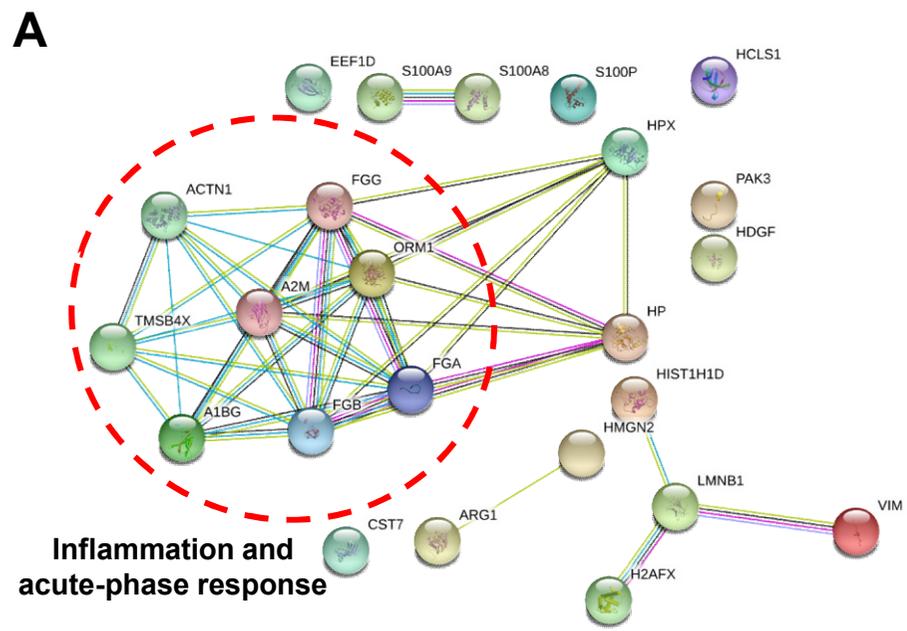
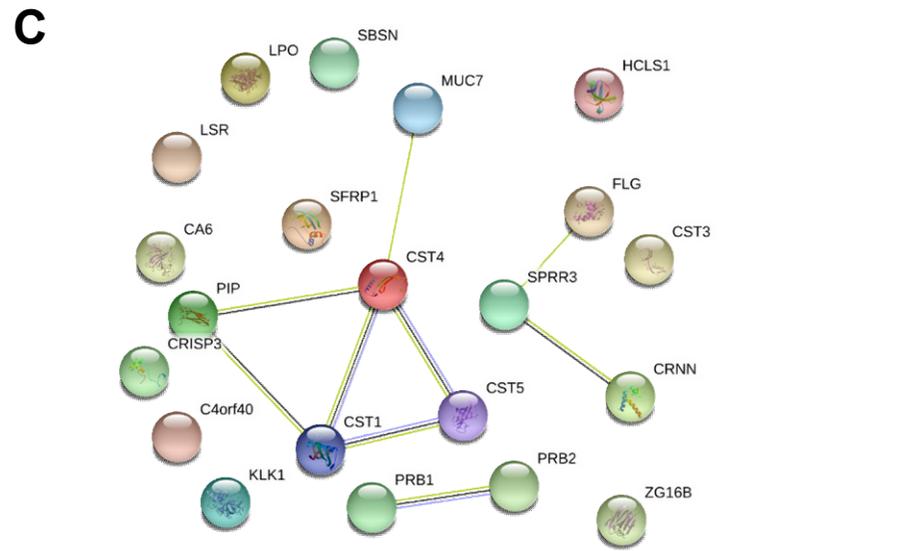


Figure 2
Click here to download Figure: new Figure 2.pdf

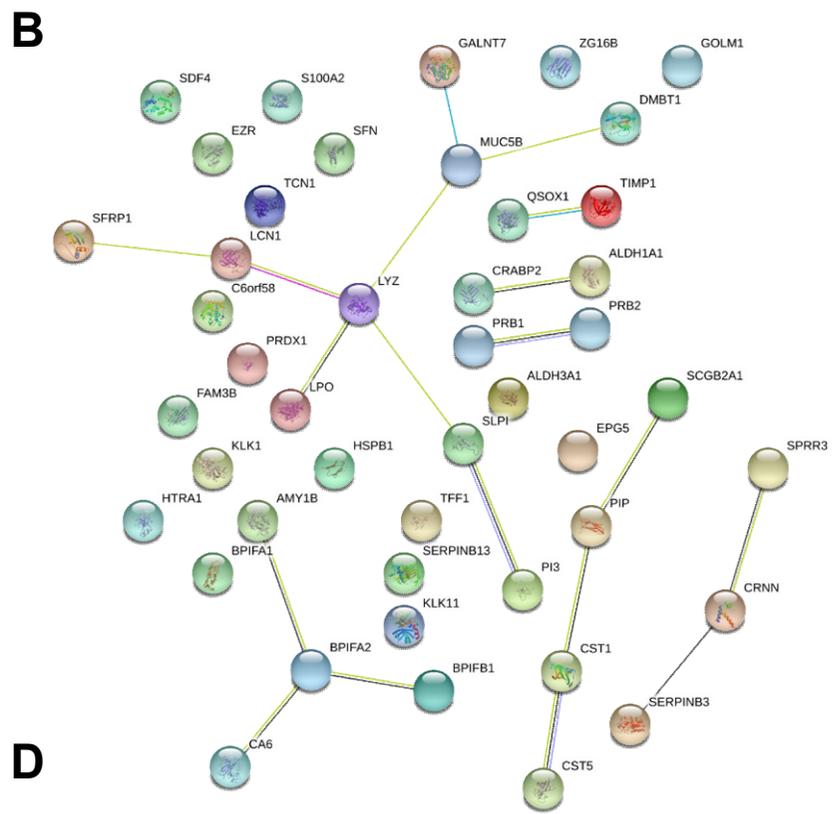




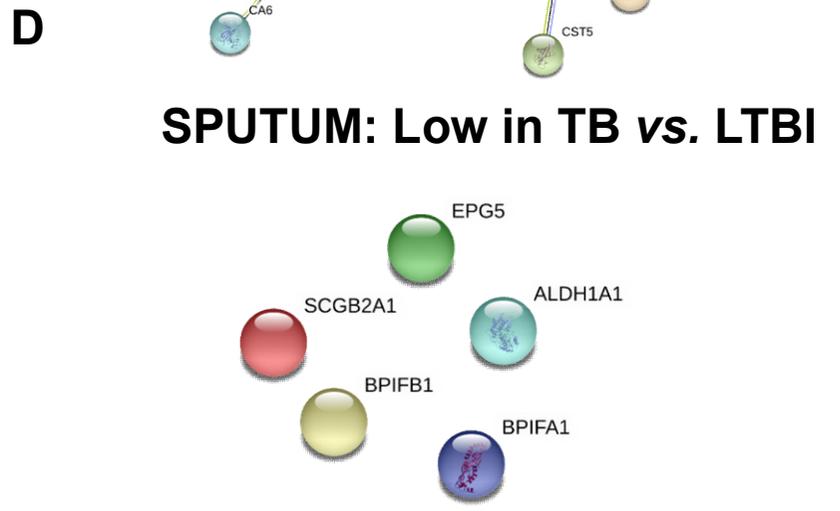
SPUTUM: High in TB vs. LTBI



SPUTUM: High in uninfected vs. LTBI



SPUTUM: Low in TB vs. LTBI



SPUTUM: High in LTBI vs. TB/uninfected

Figure 4

[Click here to download Figure: new Figure 4.pdf](#)

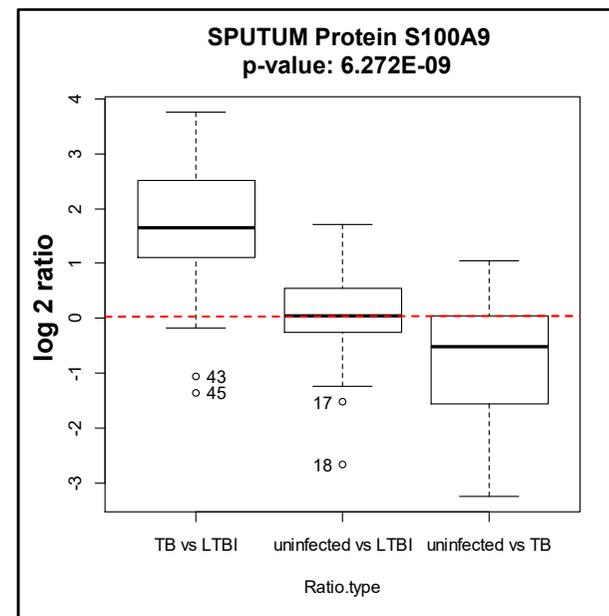
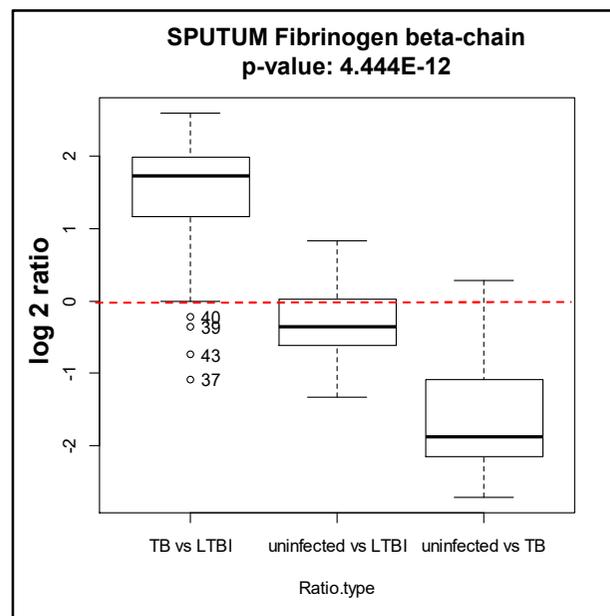
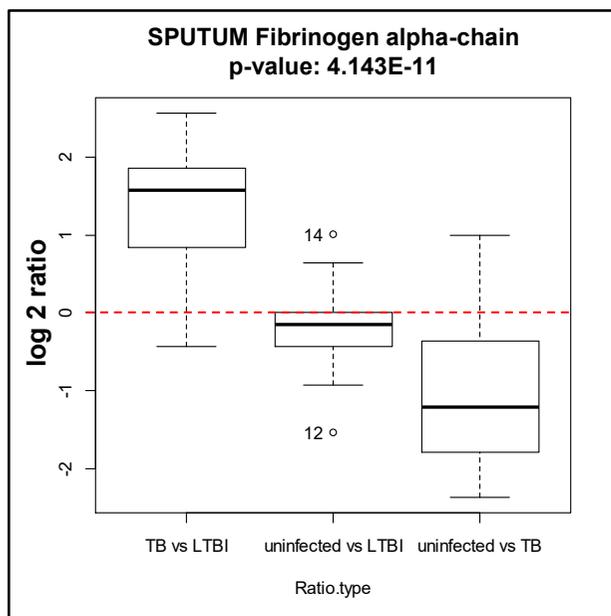
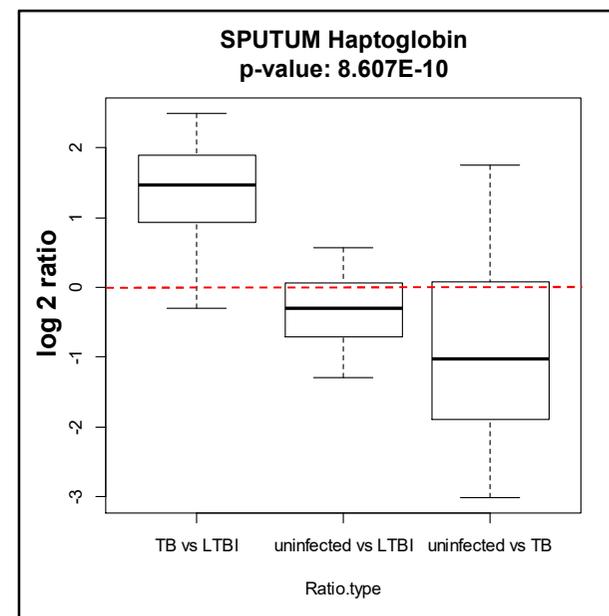
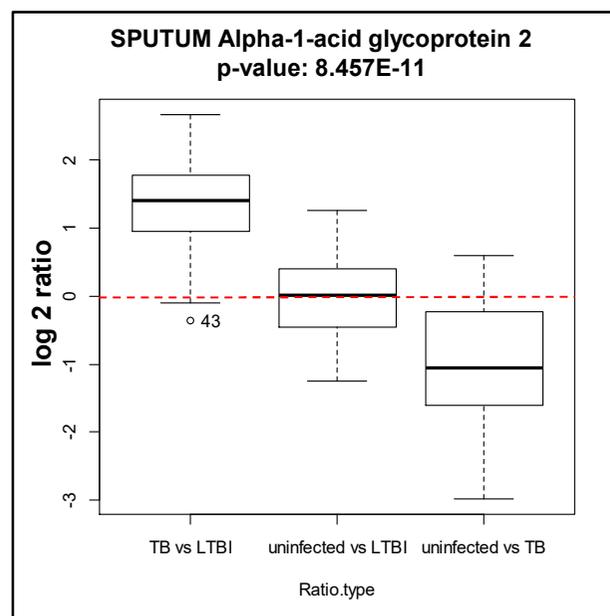
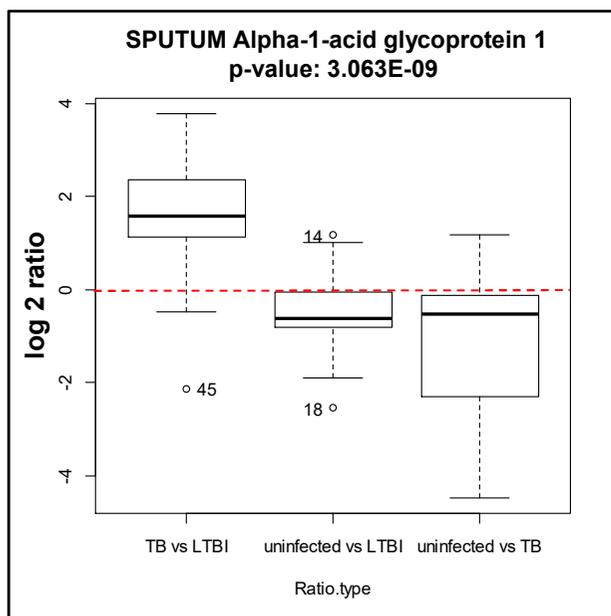


Figure 5

[Click here to download Figure: new Figure 5.pdf](#)

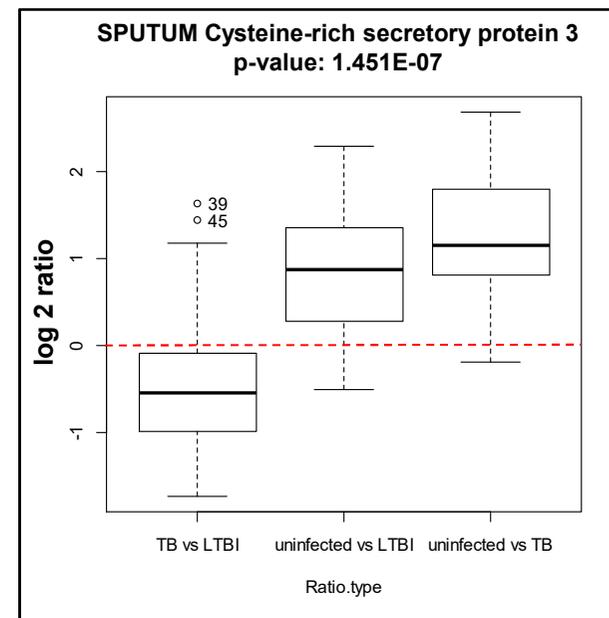
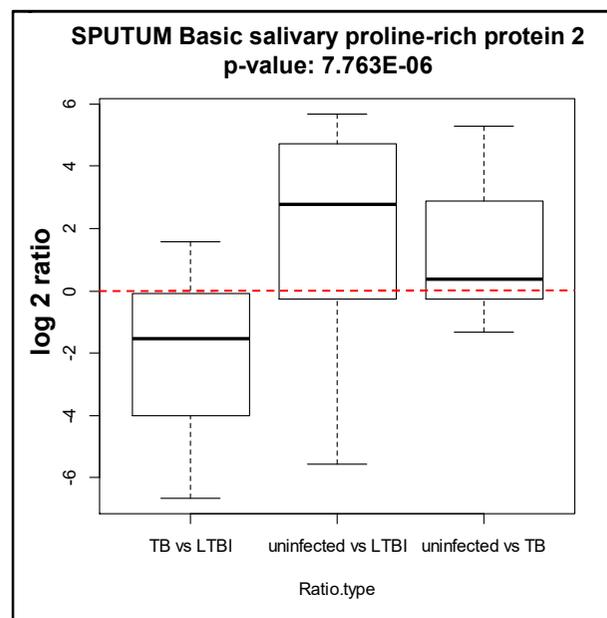
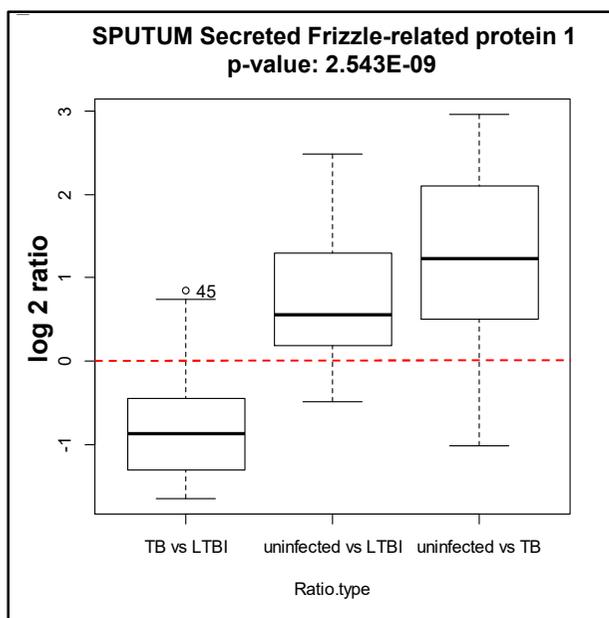
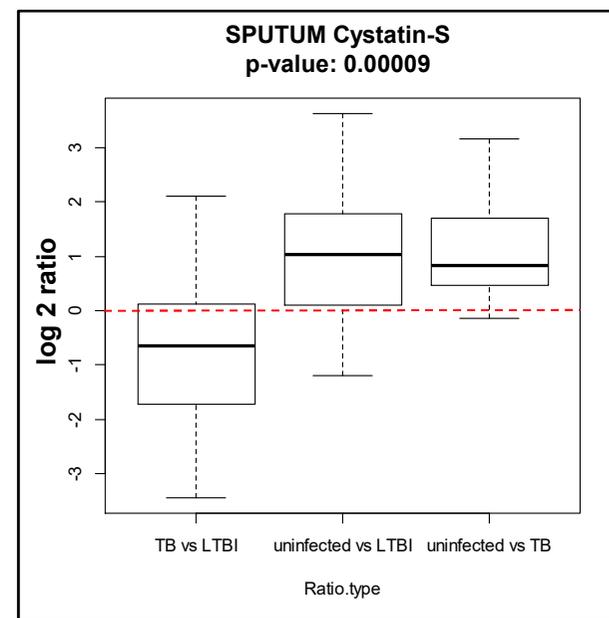
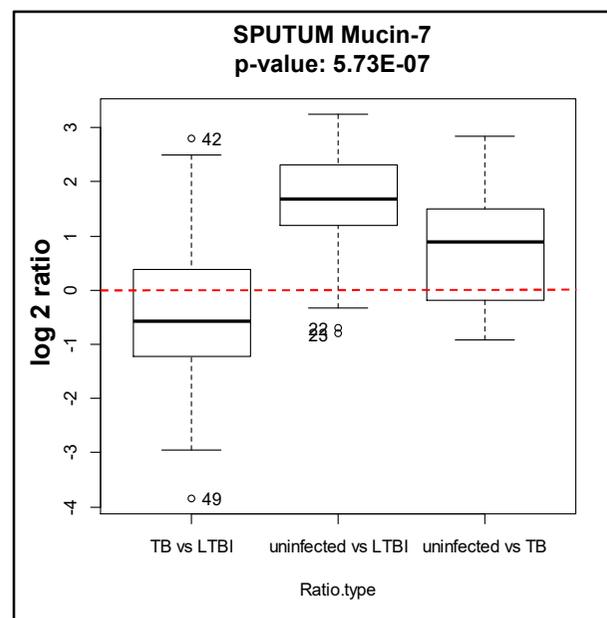
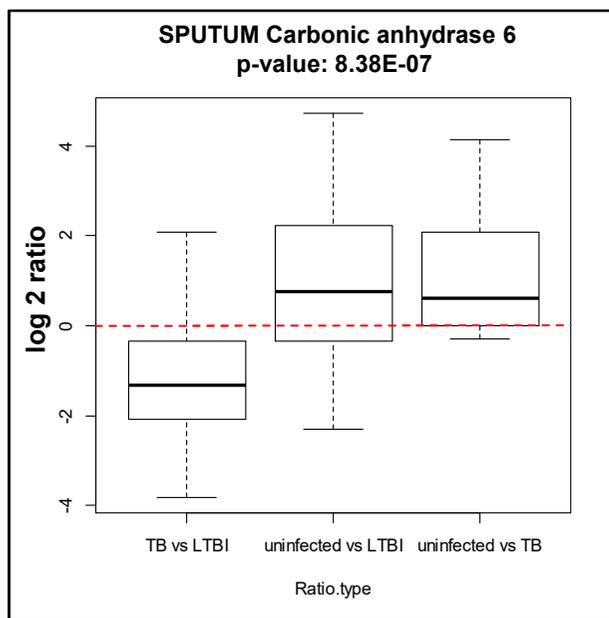


Figure 6
[Click here to download Figure: new Figure 6.pdf](#)

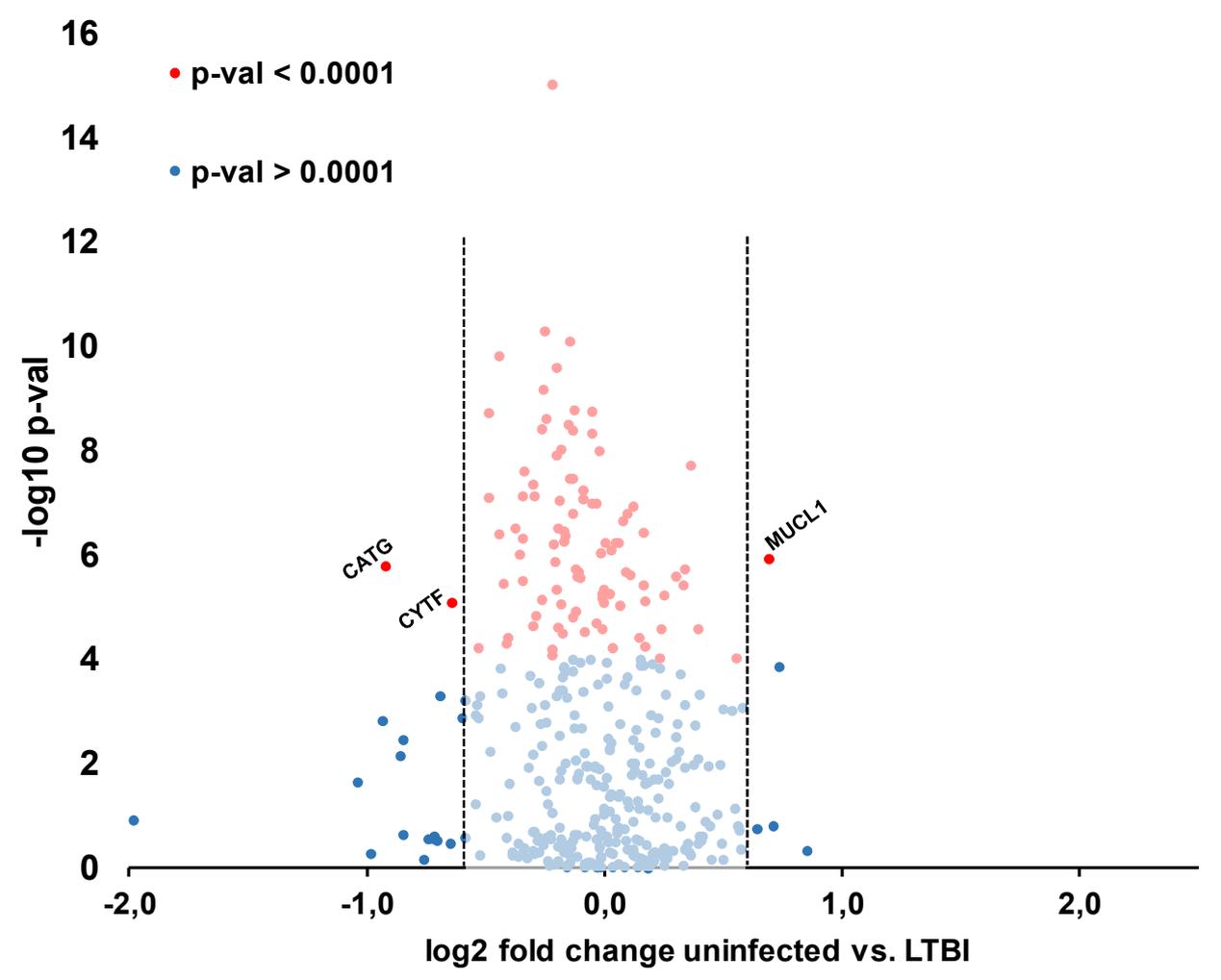
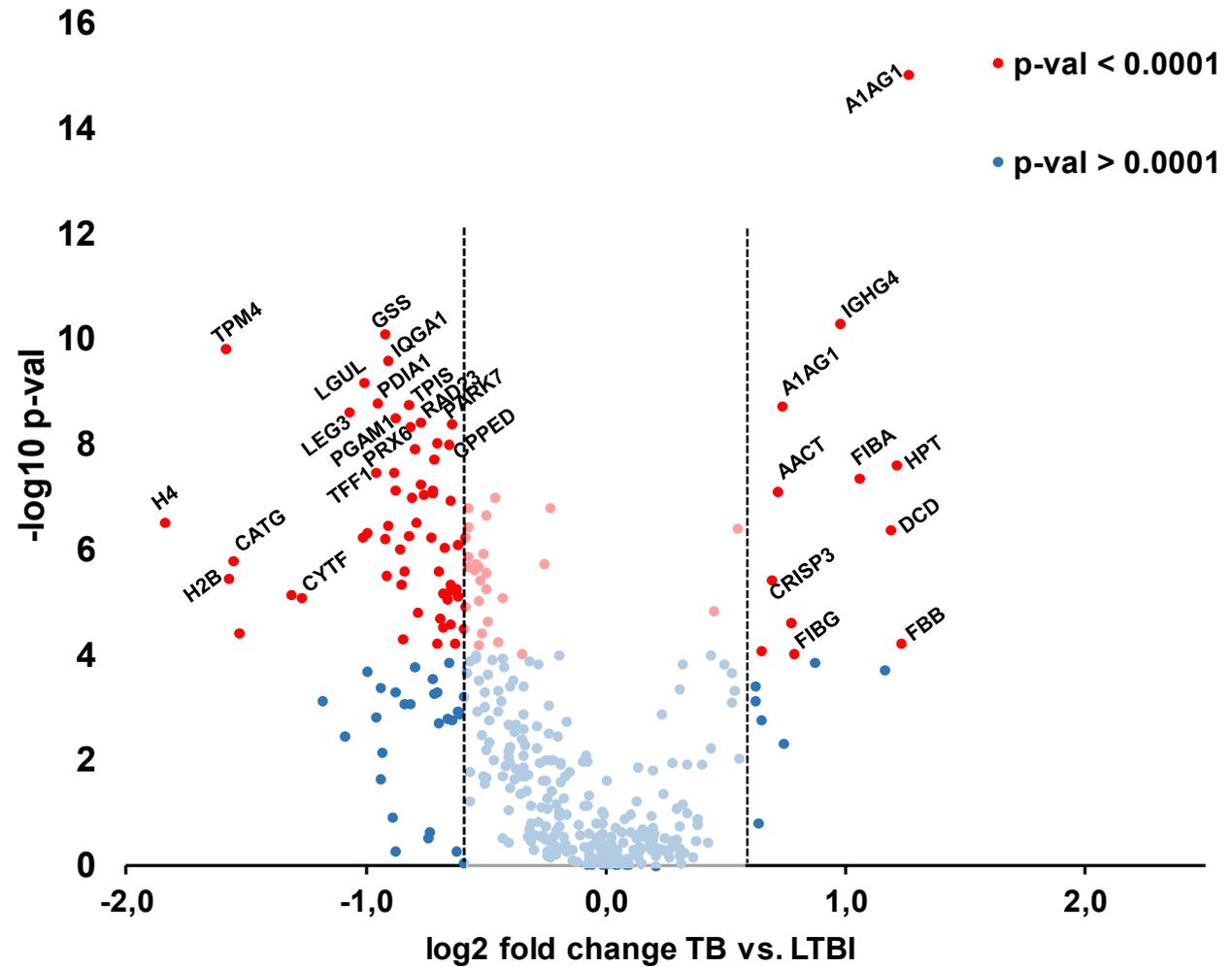


Figure 7

[Click here to download Figure: new Figure 7.pdf](#)

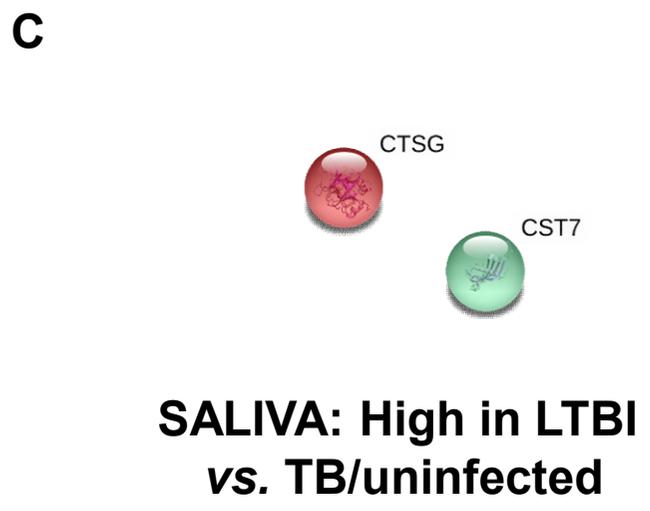
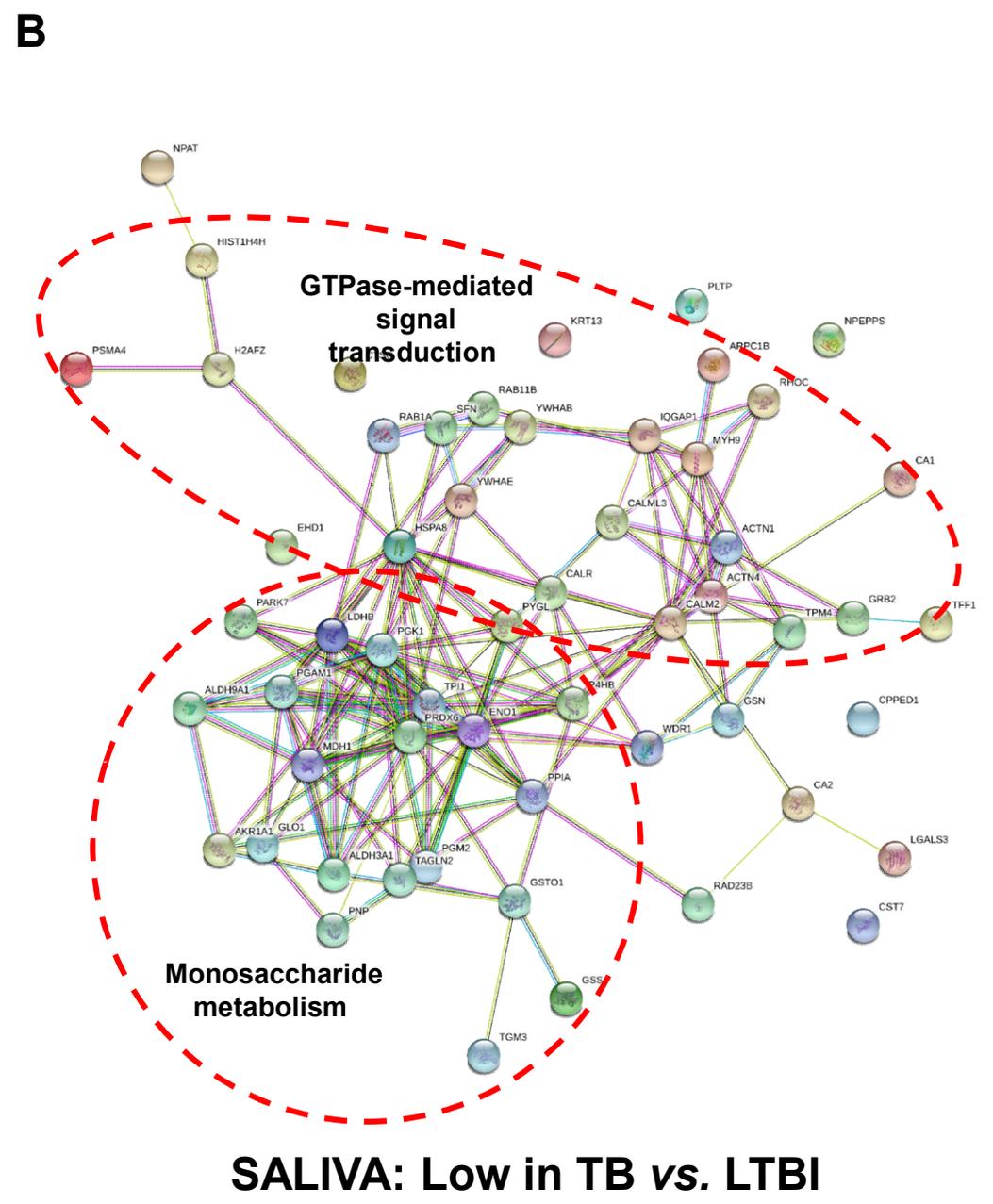
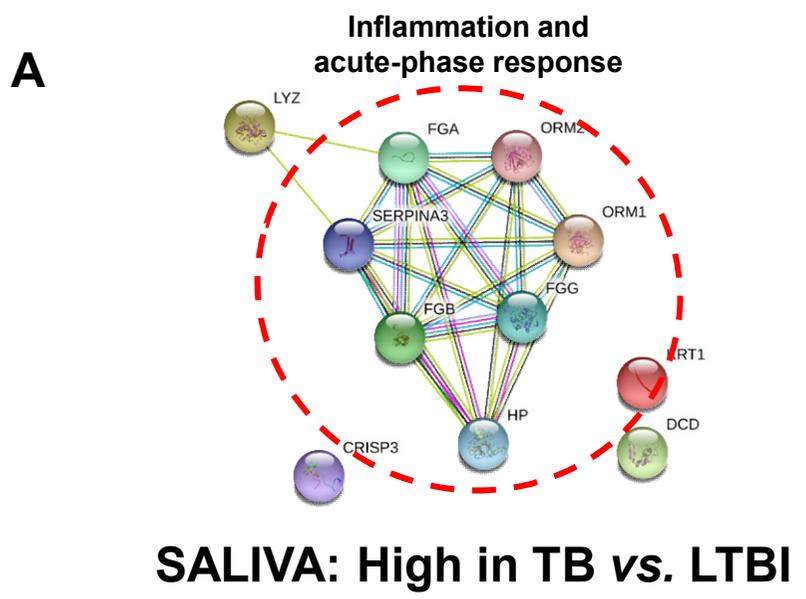


Figure 8

[Click here to download Figure: new Figure 8.pdf](#)

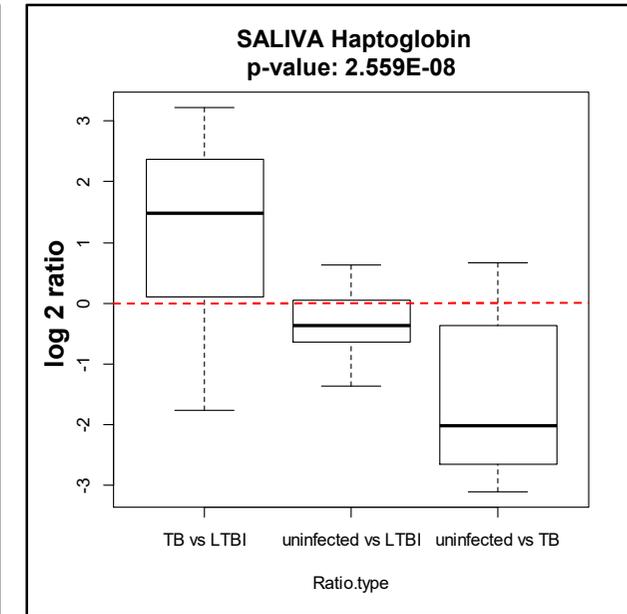
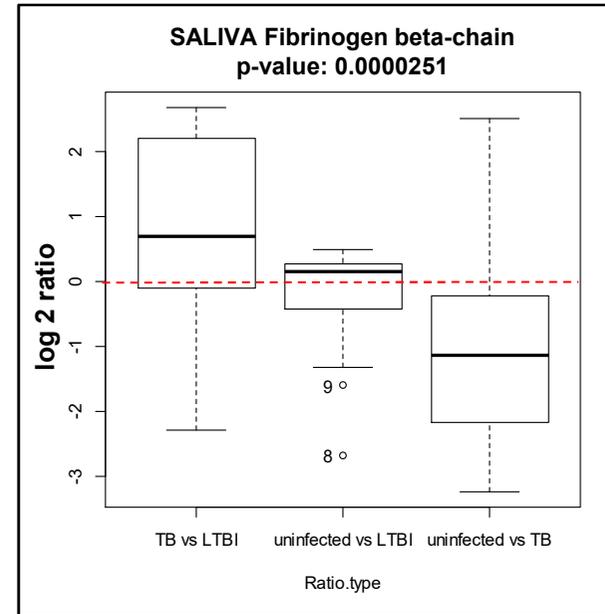
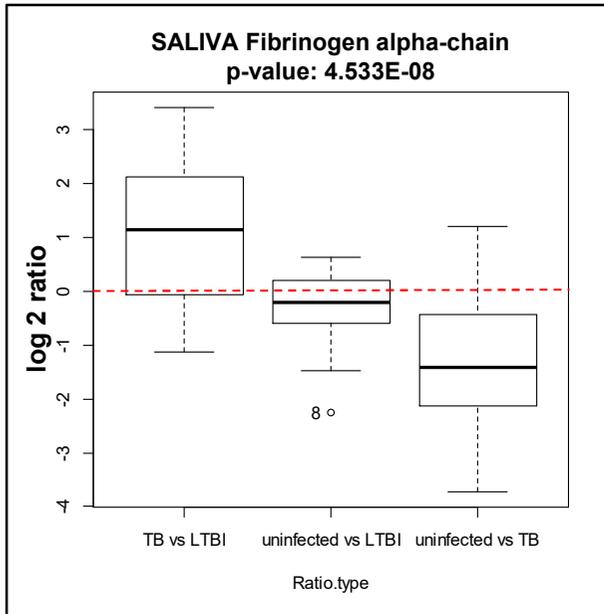
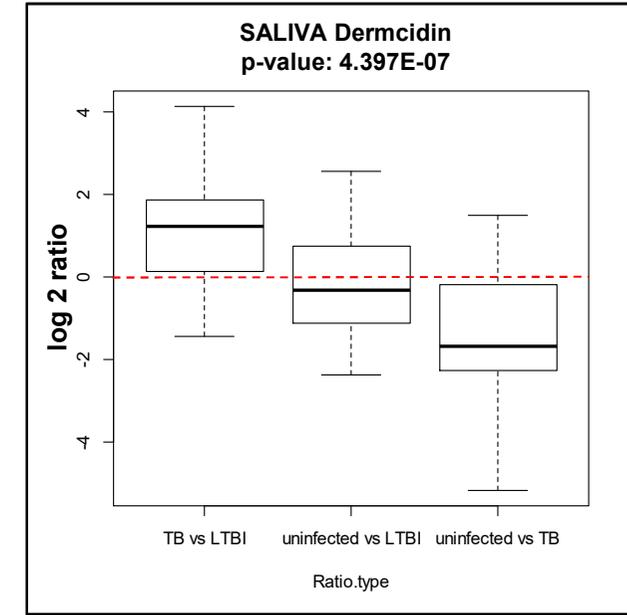
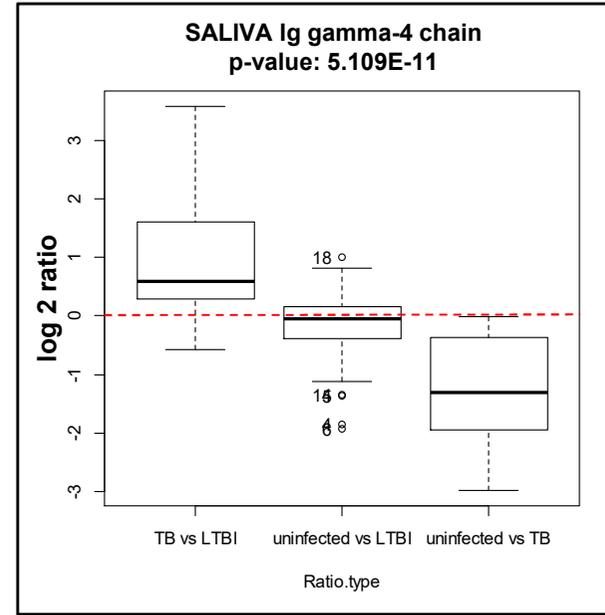
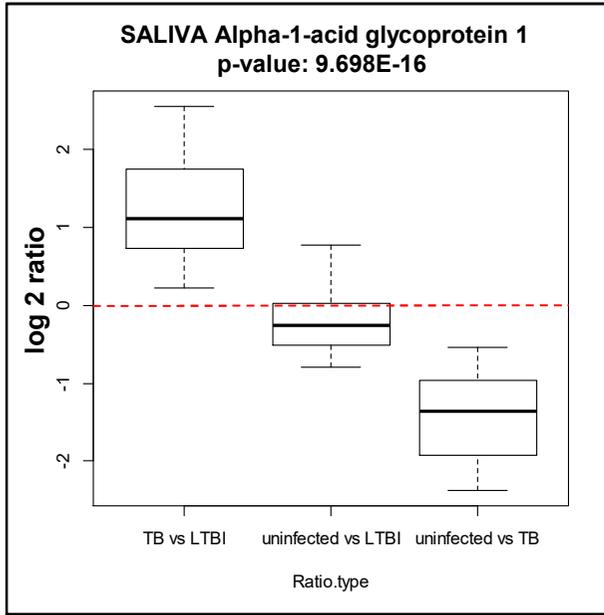


Figure 9

[Click here to download Figure: new Figure 9.pdf](#)

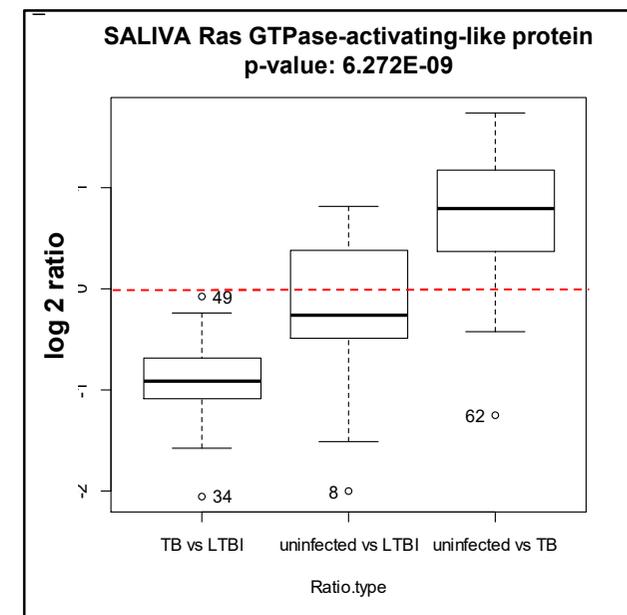
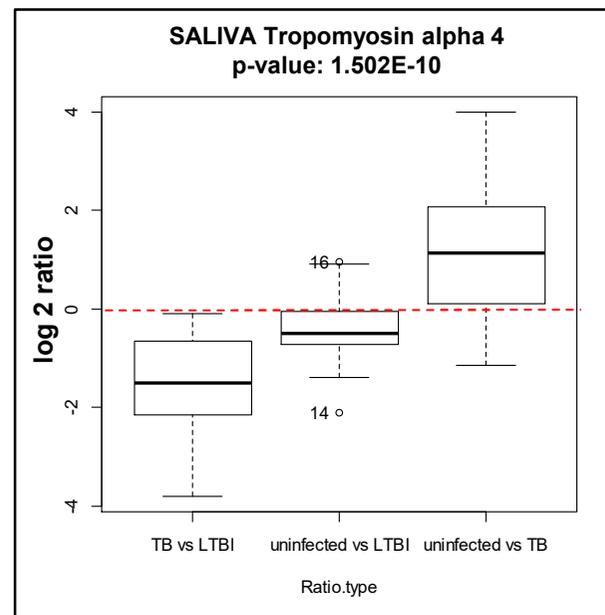
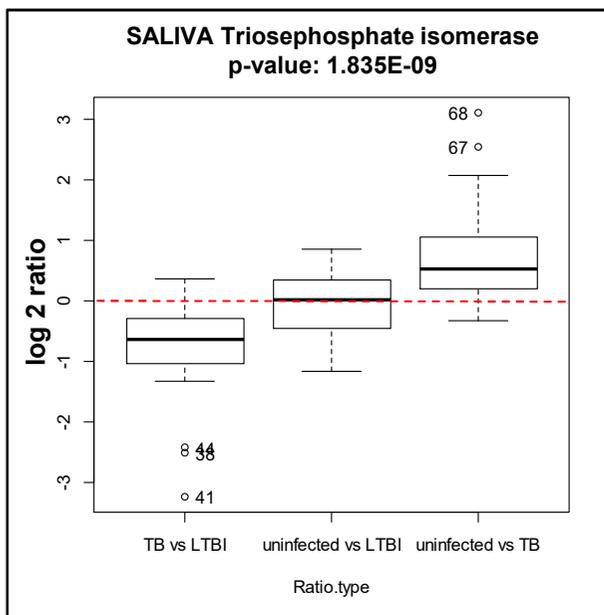
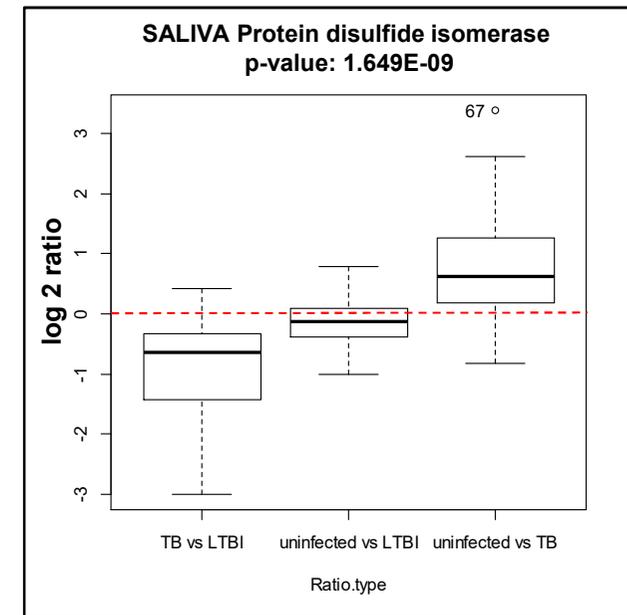
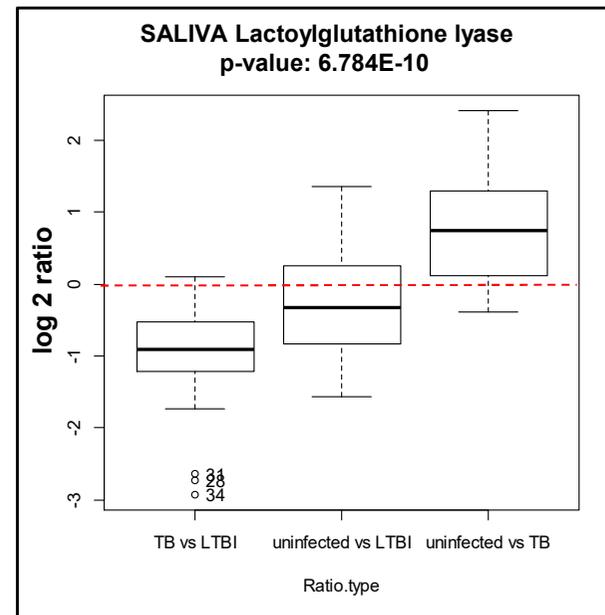
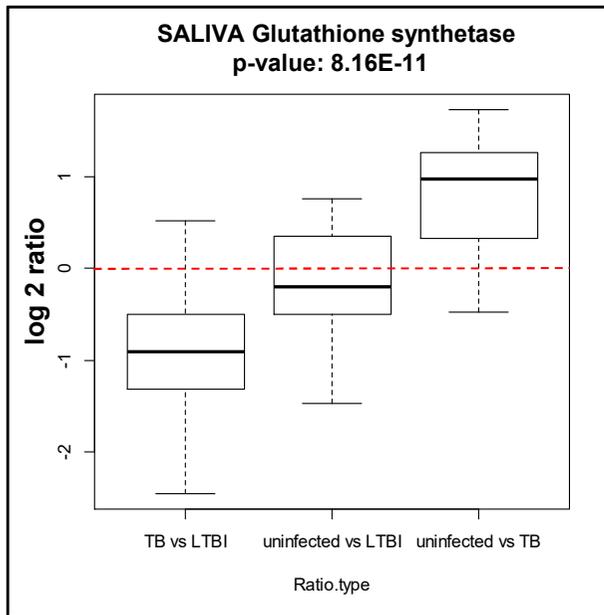
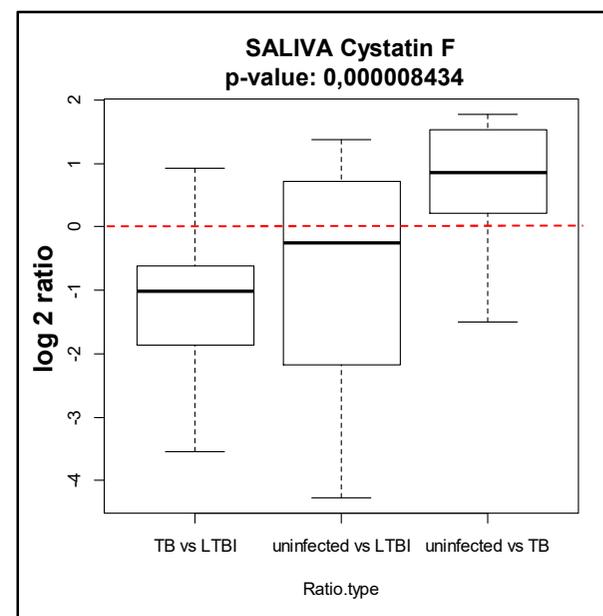
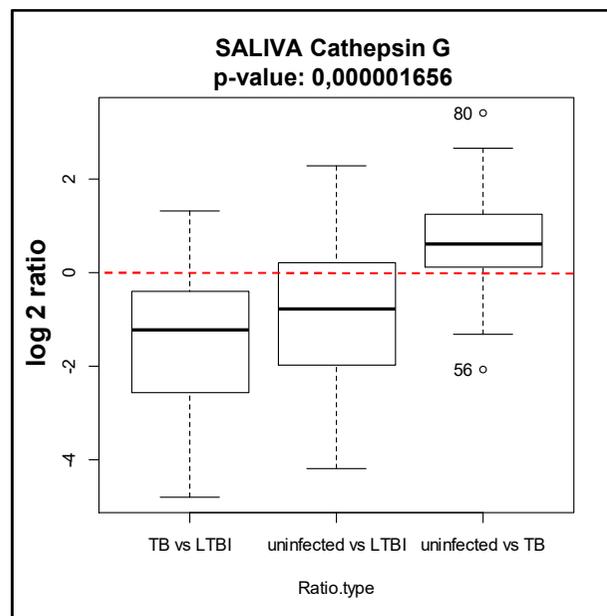
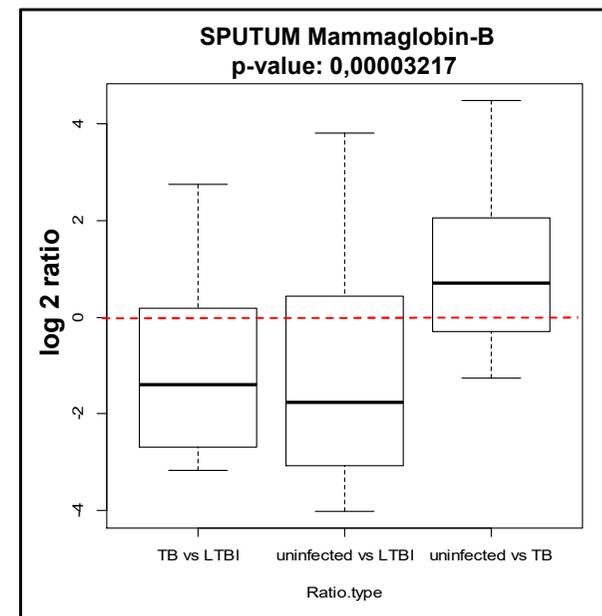
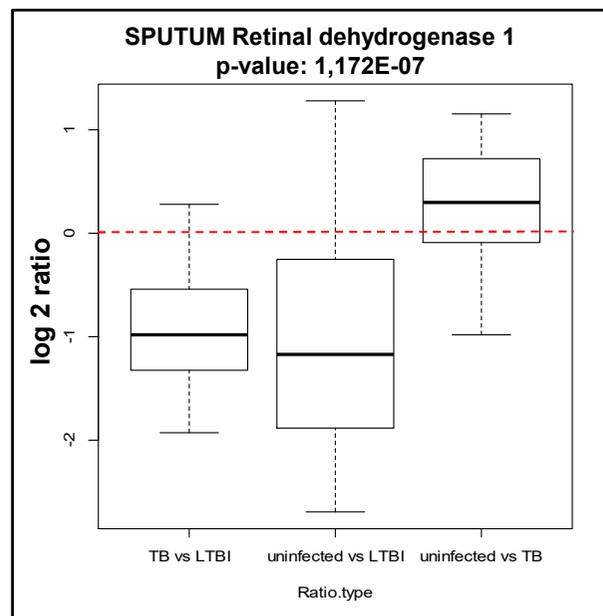
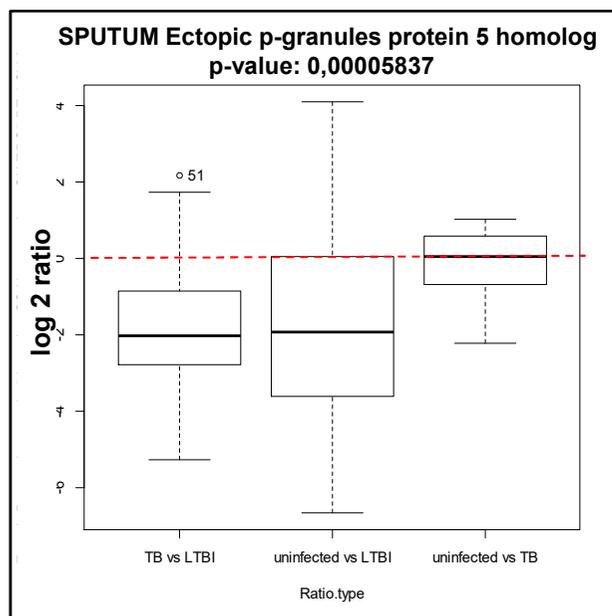


Figure 10

[Click here to download Figure: new Figure 10.pdf](#)



Supplementary figure 1

[Click here to download Supplementary material: new Supplementary figure 1.pdf](#)

Supplementary info 1

[Click here to download Supplementary material: new Supplementary Info 1 Statistical Anlysis.xlsx](#)

Supplementary info 2

[Click here to download Supplementary material: new Supplementary Info 2 Pathway analysis SPUTUM.xlsx](#)

Supplementary info 3

[Click here to download Supplementary material: new Supplementary Info 3 Pathway analysis SALIVA.xlsx](#)

***Conflict of Interest**

[Click here to download Conflict of Interest: Conflict of interest.docx](#)