nature research

https://doi.org/10.1038/s41587-020-0644-7

Supplementary information

Detection of SARS-CoV-2 in nasal swabs using MALDI-MS

In the format provided by the authors and unedited

Supporting Information

Coronavirus detection in nasal mucous secretion by MALDI-MS

Fabiane M. Nachtigall,[1] Alfredo Pereira,[2] Oleksandra S. Trofymchuk,[3] Leonardo S. Santos[2]*
Correspondence to: Dr. Leonardo S. Santos, Instituto de Química de Recursos Naturales, Universidad de Talca, Casilla 747, 3460000 Talca, Chile. Phone +56 71 201575. E-mail address: lssantos@utalca.cl
[1]Instituto de Ciencias Químicas Aplicadas, Universidad Autónoma de Chile, Talca 3467987, Chile.
[2] Instituto de Química de Recursos Naturales, Universidad de Talca, Talca 3460000, Chile.
[3] Núcleo Científico Multidisciplinario-DI, University of Talca, Talca 3460000, Chile.

Table of Contents

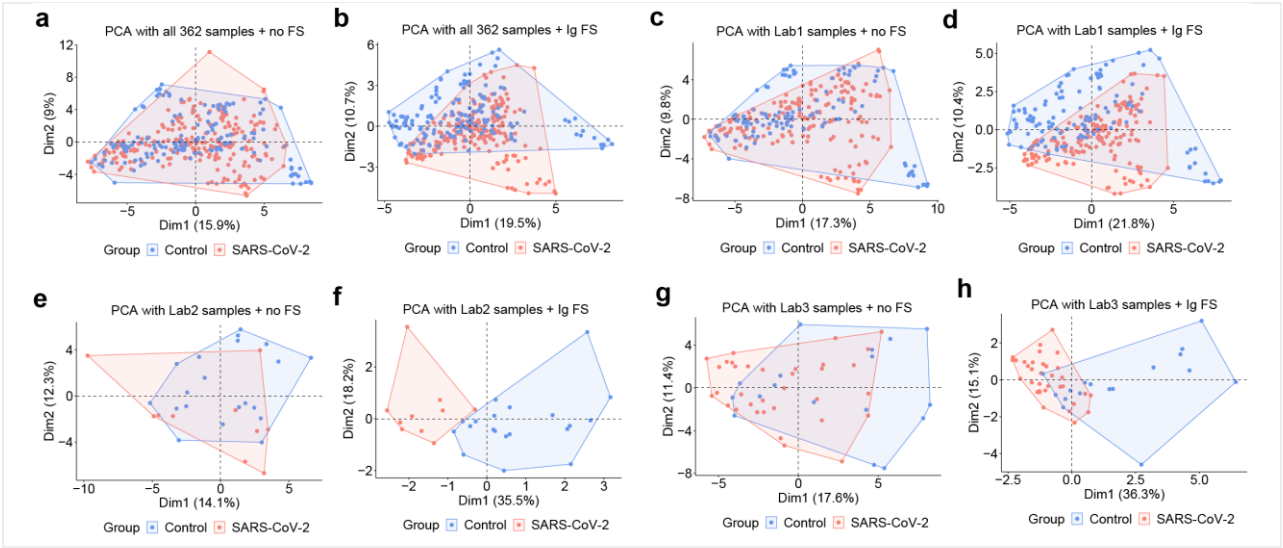

Fig. S1 Comparative analysis of SARS-CoV-2 (orange) and control (blue) samples using PCA. a, PCA of the mass spectra of the SARS-CoV-2 and control samples (362) from three labs using all 88 peaks. **b,** PCA of the mass spectra of the SARS-CoV-2 and control samples (362) from three labs using peaks selected with Ig FS method. **c**, PCA of the mass spectra of the SARS-CoV-2 and control samples from Lab 1 using all 88 peaks. **d,** PCA of the mass spectra of the SARS-CoV-2 and control samples from Lab 1 using peaks selected with Ig FS method. **e**, PCA of the mass spectra of the SARS-CoV-2 and control samples from Lab 2 using all 88 peaks. **f,** PCA of the mass spectra of the SARS-CoV-2 and control samples from Lab 2 using peaks selected with Ig FS method. **g**, PCA of the mass spectra of the SARS-CoV-2 and control samples from Lab 3 using all 88 peaks. **h,** PCA of the mass spectra of the SARS-CoV-2 and control samples from Lab 3 using peaks selected with Ig FS method.

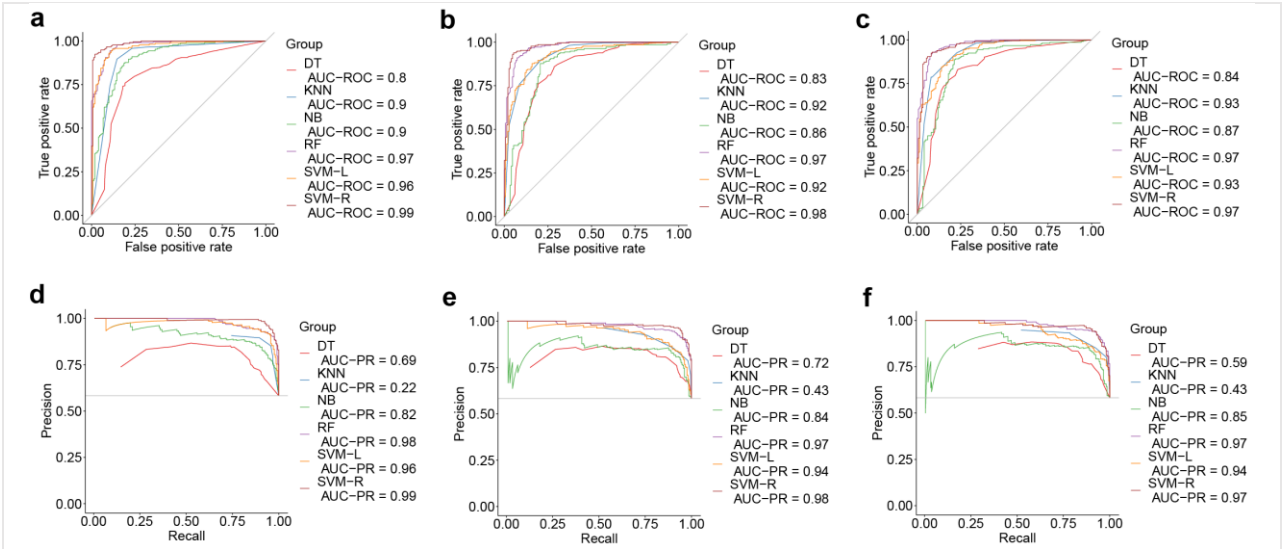

Fig. S2. **ROC and precision-recall curves obtained from 4-fold nested cross-validation with 362 samples (inter-laboratory detection). a**, ROC curve for six ML algorithms with all peaks (None FS). **b**, ROC curve for six ML algorithms with peaks selected with Ig FS. **c**, ROC curve for six ML algorithms with peaks selected with Cfs. **d**, Precision-recall curve for six ML algorithms with all peaks (None FS). **e**, Precision-recall curve for six ML algorithms with peaks selected with Ig FS. **f**, Precision-recall curves for six ML algorithms with peaks selected with Cfs.

Table S1. Two-tailed Wilcoxon rank sum test results for peaks intensities between SARS-CoV-2 and control samples. 39 peaks with a p value < 0.05 were found using two-tailed method and without multiple hypothesis test corrections. 31 peaks with a p value < 0.05 were found using FDR adjusted *p* value.[1]

#	Peak *(m/z)*	*p* value	FDR adjusted *p* value
1	3358	1.6E-15	1.4E-13
2	3095	1.3E-10	5.6E-09
3	4532	2.9E-10	8.5E-09
4	3337	6.7E-08	1.5E-06
5	3152	2.2E-07	3.2E-06
6	10444	1.9E-07	3.2E-06
7	3044	1.1E-06	1.3E-05
8	3487	4.5E-06	5.0E-05
9	11011	8.5E-06	8.3E-05
10	4138	1.0E-05	8.4E-05
11	7612	1.1E-05	8.4E-05
12	8215	7.9E-05	0.00058
13	3981	0.00013	0.00084
14	4966	0.00013	0.00084
15	3112	0.00024	0.00143
16	9956	0.00031	0.00171
17	4160	0.00042	0.00216
18	6192	0.00061	0.00300
19	3372	0.00179	0.00828
20	6964	0.00200	0.00882
21	10116	0.00247	0.01036
22	5950	0.00374	0.01497
23	7654	0.00476	0.01823
24	5236	0.00521	0.01909
25	8452	0.00713	0.02510
26	6361	0.00813	0.02753
27	3464	0.00949	0.03093
28	3443	0.01010	0.03173
29	4901	0.01234	0.03640
30	5382	0.01241	0.03640
31	3915	0.01456	0.04133
32	3754	0.01964	0.05401
33	4636	0.02068	0.05515
34	3651	0.02212	0.05725

35	3256	0.02280	0.05733
36	14692	0.02435	0.05951
37	3476	0.03161	0.07382
38	4574	0.03187	0.07382
39	5218	0.03565	0.08044

Table S2. Peaks (*m/z*) selected using Ig FS and Cfs for each set of samples (Labs). Only the peak of *m/z* of 7612 was common among laboratories. Ig FS selected peaks are ranked by information gain.

#	**All 362 samples**		**Lab1 (Argentina)**		**Lab 2 (Talca)**		**Lab 3 (Peru)**	
	Cfs	Ig FS	Cfs	Ig FS	Cfs	Ig FS	Cfs	Ig FS
1	3044	4532	3095	4532	3095	5147	3044	4428
2	3095	3358	3152	3358	3804	**7612**	3193	3044
3	3152	3152	3358	3152	4428	4473	4428	3193
4	3256	10444	4138	4138	4473	4428	5530	8215
5	3337	4160	4160	11011	5147	3095	6639	6639
6	3358	**7612**	4192	4160	**7612**	3804	**7612**	5256
7	3981	3337	4532	6192		8589	8215	5530
8	4138	11011	6192	10444				6192
9	4160	6192	6639	4551				3516
10	4192	4966	**7612**	3095				**7612**
11	4374	3095	8568	10837				7765
12	4532	4138	10444	8568				
13	6192	3044	11011	7612				
14	6964	14692	14692	4966				
15	**7612**	3487		3112				
16	8568	4192		3487				
17	10444	3981		5594				
18	11011	3651		3337				
19	14692	4374		6639				
20		8215		4811				
21		6361		10116				
22		3372		3981				
23		8568		3651				
24		3112		14692				
25		3256		8215				
26		5594		6361				
27		4393		4374				
28		4738		3372				
29		6964		4192				
30				4636				
31				5950				
32				3256				
33				4738				

Table S3. Peaks (*m/z*) selected from 4-fold nested cross-validation with 80 samples using Ig FS and Cfs. Ig FS selected peaks are ranked by information gain.

#	Fold 1		Fold 2		Fold 3		Fold 4	
	Peaks (*m/z*)		Peaks (*m/z*)		Peaks (*m/z*)		Peaks (*m/z*)	
	Cfs	Ig FS	Cfs	Ig FS	Cfs	Ig FS	Cfs	Ig FS
1	3044	11011	3044	11011	3044	11011	3044	3044
2	3443	3804	3095	3044	3095	8589	3337	11011
3	3464	3044	3337	3464	3193	6639	3464	3464
4	3804	3464	3464	3443	3464	3464	3804	6639
5	4428	4532	3476	3372	3804	3804	4428	3443
6	4901	3443	3804	3804	4428	3372	6639	3804
7	11011	4901	4192	4192	5047	3443	9956	3372
8		3372	4428	3358	5530	3044	11011	9956
9		4986	6639	4428	6639	3392		4428
10		4428	8589	6639	8589	4428		4532
11		3487	11011	3487	10444	4192		3476
12		3318	11735	4532	11011	3487		4192
13		3476		4901		4532		5594
14		4160		3095		5423		3318
15		6361		8589		5530		3337
16				3476		3193		
17				3337		3476		
18				4551		6361		
19				5423		10444		
20				11735		4551		
21						3095		
22						4901		
23						3710		
24						4986		
25						5047		
26						10837		

Table S4. Hyperparameters of ML models in the 4-Fold nested cross-validation with 80 samples. Values were found by random search using 'tuneLength = 20' in the 'train' function of 'Caret' on R.

Fold	ML Algorithms	Hyperparameters	Optimal values		
			No FS	Cfs	Ig FS
1	DT	C	0.259547975	0.058743681	0.058743681
		M	6	2	2
	KNN	k	7	9	5
	NB	laplace	0	0	0
		usekernel	FALSE	FALSE	FALSE
		adjust	1	1	1
	RF	mtry	65	1	2
	SVM-L	C	0.106012804	0.106012804	0.106012804
	SVM-R	sigma	0.002850441	0.031470579	0.004733863
		C	10.96400224	0.354916359	0.997286379
2	DT	C	0.228870888	0.228870888	0.228870888
		M	12	12	12
	KNN	k	7	13	13
	NB	laplace	0	0	0
		usekernel	FALSE	FALSE	FALSE
		adjust	1	1	1
	RF	mtry	5	1	1
	SVM-L	C	0.106012804	0.612207669	0.106012804
	SVM-R	sigma	0.004421021	0.015539616	0.005450704
		C	6.615957411	0.298823492	19.52686716
3	DT	C	0.228870888	0.228870888	0.228870888
		M	12	12	12
	KNN	k	19	13	19
	NB	laplace	0	0	0
		usekernel	FALSE	TRUE	TRUE
		adjust	1	1	1
	RF	mtry	18	2	5
	SVM-L	C	0.106012804	4.361898573	0.126744843
	SVM-R	sigma	0.009132039	0.051138823	0.012359254
		C	21.01984561	6.615957411	0.746531759
4	DT	C	0.069355084	0.259547975	0.280166373
		M	4	6	8
	KNN	k	19	15	13
	NB	laplace	0	0	0

	usekernel	FALSE	FALSE	FALSE
	adjust	1	1	1
RF	mtry	26	1	4
SVM-L	C	0.106012804	0.126744843	0.106012804
SVM-R	sigma	0.003115024	0.063352634	0.01443073
	C	10.96400224	0.296994818	0.746531759

Table S5. Peaks (*m/z*) selected from 4-fold nested cross-validation with 362 samples (inter-laboratory detection) using Ig FS and Cfs. Ig FS selected peaks are ranked by information gain.

#	Fold 1		Fold 2		Fold 3		Fold 4	
	Peaks (*m/z*)		Peaks (*m/z*)		Peaks (*m/z*)		Peaks (*m/z*)	
	Cfs	Ig FS	Cfs	Ig FS	Cfs	Ig FS	Cfs	Ig FS
1	3044	4532	3044	4532	3044	4532	3044	4532
2	3095	3152	3152	3358	3095	3358	3095	3152
3	3152	3358	3337	3152	3152	3152	3152	3358
4	3256	10444	3358	10444	3337	4160	3337	4160
5	3337	4160	3487	11011	3358	10444	3358	10444
6	3358	3337	3981	4160	3487	11011	4138	3337
7	3981	11011	4160	3487	4138	6192	4160	6192
8	4160	4966	4192	6192	4160	10116	4192	4138
9	4532	14692	4374	3044	4192	3095	4374	3095
10	4966	3095	4532	3981	4532	3337	4532	7612
11	5530	6192	6192	3337	5594	4138	6192	3044
12	6192	4138	7612	7612	6192	3487	7612	11011
13	7612	3044	7654	3372	6361	14692	10444	3112
14	8215	3256	8568	4192	7612	3044		3651
15	10444	3981	10444	4138	10116	6361		4192
16	11011	7612	14692	3095	10444	3112		14692
17	14692	8215		6361	11011	7612		4374
18		10837		14692	14692	4636		
19		4738		7654		4192		
20		5530		4374		3651		
21		3651		8568		5594		
22		4374						

Table S6. Hyperparameters of ML models in the 4-fold nested cross-validation with 362 samples (inter-laboratory detection). Values were found by random search using 'tuneLength = 20' in the 'train' function of 'Caret' on R.

Fold	ML Algorithms	Hyperparameters	Optimal values		
			No FS	Cfs	Ig FS
1	SVM-R	sigma	0.020821492	0.009976315	0.162733753
		C	40.93533689	117.718977	40.93533689
2	SVM-R	sigma	0.018462965	0.01684511	0.145994387
		C	53.9660878	26.02133193	40.93533689
3	SVM-R	sigma	0.006944261	0.379327866	0.249845419
		C	182.0196081	4.522226666	40.93533689
4	SVM-R	sigma	0.019022601	0.018961215	0.00690183
		C	40.93533689	26.02133193	230.3463207

Table S7. Peaks used to obtain the calibration function for Lab 2 spectra. The function is: $-7E\text{-}11x^3 + 2E\text{-}06x^2 + 0.973x + 51.611$ and the $R^2 = 1$.

Lab 2 Peaks (X) (*m/z*)	Lab 1 Peaks (Y) (*m/z*)
3391	3372
3463	3442
3508	3487
3736	3710
4391	4356
4409	4373
4430	4395
4589	4551
5009	4966
5281	5235
7016	6952
8529	8453
10539	10444
14828	14691

Table S8. RT-PCR Ct values for the samples used in the ML training and analysis.

Sample ID	Group	Lab	Ct
P1	Positive	2	19
P2	Positive	2	18
P3	Positive	2	28
P4	Positive	2	34
P5	Positive	2	32
P6	Positive	2	22
P7	Positive	2	25
P8	Positive	2	33
P9	Positive	2	18
P10	Positive	2	26
P11	Positive	3	28
P12	Positive	3	25
P13	Positive	3	29
P14	Positive	3	37
P15	Positive	3	36
P16	Positive	3	33
P17	Positive	3	32
P18	Positive	3	34
P19	Positive	3	26
P20	Positive	3	27
P21	Positive	3	26
P22	Positive	3	32
P23	Positive	3	33
P24	Positive	3	31
P25	Positive	3	32
P26	Positive	3	26
P27	Positive	3	25
P28	Positive	3	24
P29	Positive	3	24
P30	Positive	3	25
P31	Positive	3	20
P32	Positive	3	21
P33	Positive	3	23
P34	Positive	3	18
P35	Positive	3	19
P36	Positive	3	18
P37	Positive	3	18

P38	Positive	3	29
P39	Positive	1	28
P40	Positive	1	24
P41	Positive	1	24
P42	Positive	1	25
P43	Positive	1	35
P44	Positive	1	36
P45	Positive	1	35
P46	Positive	1	34
P47	Positive	1	34
P48	Positive	1	33
P49	Positive	1	17
P50	Positive	1	18
P51	Positive	1	17
P52	Positive	1	18
P53	Positive	1	21
P54	Positive	1	21
P55	Positive	1	19
P56	Positive	1	21
P57	Positive	1	20
P58	Positive	1	23
P59	Positive	1	23
P60	Positive	1	25
P61	Positive	1	27
P62	Positive	1	23
P63	Positive	1	24
P64	Positive	1	25
P65	Positive	1	24
P66	Positive	1	35
P67	Positive	1	33
P68	Positive	1	34
P69	Positive	1	32
P70	Positive	1	33
P71	Positive	1	32
P72	Positive	1	25
P73	Positive	1	25
P74	Positive	1	24
P75	Positive	1	23
P76	Positive	1	23

P77	Positive	1	22
P78	Positive	1	18
P79	Positive	1	19
P80	Positive	1	18
P81	Positive	1	28
P82	Positive	1	28
P83	Positive	1	27
P84	Positive	1	25
P85	Positive	1	24
P86	Positive	1	34
P87	Positive	1	33
P88	Positive	1	34
P89	Positive	1	20
P90	Positive	1	20
P91	Positive	1	23
P92	Positive	1	24
P93	Positive	1	24
P94	Positive	1	25
P95	Positive	1	24
P96	Positive	1	24
P97	Positive	1	28
P98	Positive	1	27
P99	Positive	1	27
P100	Positive	1	24
P101	Positive	1	24
P102	Positive	1	25
P103	Positive	1	34
P104	Positive	1	33
P105	Positive	1	35
P106	Positive	1	19
P107	Positive	1	23
P108	Positive	1	31
P109	Positive	1	30
P110	Positive	1	32
P111	Positive	1	21
P112	Positive	1	25
P113	Positive	1	27
P114	Positive	1	31
P115	Positive	1	26

P116	Positive	1	29
P117	Positive	1	24
P118	Positive	1	33
P119	Positive	1	34
P120	Positive	1	36
P121	Positive	1	26
P122	Positive	1	23
P123	Positive	1	25
P124	Positive	1	27
P125	Positive	1	29
P126	Positive	1	32
P127	Positive	1	32
P128	Positive	1	34
P129	Positive	1	35
P130	Positive	1	20
P131	Positive	1	22
P132	Positive	1	27
P133	Positive	1	28
P134	Positive	1	31
P135	Positive	1	26
P136	Positive	1	27
P137	Positive	1	35
P138	Positive	1	24
P139	Positive	1	24
P140	Positive	1	29
P141	Positive	1	25
P142	Positive	1	26
P143	Positive	1	35
P144	Positive	1	26
P145	Positive	1	26
P146	Positive	1	16
P147	Positive	1	17
P148	Positive	1	17
P149	Positive	1	18
P150	Positive	1	18
P151	Positive	1	17
P152	Positive	1	29
P153	Positive	1	30
P154	Positive	1	21

P155	Positive	1	21
P156	Positive	1	33
P157	Positive	1	18
P158	Positive	1	17
P159	Positive	1	17
P160	Positive	1	25
P161	Positive	1	25
P162	Positive	1	26
P163	Positive	1	29
P164	Positive	1	27
P165	Positive	1	24
P166	Positive	1	34
P167	Positive	1	29
P168	Positive	1	26
P169	Positive	1	25
P170	Positive	1	22
P171	Positive	1	23
P172	Positive	1	36
P173	Positive	1	35
P174	Positive	1	25
P175	Positive	1	23
P176	Positive	1	22
P177	Positive	1	21
P178	Positive	3	24
P179	Positive	3	21
P180	Positive	3	33
P181	Positive	3	24
P182	Positive	3	25
P183	Positive	3	26
P184	Positive	3	18
P185	Positive	3	19
P186	Positive	3	26
P187	Positive	3	32
P188	Positive	3	19
P189	Positive	3	36
P190	Positive	3	30
P191	Positive	3	37
P192	Positive	3	33
P193	Positive	3	36

P194	Positive	3	34
P195	Positive	3	24
P196	Positive	3	30
P197	Positive	3	32
P198	Positive	3	18
P199	Positive	3	27
P200	Positive	3	32
P201	Positive	3	31
P202	Positive	3	28
P203	Positive	3	25
P204	Positive	3	24
P205	Positive	3	19
P206	Positive	3	22
P207	Positive	3	18
P208	Positive	3	21
P209	Positive	3	25
P210	Positive	3	24
P211	Positive	3	34
N1	Negative	2	>40
N2	Negative	2	>40
N3	Negative	2	>40
N4	Negative	2	>40
N5	Negative	2	>40
N6	Negative	2	>40
N7	Negative	2	>40
N8	Negative	2	>40
N9	Negative	2	>40
N10	Negative	2	>40
N11	Negative	2	>40
N12	Negative	2	>40
N13	Negative	2	>40
N14	Negative	2	>40
N15	Negative	2	>40
N16	Negative	2	>40
N17	Negative	2	>40
N18	Negative	2	>40
N19	Negative	2	>40
N20	Negative	2	>40
N21	Negative	1	>40

N22	Negative	1	>40
N23	Negative	1	>40
N24	Negative	1	>40
N25	Negative	1	>40
N26	Negative	1	>40
N27	Negative	1	>40
N28	Negative	1	>40
N29	Negative	1	>40
N30	Negative	1	>40
N31	Negative	1	>40
N32	Negative	1	>40
N33	Negative	1	>40
N34	Negative	1	>40
N35	Negative	1	>40
N36	Negative	1	>40
N37	Negative	1	>40
N38	Negative	1	>40
N39	Negative	1	>40
N40	Negative	1	>40
N41	Negative	1	>40
N42	Negative	1	>40
N43	Negative	1	>40
N44	Negative	1	>40
N45	Negative	1	>40
N46	Negative	1	>40
N47	Negative	1	>40
N48	Negative	1	>40
N49	Negative	1	>40
N50	Negative	1	>40
N51	Negative	1	>40
N52	Negative	1	>40
N53	Negative	1	>40
N54	Negative	1	>40
N55	Negative	1	>40
N56	Negative	1	>40
N57	Negative	1	>40
N58	Negative	1	>40
N59	Negative	1	>40
N60	Negative	1	>40

N61	Negative	1	>40
N62	Negative	1	>40
N63	Negative	1	>40
N64	Negative	1	>40
N65	Negative	1	>40
N66	Negative	1	>40
N67	Negative	1	>40
N68	Negative	1	>40
N69	Negative	1	>40
N70	Negative	1	>40
N71	Negative	1	>40
N72	Negative	1	>40
N73	Negative	1	>40
N74	Negative	1	>40
N75	Negative	1	>40
N76	Negative	1	>40
N77	Negative	1	>40
N78	Negative	1	>40
N79	Negative	1	>40
N80	Negative	1	>40
N81	Negative	1	>40
N82	Negative	1	>40
N83	Negative	1	>40
N84	Negative	1	>40
N85	Negative	1	>40
N86	Negative	1	>40
N87	Negative	1	>40
N88	Negative	1	>40
N89	Negative	1	>40
N90	Negative	1	>40
N91	Negative	1	>40
N92	Negative	1	>40
N93	Negative	1	>40
N94	Negative	1	>40
N95	Negative	1	>40
N96	Negative	1	>40
N97	Negative	1	>40
N98	Negative	1	>40
N99	Negative	1	>40

N100	Negative	1	>40
N101	Negative	1	>40
N102	Negative	1	>40
N103	Negative	1	>40
N104	Negative	1	>40
N105	Negative	1	>40
N106	Negative	1	>40
N107	Negative	1	>40
N108	Negative	1	>40
N109	Negative	1	>40
N110	Negative	1	>40
N111	Negative	1	>40
N112	Negative	1	>40
N113	Negative	1	>40
N114	Negative	1	>40
N115	Negative	1	>40
N116	Negative	1	>40
N117	Negative	1	>40
N118	Negative	1	>40
N119	Negative	1	>40
N120	Negative	1	>40
N121	Negative	1	>40
N122	Negative	1	>40
N123	Negative	1	>40
N124	Negative	1	>40
N125	Negative	1	>40
N126	Negative	1	>40
N127	Negative	1	>40
N128	Negative	1	>40
N129	Negative	1	>40
N130	Negative	1	>40
N131	Negative	1	>40
N132	Negative	1	>40
N133	Negative	1	>40
N134	Negative	1	>40
N135	Negative	3	>40
N136	Negative	3	>40
N137	Negative	3	>40
N138	Negative	3	>40

N139	Negative	3	>40
N140	Negative	3	>40
N141	Negative	3	>40
N142	Negative	3	>40
N143	Negative	3	>40
N144	Negative	3	>40
N145	Negative	3	>40
N146	Negative	3	>40
N147	Negative	3	>40
N148	Negative	3	>40
N149	Negative	3	>40
N150	Negative	3	>40
N151	Negative	3	>40

Reference

1. Whitin, J. C., Jang, T., Merchant, M., Yu, T. T-S., Lau, K., Recht, B., Cohen , H. J. & Recht, L. Alterations in Cerebrospinal Fluid Proteins in a Presymptomatic Primary Glioma Model. *PlosOne* **7**, e49724 (2012).