Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 4 September 2020 doi:10.20944/preprints202007.0141.v3

Article

Possible Cross-Reactivity Between SARS-CoV-2 Proteins, CRM197 and Proteins in Pneumococcal Vaccines May Protect Against Symptomatic SARS-CoV-2 Disease and Death

Robert Root-Bernstein [1*]

[1] Michigan State University, Department of Physiology, East Lansing, MI 48824 USA; rootbern@msu.edu
* Correspondence: rootbern@msu.edu

Received: date; Accepted: date; Published: date

Abstract: Various studies indicate that vaccination, especially with pneumococcal vaccines, protects against symptomatic cases of SARS-CoV-2 infection and death. This paper explores the possibility that pneumococcal vaccines in particular, but perhaps other vaccines as well, contain antigens that might be cross-reactive with SARS-CoV-2 antigens. Comparison of the glycosylation structures of SARS-CoV-2 with the polysaccharide structures of pneumococcal vaccines yielded no obvious similarities. However, while pneumococcal vaccines are primarily composed of capsular polysaccharides, some are conjugated to CRM197, a modified diphtheria toxin, and all contain about three percent protein contaminants, including the pneumococcal surface proteins PsaA, PspA and probably PspC. All of these proteins have very high degrees of similarity, using very stringent criteria, with several SARS-CoV-2 proteins including the spike protein, membrane protein and replicase 1a. CRM197 is also present in Hib and meningitis vaccines. Equivalent similarities were found at lower rates, or were completely absent, among the proteins in diphtheria, tetanus, pertussis, measles, mumps, rubella, and poliovirus vaccines. Notably, PspA and PspC are highly antigenic and new pneumococcal vaccines based on them are currently in human clinical trials so that their effectiveness against SARS-CoV-2 disease is easily testable.

Keywords: COVID-19; SARS-CoV-2; pneumococcal; Streptococcus pneumoniae; vaccine; vaccination; cross-reactivity; similarity; protection; CRM197; PspA; PsaA; PspC; BCG; poliovirus; measles-mumps-rubella; diphtheria-tetanus-pertussis; meningococcus

---

1. Introduction

Various studies have indicated that some vaccines may protect against symptomatic SARS-CoV-2 infection and death. A very significant inverse correlation has been found between rates of pneumococcal vaccination at both national and local population levels and rates of SARS-CoV-2 infections and death [1]. No such correlations were found in that study to the tuberculosis vaccine BCG (Bacillus Calmette Guerin), Haemophilus influenzae type B (Hib), diphtheria-tetanus-pertussis, measles-mumps-rubella, or poliovirus vaccinations. The results were controlled for percent of the population over 65 years of age, percent of obese individuals, percent of diabetics and the sum of these factors. Pneumococcal vaccination with PCV13 was again found to be very significantly protective in a study of 137,037 individuals for whom vaccination records were available [2] and other recent vaccinations also provided apparent protection against SARS-CoV-2 after controlling for other variables. The purpose of this paper is to provide a possible mechanism for how pneumococcal and other vaccines might protect against SARS-CoV-2.

The specific hypothesis tested here is that antigens in pneumococcal vaccines induce antibodies protective against SARS-CoV-2 by means of cross-reactivity with similar SARS-CoV-2 antigens. I

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 4 September 2020 doi:10.20944/preprints202007.0141.v3

have treated all other vaccines as controls. There are two types of antigens that might play such a role, one being the capsular polysaccharide antigens in current pneumococcal vaccines and the other the proteins that they contain. An extensive search for polysaccharide structures comparing SARS-CoV-2 glycosylated proteins [3] and S. pneumoniae serotypes [4] failed to identify any obvious similarities. SARS-CoV-2 glycosylations are composed mainly of various arrangements of N-acetylglucosamine, mannose, galactose and N-acetylneuraminic acid, with fucose appearing in about half of the polysaccharides [3]. While N-acetylglucosamine and some mannose derivatives appear in pneumococcal polysaccharides, N-acetylneuraminic acid does not appear in any and only pneumococcal serogroups 4, 5, 12 and 46 contain polysaccharides composed of both mannose and fucose or N-acetylglucosamine and fucose [4]. These pneumococcal polysaccharides do not, however, appear to share any obvious structural similarities with SARS-CoV-2 polysaccharides. While identity of polysaccharide structures is probably not required for antigenic cross-reactivity, with no obvious structural homologies, the search then shifted to possible protein similarities.

While current pneumococcal vaccines are composed primarily of capsular polysaccharides, they also contain one or both of two types of proteins. The polysaccharide component is never pure, generally containing around three percent of the cell surface proteins to which the polysaccharides are attached [5-7]. Proteins identified in pneumococcal vaccines include pneumococcal surface protein A (PspA) and pneumococcal surface adhesin A (PsaA) [8,9]. Because the presence of PsaA was identified only by immunological methods and PsaA cross-reacts strongly with an additional pneumococcal surface protein, PspC (also known as CbpA and SpsA) [10, 11], it is likely that PspC is also present in capsular polysaccharide-based pneumococcal vaccines. Additionally, pneumococcal conjugate vaccines covalently attach the polysaccharides to a modified diphtheria toxin protein called Cross-Reactive Material 197 (CRM197) which is also present in Hib and meningitis vaccines [12].

This study reports that SARS-CoV-2 proteins contain many regions that mimic sequences within pneumococcal surface proteins as well as CRM197 (which is also found in *Haemophilus influenzae* type B [Hib] vaccine and meningitis vaccine) as well as rubella proteins but much less frequently to proteins present in other vaccines..

2. Materials and Methods

In order to ascertain whether PspA, PsaA, PspC and CRM197 have regions of significant similarity to SARS-CoV-2 proteins, LALIGN (at www.expasy.org) was employed to perform pair-wise protein comparisons. The parameters chosen were 20 best alignments to show; BLOSUM80 (in order to maximize small, local similarities); E = 10; gap penalty of -10.0 (to maximize continuous sequence similarities as are recognized by human leukocyte antigens and T cell receptors). SARS-CoV-2 sequences were retrieved from https://viralzone.expasy.org/8996 as HTML files or using the accession numbers from the UniProtKB database (UniProtKB accession numbers P0DTC1-P0DTC9). Streptococcus pneumoniae PspA, PsaA and PspC sequences were retrieved as accession numbers (provided in the Tables below) from the UniProtKB database. Because different streptococcal serotypes have slightly different versions of these proteins, several were randomly selected for each search and the sequences similarities displayed in FIGURE 1 are representative of several serotype results. The accession numbers for the pneumococcal vaccines, CRM197 and the control vaccine proteins are listed in TABLE 1.

Table 1. UniProtKB accession numbers for viral and bacterial proteins used in this study.

STREPTOCOCCUS PNEUMONIAE	O34097, Q9LAZ1, B2IRK1, Q9LAY4	pspA, Pneumococcal surface protein A
	P0A4G2 P0A4G3 P42363 Q04JB8	psaA, Pneumococcal surface protein, Manganese ABC transporter substrate protein
	Q9KK40 Q9FDQ1 Q9KK37	pspC, Pneumococcal Surface protein PspC

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 4 September 2020 doi:10.20944/preprints202007.0141.v3

	Q9KK24	
	Q8DRK2	Pneumococcall Gram-positive anchor protein
MUMPS	P11235|	HN_MUMPM (HN)RecName: Full=Hemagglutinin-neuraminidase
	P30929	L_MUMPM (L)RecName: Full=RNA-directed RNA polymerase L
	P09458	FUS_MUMPR (F)RecName: Full=Fusion glycoprotein F0
	P30928	V_MUMPM (P/V)RecName: Full=Non-structural protein V
	P22112	SH_MUMPM (SH)RecName: Full= Small hydrophobic protein
MEASLES	P08362	HEMA_MEASE (H)RecName: Full=Hemagglutinin glycoprotein
	Q89933	NCAP_MEASF (N)RecName: Full=Nucleoprotein
	P12576	L_MEASE (L)RecName: Full=RNA-directed RNA polymerase L
	Q786F3	FUS_MEASC (F)RecName: Full=Fusion glycoprotein F0
	P0C774	V_MEASC (P/V)RecName: Full=Non-structural protein V
RUBELLA	P08563	POLS_RUBVM RecName: Full=Structural polyprotein (contains spike protein E1, spike protein E2, capsid protein) 1063 aa
	Q86500	POLN_RUBVM RecName: Full=Non-structural polyprotein p200 (contains p90, p150 and p200 proteins) 2116 aa
POLIO	P03301	P03301|POLG_POL1S RecName: Full=Genome polyprotein; 2209 aa CONTAINS: P3; Protein 3AB; P1; Capsid protein VP0; Capsid protein VP4; Capsid protein VP2; Capsid protein VP3
PERTUSSIS	P04977	TOX1_BORPE (ptxA)RecName: Full=Pertussis toxin subunit 1
	P04978	TOX2_BORPE (ptxB)RecName: Full=Pertussis toxin subunit 2
	P04979	TOX3_BORPE (ptxC)RecName: Full=Pertussis toxin subunit 3
	P0A3R5	TOX4_BORPE (ptxD)RecName: Full=Pertussis toxin subunit 4
	P04981	TOX5_BORPE (ptxE)RecName: Full=Pertussis toxin subunit 5
	P35077	FHAC_BORPE (fhaC)RecName: Full=Filamentous hemagglutinin transporter protein FhaC
	P14283	PERT_BORPE (prn)RecName: Full=Pertactin autotransporter
	P05788	FM2_BORPE (fim2)RecName: Full=Serotype 2 fimbrial subunit
	P17835	FM3_BORPE (fim3)RecName: Full=Serotype 3 fimbrial subunit
TETANUS	P04958	TETX_CLOTE (tetX)RecName: Full=Tetanus toxin
DIPHTHERIA	Q5PY51	Q5PY51_CORDP SubName: Full=Diphtheria toxin
	Q6NK15	Q6NK15_CORDI (tox)SubName: Full=Diphtheria toxin
MENINGOCOCCUS	0DH58	OMPA_NEIMB (porA)RecName: Full=Major outer membrane protein
SARS-CoV-2	P0DTC1	P0DTC1 Replicase polyprotein 1a (pp1a)
	P0DTC2	P0DTC2 Spike glycoprotein (S)
	P0DTC3	P0DTC3 Protein 3a (NS3a)
	P0DTC4	P0DTC4 Envelope small membrane protein (E)
	P0DTC5	P0DTC5 Membrane protein (M)
	P0DTC6	P0DTC6 Non-structural protein 6 (NS6)
	P0DTC7	P0DTC7 Protein 7a (NS7a)
	P0DTC8	P0DTC8 Non-structural protein 8 (NS8)
	P0DTC9	P0DTC9 Nucleoprotein (N)
	P0DTD1	P0DTD1 Replicase polyprotein 1ab (pp1ab)
	P0DTD2	P0DTD2 Protein 9b (NS9B)
	P0DTD3	P0DTD3 Uncharacterized protein 14 (NS14)
	P0DTD8	P0DTD8 Protein 7b (NS7b)
Mycobacterium tuberculosis	MYCTU_ UP000001584	M. tuberculosis (strain ATCC 25618 / 3,997 protein sequences; 1,332,562 total letters
Bordetella pertussis	BORPE_ UP000002676	B. pertussis strain Tohama I / ATCC BAA-589 / NCTC 13251; 3260 proteins sequences
Escherichia coli K12	ECOLI_ UP000000625	Escherichia coli K12, 4,403 protein sequences
Clostridium leptum	9CLOT_ UP000018168	Clostridium leptum CAG:27 proteome; 2,482 protein sequences
Lactobacillus paracasei	LACP3_ UP000001651	Lactobacillus paracasei strain ATCC 334 / BCRC; 2708 protein sequences
Lactococcus lactis	LACLA_ UP000002196	Lactococcus lactis subsp. lactis (strain IL1403); 2,225 protein sequences

The LALIGN results were culled by applying the criterion that any sequence similarity reported must have an E value less than either 0.1 (TABLE 2) or 1.0 (TABLE 3), a Waterman-Eggert score of more than 50, and a region containing at least six out of ten identities. These criteria are based on a number of experimental studies involving the average length of peptide recognized by major histocompatibility (MHC) receptors and T cell receptors (TCR) which is about 10 consecutive amino acids [13-15] and the degree of similarity between two antigens that is likely to induce cross-reactive immune responses, which generally consists of at least five consecutive identical amino acids or six identities distributed within a 10 amino acid sequence [14, 16-20].

In essence, setting the E value to 0.1 or 1.0 determines how many matches the BLAST program will yield. The lower the E value, the less matches BLAST will yield because a lower E value limits the matches to those with rare combinations of amino acids such as methionines, tryptophans, tyrosines, cysteines, etc., rather than ones made up of sequences of very common amino acids such as glycine, alanine, valine and leucine, which appear at high rates in almost all proteins. In this case, keeping the E value low also selects of matching sequences that have a high probability of being antigenic since the immune system is more sensitive to rare amino acids than to common ones. Conversely, the lower the Waterman-Eggert score, the less amino acid matches are likely to be found in a pair of sequences. Thus, limiting the Waterman-Eggert score to more than 50 provides reasonable assurance that any sequence that appears in the BLAST search will display a high proportion of amino acid identities and similarities. Experience shows [16-20] that the combination of low E value and high Waterman-Eggert score tends to yield reasonably short sequences of high similarity, which is emphasized by using BLOSUM80. Despite using these boundary conditions, however, experiences shows about half of the sequences that BLAST yields are unlikely to be antigenically cross-reactive. As noted above, TCR and MHC recognize short peptide sequences averaging about 10 amino acids in length [13-15] and experimental evidence has shown that within such peptides, sequences of five contiguous identical amino acids or six non-contiguous amino acids are generally required for two peptides to elicit cross-reactive T cell or B cell responses [14, 16-20]. Thus, the BLAST results were culled for sequences meeting the latter criteria. Thus, by employing the tried-and-tested set of parameters just described, previous experimentation demonstrates that the resulting matches have a high probability of being recognized as cross-reactive antigens.

As controls for the LALIGN results, all thirteen SARS-Cov-2 proteins were used to search for similarities to bacterial proteins found in diphtheria, pertussis, and tetanus vaccines (TABLE 1) and viral proteins incorporated into the measles, mumps, rubella and polio vaccines. The only identified proteins in Hib and meningitis vaccines are CRM197 or meningococcal outer membrane complex protein, so these were also examined for similarities to SARS-CoV-2 proteins (TABLES 1 and 2). The same criteria used above were used to screen the results for sequences having at least six identities in a span of ten amino acids.

Bacillus Calmette Guerin (BCG) vaccine could not be searched as were the other vaccines. BCG is a version of Mycobacterium bovis consisting of 3891 proteins. It has no integrated, searchable proteome on BLAST (www.expasy.org); instead, each protein is separately listed in the UniProt database (https://www.uniprot.org/uniprot/?query=taxonomy:410289). *M. tuberculosis* ([MYCTU_UP000001584] Mycobacterium tuberculosis (strain ATCC 25618 / comprised of 3,997 sequences) was substituted for BCG since they are highly cross-reactive. Since searching nearly 4000 proteins using the LALIGN method listed above was unreasonable, the complete proteome was searched instead and BLAST was used with the parameters set similarly (BLOSUM80; E = 10; filter low complexity regions; no gaps permitted; show best 100 matches). As with the other microbial comparisons, the results were hand curated to eliminate any sequences failing to meet the six-in-ten antigenic-cross-reactivity criterion and an E value of less than 1.0 (rather than 0.1,because this value gave equivalent length and quality of matches to the LALIGN searches) and a Waterman-Eggert score of at least 50.

Bordetella pertussis vaccines come in two forms; one is acellular (which is the form tested above using LALIGN) but there are also whole-cell pertussis vaccines, so the same BLAST procedure used to examine *M. tuberculosis* was used to examine *Bordetella pertussis* UP000002676. Taxonomy, 257313 - (strain Tohama I / ATCC BAA-589 / NCTC 13251) comprised of 3260 protein sequences.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 4 September 2020 doi:10.20944/preprints202007.0141.v3

As controls for the whole-bacteria BLAST searches, two human commensal bacteria, *Escherichia coli* (Escherichia coli K12 UP000000625, 4,403 protein sequences) and *Clostridium leptum* ([UP000018168] Clostridium leptum CAG:27 proteome, 2,482 protein sequences), as well as the probiotics Lactococcus lactis ([LACLA_UP000002196] *Lactococcus lactis* subsp. lactis (strain IL1403) 2,225 protein sequences) and *Lactobacillus paracasei* ([LACP3_UP000001651] Lactobacillus paracasei strain ATCC 334 / BCRC, 2708 protein sequences), were tested for similarities to SARS-CoV-2 proteins.

3. Results

Results of the LALIGN similarity searches that satisfy the criteria of at least six identical amino acids in a sequence of ten amino acids and a Waterman-Eggert score of 50 or greater are found in TABLES 2 and 3 and in the FIGURES. Results with E values of 0.1 or less are summarized in TABLE 2 and FIGURES 1-4. Those that satisfy a W-E score of 50 or greater and an E value of 1.0 or less are summarized in TABLE 3 but sequences are not provided as they are too numerous.

TABLE 2 demonstrates that pneumococcal proteins psaA, pspA and psPc present a very large number of high-quality sequence matches with various SARS-CoV-2 proteins. All of these matches are provided in FIGURE 1. Twenty-one significant similarities were observed, ten of which are indicated in the figure in bold type as sequences that repeat within pair of proteins. Note that a significant sequence similarity was also found between SARS-CoV-2 proteins and the S. pneumoniae GRAM positive anchor protein (Q8DRK2), which serves as an anchor site for capsular polysaccharides. It is not known at this time whether this protein is among those contaminating capsular polysaccharide preparations but because of its association with polysaccharide anchoring, it is likely to be such a contaminant of the polysaccharide material used in pneumococcal vaccines. Each of the four streptococcal proteins was tested against each of the SARS-CoV-2 proteins yielding 52 pairwise tests. Six of these combinations yielded one or more matches that satisfied all similarity criteria employed here. An additional 30 matches between these pneumococcal proteins and SARS-CoV-2 proteins was found when E was relaxed to 1.0 (TABLE 3) for a total, including the CRM197 matches, of 61.

Table 2. Summary of LALIGN searches set to E = 0.1 comparing SARS-CoV-2 proteins (left-hand column) with vaccine proteins (see TABLE 1 for list of individual proteins). PNEUM = pneumococcal; CRM197 = Cross-Reactive Material 197; Acell PERT = acellular pertussis vaccine; DIPH = diphtheria vaccine; TET = tetanus vaccine; Whole PERT = whole cell pertussis vaccine; BCG = Bacillus Calmette-Guerin, here represented by *M. tuberculosis*. Avg/Pro= average number of matches per protein.

LALIGN E = 0.1	PNEUM	CRM 197	RUB-ELLA	MEAS-LES	MUMPS	Acell PERT	DIPH	TET	POLIO	Men-ingitis
P0DTC1 Repl 1a	15	0	2	2	0	2	0	0	0	0
P0DTC2 Spike Prot	4	0	0	0	0	0	0	0	0	0
P0DTC3 Prot 3a	0	0	0	0	0	0	0	1	0	0
P0DTC4 Env Prot	0	0	0	0	0	0	0	0	0	0
P0DTC5 Memb Prot	0	1	2	0	0	0	0	0	0	0
P0DTC6 NS6 Prot	0	0	0	0	0	0	0	0	0	0
P0DTC7 Prot 7a	0	0	0	0	0	0	0	0	0	0
P0DTC8 NS8 Prot	0	0	0	0	0	0	0	0	0	0
P0DTC9 Nucleoprot	2	0	0	0	0	0	0	0	0	0
P0DTD1 Repl 1ab	0	0	0	0	0	0	0	0	0	0

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 4 September 2020 doi:10.20944/preprints202007.0141.v3

P0DTD2 NS9b Prot	0	0	0	0	0	0	0	0	0	0
P0DTD3 NS Prot 14	0	0	0	0	0	0	0	0	0	0
P0DTD8 Prot 7b	0	0	0	0	0	0	0	0	0	0
Total Matches	21	1	4	2	0	2	0	1	0	0
# Proteins	4	1	6	5	5	9	1	1	7	1
Avg/Prot	5.2	1.0	0.7	0.4	0	0.2	0	1.0	0	0

One significant match at E=0.1 was also found between CRM197 and the membrane protein (P0DTC5) of SARS-CoV-2 (TABLE 2 and FIGURE 1) with an additional nine matches at E=1.0 (TABLE 3). However, no significant similarities at E=0.1 between meningococcal outer membrane protein complex and any SARS-CoV-2 protein (TABLE 2) and only five when E was relaxed to 1.0 (TABLE 3).

Table 3. Summary of LALIGN searches set to E = 1.0 comparing SARS-CoV-2 proteins (left-hand column) with vaccine proteins (see TABLE 1 for list of individual proteins). & Note that the BLAST searches on Whole PERT and BCG were set to E=10 because of the much larger size of the entire genome as compared with the average of 17 proteins searched for the other vaccines. PNEUM = pneumococcal; CRM197 = Cross-Reactive Material 197; Acell PERT = acellular pertussis vaccine; DIPH = diphtheria vaccine; TET = tetanus vaccine; Whole PERT = whole cell pertussis vaccine; BCG = Bacillus Calmette-Guerin, here represented by M. tuberculosis. Avg/Pro= average number of matches per protein.

LALIGN E = 1.0	PNEUM	CRM 197	RUB-ELLA	MEAS-LES	MUMPS	Acell PERT	DIPH	TET	POLIO	Men-ingitis
P0DTC1 Repl 1a	26	4	18	9	6	2	3	1	3	3
P0DTC2 Spike Prot	4	0	5	2	2	0	0	6	1	2
P0DTC3 Prot 3a	2	0	6	1	2	0	0	1	1	0
P0DTC4 Env Prot	0	0	1	0	0	0	0	0	0	0
P0DTC5 Memb Prot	7	2	0	0	1	2	2	1	1	0
P0DTC6 NS6 Prot	0	1	1	0	0	0	0	0	0	0
PODTC7 Prot 7a	0	0	0	0	0	0	0	0	0	0
P0DTC8 NS8 Prot	2	0	0	0	0	0	0	0	0	0
P0DTC9 Nucleoprot	4	1	0	0	1	0	0	0	2	0
P0DTD1 Repl 1ab	6	2	3	0	0	2	0	0	0	0
P0DTD2 NS9b Prot	0	0	0	0	0	0	0	0	0	0
P0DTD3 Prot NS14	0	0	0	0	0	0	0	0	0	0
P0DTD8 Prot 7b	0	0	0	0	0	0	0	0	0	0
Total Matches	51	10	34	12	12	6	5	9	8	5
# Proteins	4	1	6	5	5	9	1	1	7	1
Avg/Prot	12.8	10.0	5.7	2.4	2.4	0.7	5.0	9.0	1.1	5

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 4 September 2020 doi:10.20944/preprints202007.0141.v3

Figure 1. Similarities between the four known or probable pneumococcal vaccine protein contaminants PsaA, PspA, PspC and Gram-positive anchor protein and SARS-CoV-2 proteins as well as CRM197, the modified diphtheria toxin to which pneumococcal conjugate vaccines are attached. Multiple variants for each protein were examined and results provided here are representative of results at E = 0.1.

```
COVID-19 Replicase 1AB 7096 aa vs S. pneumoniae
pspA O34097 653 aa
Waterman-Eggert score: 100;  35.6 bits; E(1) <
8.7e-05
                        90        100        110        120
SP pspA O34097   EKIAEATKEVQQAYLAYLQASNESQRKEADKKIK
                 :||||  ||  : :::  : | | |||: |||||
COVID Rep1A      QKIAEIPKEEVKPFITESKPSVE-QRKQDDKKIK
                    1200       1210       1220
1230

COVID-19 Spike Protein 1273 aa vs. S. pneumoniae
pspA Q9LAZ1 395 aa
Waterman-Eggert score: 60;  23.2 bits; E(1) <  0.05
                     260
SP pspA Q9LAZ1   PLQSKLDTKKAKLSK
                 ||| |||: | :|:|
COVID SP         PLQPELDSFKEELDK
              1140      1150

COVID-19 SP 1273 aa vs. S. pneumoniae pspA B2IRK1
609 aa
Waterman-Eggert score: 62;  23.9 bits; E(1) <
0.049
                 200       210       220
SP pspA B2IRK1   QAKIAELENQVHRLEQDLKDINES
                 :| :: ::::: ||::  |::|||
COVID SP         NASVVNIQKEIDRLNEVAKNLNES
                     1180      1190

COVID-19 Nucleoprotein P59595 422 aa vs. S.
pneumoniae pspA  Q9LAY4
Waterman-Eggert score: 72;  22.7 bits; E(1) <
0.027
                   110       120       130
SP pspA Q9LAY4   QKAFLILREAQEQLSKRPNNKKTAAQQ
                 |::  : :::  : ||:| :|:||::|
COVID NP         QQGQTVTKKSAAEASKKPRQKRTATKQ
                      250       260

COVID-19 Nucleoprotein 417 aa vs. S. pneumoniae
pspA Q9LAZ1
Waterman-Eggert score: 67;  22.1 bits; E(1) <
0.037
              100       110       120       130
SP pspA Q9LAZ1   QLKLKKYLDGRNLSNSSVLKKEMEEAEKKDKEKQ
                 ||: |     |::     :| ||    || || ::|:
COVID NP         QLESKMSGKGQQQQGQTVTKKSAAEASKKPRQKR
                   230       240       250       260

COVID-19 Spike Protein 1273 aa vs. S. pneumoniae
psaA 309 aa
Waterman-Eggert score: 76;  27.2 bits; E(1) <
0.0025
              20        30        40
SP psaA P0A4G2   CASGKKDTTSGQKLKVVATNSIIA
                 ||| : :|:| :: : ||::||||
COVID SP         CASYQTQTNSPRRARSVASQSIIA
                        680       690

COVID-19 Replicase 1AB 7096 aa vs S. pneumoniae
psaA 310 aa
Waterman-Eggert score: 62;  22.7 bits; E(1) <
0.028
                        110       120
SP psaA P42363   FTKLVKNANKVENKDYFAASDGVEV
                 ::| ::|| :||: || |:: /::|
COVID Rep1AB     LNKATNNAMQVESDDYIATNGPLKV
                   1080      1090
```

```
COVID-19 Replicase 1AB 7096 aa vs S. pneumoniae
pspC 792 aa
Waterman-Eggert score: 98;  32.6 bits; E(1) <
0.00053
                    1210      1220      1230
COVID19 Repla    EIPKEEVKPFITESKPSVEQRKQDDKK
                 | || |||| : : ||    : : ||||
SP pspC F2WWN4   EKPKPEVKPQLEKPKPDNSKPQADDKK
                      710       720       730

COVID-19 Replicase 1AB 7096 aa vs S. pneumoniae
pspC 792 aa
Waterman-Eggert score| 79;  26.9 bits; E(1) <
0.027
                    1210      1220      1230
COVID19 Repla    EIPKEEVKPFITESKPSVEQRKQDDK
                 | || |||| : : || |: : :  |
SP pspC F2WWN4   EKPKPEVKPQLEKPKPEVKPQPEKPK
               660       670       680

COVID-19 Replicase 1AB 7096 aa vs S. pneumoniae
pspC 792 aa
Waterman-Eggert score: 77;  26.3 bits; E(1) <
0.039
                        1210      1220
COVID19 Repla      EIPKEEVKPFITESKPSVE
                   | || |||| : / || |:
SP pspC F2WWN4     EKPKPEVKPQLEKPKPEVK
                         670       680
(ADDITIONAL SIMILARITY TO 653-671)

COVID-19 Replicase 1AB 7096 aa vs S. pneumoniae
pspC 792 aa
Waterman-Eggert score| 72;  24.9 bits; E(1) <  0.1
                      1210      1220      1230
COVID19 Repla      EIPKEEVKPFITESKPSVEQRKQDDK
                   | || ||||   / || |: \ :  |
SP pspC F2WWN4     EKPKPEVKPQPEKPKPEVKPQPEKPK
                        560       570
(ADDITIONAL SIMILARITIES TO 565-590; 576-601; 587-
612; 598-623;
609-634; 620-645; 631-656; 681-707)

COVID19 Spike Protein 1273 bp vs. S. pneumoniae
Gram-positive anchor protein Q8DRK2 1161 aa
Waterman-Eggert score: 72;  26.7 bits; E(1) <
0.013
                        280       290
SP GPAP Q8DRK2      GKADLTNLVATKNVDININGL
                    |\   |||| :| |: | |||
COVID19 SPIKE PROT GPKKSTNLVKNKCVNFNFNGL
                      530       540

------------------------------------------------------------

CRM197 Q6NK15 560 aa vs. COVID19 MEMBRANE
PROT 222 aa
Waterman-Eggert score: 51;  20.3 bits; E(1)
<  0.09
                 380
CRM197             DIGFAAYN
                   | |||||:
COVID19 MEMB       DSGFAAYS
                 190
```

Figure 2. Similarities between nine SARS-CoV-2 proteins and 32 proteins from measles, mumps, rubella, polio, Hib, meningitis, diphtheria, pertussis and tetanus vaccines (TABLE 1). 288 pairwise combinations were searched. Only similarities satisfying criteria laid out in Methods are shown with E = 0.1. .

```
RUBELLA NON-STRUCTURAL PROT Q86500 2116 aa
vs. COVID19 REPL1a 4405 aa
Waterman-Eggert score: 83;  29.9 bits; E(1)
<  0.009
                           840        850
RUB NSP Q86500        VVVNAANEGLLAGSGVCGAI
                      |||||||  |  | || ||:
COVID19 REPL1a        VVVNAANVYLKHGGGVAGAL
                        1060       1070

RUBELLA NON-STRUCTURAL PROT Q86500 2116 aa
vs. COVID19 REPL1a 4405 aa
Waterman-Eggert score| 73;  26.8 bits; E(1)
<  0.076
                         940        950
RUB NSP Q86500        PLLGAGVYGWSAAESLRAALAATR
                      ||| ||::|     :|||  : ::|
COVID19 REPL1a        PLLSAGIFGADPIHSLRVCVDTVR
                      1150       1160       1170

RUBELLA STRUCTURAL POLYPROT P08563 1063 aa
vs. COVID19 MEMBRANE PROT 222 aa
Waterman-Eggert score: 54;  21.2 bits;
E(1) <  0.093
                              10
RUB SPP P08563          STTPITMEDLQKALE
                        |: \||:|:|:| ||
COVID19 MEMB            SNGTITVEELKKLLE
                             10

RUBELLA NON-STRUCTRUAL PROT Q86500 2116 aa
vs. COVID19 MEMBRANE PROT 222 aa
Waterman-Eggert score: 58;  22.2 bits;
E(1) <  0.091
                           200
RUB NSP Q86500        LWPVALAAHV
                      ||||:||  |
COVID19 MEMB          LWPVTLACFV
                          60

TETANUS TOXOID P04958 1315 aa vs. COVID19
PROTEIN 3a 275 aa
Waterman-Eggert score| 64;  23.7 bits; E(1)
<  0.026
                    1120       1130
TETANUS TOX P04958 NPLRYDTEYYL
                   ||| ||: |:|
COVID19 PROT 3a    NPLLYDANYFL
                     140
```

```
MEASLES HEMA P08362 617 aa vs. COVID19 REPL1a
4405 aa
Waterman-Eggert score| 70;  25.8 bits; E(1)
 <  0.045
                             530        540
MEASLE HEM P08362    YVLATYDTSRVEHAVVYYVYSPS
                     ||| : || |||     |::  ||
COVID19 REPL1a       YVLPNDDTLRVEAFEYYHTTDPS
                  1620       1630       1640

MEASLES FUSION GLYCOPROTEIN Q786F3 550 aa vs.
COVID19 PROT REPL1A 4405 aa
Waterman-Eggert score| 65;  24.5 bits; E(1)
 <  0.097
                          500
MEASLES FGP Q786F3   IVYILIAVCLGGLI
                     | ::|::|||| ||
COVID19 REPL1A       IWFLLLSVCLGSLI
                           2240

PERTUSSIS TOXIN 1 P04977 269 aa vs. COVID19
REPL1a 4405 aa
Waterman-Eggert score| 69;  25.3 bits; E(1)
<  0.029
                           220
PERT TOX1 P04977     YTSRRSVASIVGTL
                     |||: :|||:: ||
COVID19 REPL1a       YTSKTTVASLINTL
                       1430

PERTUSSIS TOXIN 4 P0A3R5 152 aa vs. COVID19
Repla 4405 aa
Waterman-Eggert score| 62;  23.2 bits; E(1)
<  0.069
                          10         20         30
PERT TOX4 P0A3R5     FPTRTTAPGQGGARRSRVRALAWLLA
                     |  | || : |          |  |||| |
COVID19 REPL1A       FVDRQTAQAAGTDTTITVNVLAWLYA
                   3450       3460       3470
```

FIGURE 2 displays the results for the pairwise tests of the thirteen SARS-CoV-2 proteins with the additional bacterial and viral proteins listed in TABLE 1 that are present in measles, mumps, rubella, polio, diphtheria, pertussis, and tetanus vaccines, for a total of 32 microbial proteins. Of these, six yielded one or more significant similarities for a total of nine matches out of 416 possible pairwise combinations (TABLE 2). When the E value was relaxed to 1.0 (TABLE 3), an additional 81 matches were found, most notably between rubella vaccine proteins and SARS-CoV-2 proteins.

Results from the BLAST searches on whole bacteria are presented in TABLE 4. The 3997 M. tuberculosis proteins yielded five significant similarities at an E value of 1.0 or less when compared with the 13 SARS-CoV-2 proteins (51,961 combinations) (FIGURE 3). These matches are of roughly equivalent quality to those of the LALIGN searches conducted on the other vaccine proteins described above. The sequences are listed in FIGURE 3. Raising the E value to 10 and lowering the

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 4 September 2020 doi:10.20944/preprints202007.0141.v3

Waterman-Eggert (W-E) score to 40 increased the total number of matches (still including at least six identities in a stretch of 10 amino acids) to 36. These matches appear to be equivalent in quality to those found for E=1.0 for the LALIGN searches. Similarly, the whole pertussis proteome (3260 proteins) yielded only six matches at E=0.1 and the W-E score at 50 (TABLE 2 and FIGURE 4), which increased to 55 when the W-E score was lowered to 40 and E was raised to 1.0 (TABLE 3). However, these results do not differ significantly from those obtained from commensal and probiotic control bacteria (TABLE 4): the average number of matches per protein for the tuberculosis and pertussis bacteria at E = 1.0 was 0.0015 and at E = 10.0, 0.013 whereas the average number of matches per protein for the control bacteria at E = 1.0 was 0.0015 and at E = 10.0, 0.014. These results suggest that the rate of matches between *M. tuberculosis* and SARS-CoV-2 is what can be expected as the result of randomness rather than any of the tested bacteria expressing particular proteins sequences of relevance to the current study.

FIGURE 3: SARS-CoV-2 protein similarities with Mycobacterium tuberculosis (Mtb). Note that BCG, unlike the vaccines in Figures 1 and 2 that are composed of one to seventeen proteins, is composed of 3993 proteins so that even given the somewhat larger number of significant similarities displayed here, the probability of them being major antigens is extremely small. Note also that because of the size of the BCG proteome, BLAST (rather than LALIGN,as in Figures 1 and 2), was used to find these similarities and a cut-off value for significance of E=1.0 rather than 0.1 was used.

SARS-CoV-2 P0DTC1 (Repl 1a) vs Mtb P9WK29, uncharacterized protein Rv1899c

```
Waterman-Eggert score (80),   Expect = 6e-04
P0DTC1 1051   KVKPTVVVNAANVYLKHGGGVAGALNKATNNAMQVES   1087
              K++     + NAAN  L+H GGVA A+ +A     +Q ES
Mtb    201    KLELDAITNAANTRLRHAGGVAAAIARAGGPELQRES   237
```

SARS-CoV-2 P0DTD1 (Repl 1b) vs Mtb P96287 AAA domain-containing protein

```
Waterman-Eggert score (72),   Expect = 0.016
PODTD1 5602   STLQGPPGTGKSHFAIGLAL   5621
              S     PPGTGK+H A+GLA+
Mtb    84     SCFWAPPGTGKTHLAVGLAI   103
```

SARS-CoV-2 P0DTC2 (Spike Protein) vs Mtb P9WK23 4-alpha-glucanotransferase

```
Waterman-Eggert score (56),   Expect = 0.91
PODTC2 222  ALEPLVDLPIGINITRFQTLLALHRSYLTPGDSSSGWTAGAAAYYVGYLQPRT   274
            A+   LVDLP      + R +T +   H    L    D  S W A  AA   + +   PR+
Mtb    256  AIPELVDLPKRGRVQRLRTNVQQHADQLDTIDRDSAWAAKRAALKLVHRVPRS   308
```

SARS-CoV-2 P0DTC7 (Protein 7a) vs Mtb P9WJ63 16S/23S rRNA (cytidine-2'-O)-methyltransferase TlyA

```
Waterman-Eggert score (49),   Expect = 0.81
PODTC7 68   PDGVKHVYQLRARSV   82
            P GV H  QLRARSV
Mtb    194  PGGVVHDPQLRARSV   208
```

SARS-CoV-2 P0DTC9 (NucleoProtein) vs Mtb I6X9V3 GCV_T domain-containing protein

```
Waterman-Eggert score (55),   Expect = 0.83
PODTC9 80   PDDQIGYYRRATRRIRGG   97
            P D +G   RRA   R+RGG
Mtb    349  PADDVGAGRRAVERLRGG   366
```

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 4 September 2020 doi:10.20944/preprints202007.0141.v3

FIGURE 4: SARS-CoV-2 protein similarities with Bordetella pertussis polyprotein (UniProte accession number UP000002676). Note that whole B. pertussis is used as a vaccine. It is comosed of 3260 proteins so that the probability that the matches shown are major antigens is extremely small. Note also that because of the size of the size of the B. pertussis proteome, BLAST (rather than LALIGN,as in Figures 1 and 2), was used to find these similarities and a cut-off value for significance of E=1.0 rather than 0.1 was used, as was the case with M. tuberculosis (FIGURE 3) as well.

SARS-CoV-2 P0DTD1 (Repli 1b) vs. B. pertussis Q7VXF9 MOSC domain-containing protein

```
Watermann-Eggert score (61),  Expect = 0.62
PERTUSSIS          5761   FLGTCRRCPAEIVDTVSALVYD  5782
                          F+  C RCP    VD V+A VYD
P0DTD1 Rep1ab      225    FVKPCTRCPMSNVDQVTAEVYD  246
```

SARS-CoV-2 P0DTC5 (Membrane) vs. B. pertussis Q7VT43 Amidase domain protein

```
Watermann-Eggert score (57),  Expect = 0.51
PERTUSSIS Q7VT43  250  TPGDSSSGWTAGAAA  264
                        TPGDSSSG  A++AA
COVID19 Membrane  143  TPGDSSSGSAAAVAA  157
```

SARS-CoV-2 P0DTC5 (Membrane) vs. B. pertussis Q7VV25 Putative export protein

```
Watermann-Eggert score (51),  Expect = 0.45
PERTUSSIS Q7VV23     46  LYIIKLIFLWLLWPVTLACF  65
                         L ++ + F WLLWP    A F
COVID19 Membrane     16  LIVVTIAFAWLLWPFYGAVF  35
```

SARS-CoV-2 P0DTC6 (NS6) vs. B. pertussis Q7VVU5 Succinate-CoA ligase

```
Watermann-Eggert Score (45),  Expect = 0.72
PERTUSSIS Q7VVU5    49   YSQLDEEQPMEID  61
                         Y  LDEE P EI+
COVID19 NS6         232  YRDLDEEDPAEIE  244
```

SARS-CoV-2 P0DTC9 (Nucleoprotein) vs. B. pertussis Q7VVM8 MFS domain protein

```
Watermann-Eggert Score (56),  Expect = 0.47
PERTUSSIS Q7VVM8    305  AQFAPSASAFFGMSRIG  321
                         A F PSA AFFG S +G
COVID 19 NUCL       312  AVFTPSALAFFGASLVG  328
```

SARS-CoV-2 P0DTD2 (Protein 9b) vs. B. pertussis Q7VUM1 HTH lysR-domain protein

```
Watermann-Eggert Score (51),  Expect = 0.38
PERTUSS Q7VUM1      11  ALRLVDPQIQLAVTRMENAVG  31
                        AL L   P +  A+ R+E AVG
COVID19 PROT 9B     42  ALHLSQPAVSQALKRLEQAVG  62
```

TABLE 4: Summary of BLAST search result matches between SARS-CoV-2 proteins (left-hand column) and whole bacteria at E = 1.0 (shaded left-hand columns) and E = 10.0 (unshaded right-hand columns): *Bordetella pertussis* (whole PERT); *Mycobacterium tuberculosis* (BCG); *Clostridium leptum* (C. lept); *Escherichia coli* (E. coli); *Lactococcus lactis* (L. lact); *Lactobacillus paracasei* (L. para). Note that in contrast to the LALIGN searches (TABLES 2 and 3) the BLAST searches were set to E=1 or E = 10 because of the much larger size of the entire genome as compared with the average of 17 proteins searched for the other vaccines (compare sequences in FIGURES 3 and 4 to FIGURES 1 and 2). Avg/Pro= average number of matches per protein.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 4 September 2020 doi:10.20944/preprints202007.0141.v3

BLAST, E = 1.0 and 10.0	Whole PERT	BCG	C. lept	E. coli	L. lact	L. para	Whole PERT	BCG	C. lept	E. coli	L. lact	L. para
P0DTC1 Repl 1a	0	5	0	2	0	1	5	4	4	6	4	3
P0DTC2 Spike Protein	1	0	0	0	0	1	9	4	1	6	1	4
P0DTC3 Protein 3a	0	0	0	1	2	0	10	6	7	5	4	4
P0DTC4 Env Protein	0	0	0	0	0	0	2	0	2	1	0	0
P0DTC5 Memb Prot	1	0	1	1	1	1	2	6	4	6	9	10
P0DTC6 NS6 Protein	1	0	0	0	1	0	4	1	2	0	10	5
P0DTC7 Protein 7a	0	0	1	1	0	1	3	2	2	5	6	3
P0DTC8 NS8 Protein	0	0	0	1	0	0	2	1	2	3	1	3
P0DTC9 Nucleoprot	1	0	0	0	0	0	7	4	1	3	2	2
P0DTD1 Repl 1ab	1	0	0	0	1	0	5	4	3	3	4	5
P0DTD2 NS9b	0	0	0	0	0	0	0	3	4	5	3	3
P0DTD3 NS Protein 14	0	0	0	0	0	0	0	1	1	3	0	2
P0DTD8 Protein 7b	1	0	0	0	0	0	6	0	0	0	0	0
Total Matches	6	5	2	6	5	4	55	36	33	46	44	42
# Proteins	3260	3997	2482	4403	2225	2708	3260	3997	2482	4403	2225	2708
Avg/Prot	0.002	0.001	0.001	0.001	0.002	0.002	0.017	0.009	0.013	0.010	0.020	0.016

Large differences in the number of matches was found between pneumococcal proteins and those from other protein antigen vaccines for LALIGN E=0.1 group (TABLE 2, FIGURES 1 and 2). All four of the pneumococcal proteins and the CRM197 protein had significant similarities (i.e., meeting the similarity criteria laid out in the Methods) to at least one of the thirteen SARS-CoV-2 proteins. Altogether, seven of the 65 possible permutations of pneumococcal protein pairs yielded significant similarities, or 10.8 percent. In contrast, only eight of the 35 viral and bacterial vaccine proteins other than whole-cell pertussis and M. tuberculosis had significant matches to any of the nine SARS-CoV-2 proteins (1.8% of the 455 pairwise comparisons). The four pneumococcal proteins yielded 21 significant matches with SARS-CoV-2 proteins, for an average of 5.25 per pneumococcal protein, while the 35 other vaccine proteins yielded only nine significant matches, for an average of 0.26 per protein. In other words, at the E = 0.1 criterion, the probability of a match leading to cross-reactivity is over 20 times more likely for pneumococcal proteins than for those from other vaccines.

The E = 1.0 data (TABLE 3) yielded similar results. The pneumococcal proteins exhibited a total of 61 matches (including CRM197) with SARS-CoV-2 proteins for an average of 12.2 matches per protein. The rest of the vaccines (other than whole cell pertussis and BCG) exhibited 90 total matches spread out over 35 proteins for an average of 2.5 matches per protein. The 61 pneumococcal matches were found among 23 of the 65 permutations with SARS-CoV-2 proteins, or 35.2 percent. In contrast, the 90 other vaccine matches were spread out over 53 of the 455 pairwise permutations, representing 11.6 percent of the possibilities. In other words, using the E = 1.0 criterion as a cutoff, it is three times more likely that pneumococcal proteins will result in a cross-reactive match than for other proteins. In this instance, rubella antigens account for more than thirty percent of the non-pneumococcal matches making rubella the next best candidate for protecting against SARS-CoV-2 infection.

The whole cell vaccines were treated separately from the limited antigen vaccines because BLAST was used rather than LALIGN to perform the searches and because the the average number

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 4 September 2020 doi:10.20944/preprints202007.0141.v3

of individual vaccine protein matches to SARS-CoV-2 was very different: for the whole cell bacteria, it was 0.0145 with an SD of 0.0055 whereas the individual vaccine proteins (using the E=1.0 data in Table 3) is about 5.4 with a standard deviation of 2.6. For *M. tuberculosis*, for example, the best rate of matches was 40 out of 51,961 combinations [E = 10], or 0.08 percent, with an average of one match per 100 M. tuberculosis proteins. At worst, using E=1.0, there were only 5 matches out of 51,961 combinations or 0.01 percent with one match per every 800 M. tuberculosis proteins. The pertussis results were very similar. On a per-protein basis, these two bacteria resulted in rates of matches that were two orders of magnitude lower than the other proteins tested (TABLES 2 and 3). Thus, the percent of whole-bacteria matches (TABLE 4) is clearly very much lower than the percent of matches for the limited-antigen vaccines listed in TABLES 2-3. The paucity of matches on a per-protein basis resulting from the tuberculosis and pertussis bacteria comparisons is itself noteworthy, strongly suggesting that the quality of matches reported in FIGURES 1 and 2 for the other vaccines are intrinsically extraordinary and the pneumococcal (both E = 0.1 and E = 1.0) and rubella (E = 1.0) results particularly so.

4. Discussion

The Results of this study indicate that while pneumococcal vaccines are primarily composed of polysaccharides there are no obvious structural homologies between these polysaccharides and SARS-CoV-2 glycosylations. The absence of such homologies does not rule out antigenic cross-reactivity between these polysaccharides but makes their identification difficult using anything other direct tests of whether SARS-CoV-2 antibodies recognize pneumococcal polysaccharides or whether pneumococcal antibodies recognize SARS-CoV-2. Such tests might be worth conducting if only as controls for studies of possible cross-reactivity between proteins found in pneumococcal vaccines and SARS-CoV-2 proteins.

CRM197, which is used to conjugate pneumococcal polysaccharides in conjugate vaccines such as the Prevnar series, and pneumococcal proteins known to contaminate the vaccines significantly, both mimic SARS-CoV-2 proteins (FIGURE 1), satisfying rigid similarity and antigenicity constraints, though there are many more high-quality matches between the pneumococcal proteins than with CRM197. The Results point specifically to potential cross-reactivity between SARS-CoV-2 proteins and the pneumococcal proteins PspA and PsaA, which are known to contaminate polysaccharide-based pneumococcal vaccines [5-7] as well as PspC, which it is reasonable to assume is another such contaminant since it derives from the same outer membrane protein complex and is highly cross-reactive with the antibodies against PspA used to demonstrate the presence of PspA in vaccines [10, 11]. Such cross-reactivity would be consistent with epidemiological studies suggesting a protective effect of pneumococcal vaccination against SARS-CoV-2 [1,2]. Since the CRM197 protein is used to conjugate some *Haemophilus* and meningitis vaccines, these vaccines may also provide some cross-reactive protection against SARS-Cov-2 proteins (FIGURE 1), a result that is consistent with the findings of Pawlowski, et al. [2]. Further clinical and experimental tests of whether these vaccines elicit antibodies cross-reactive with SARS-CoV-2 proteins are clearly needed.

It is important to emphasize that the fact that a microbe expresses an antigen that is sequentially similar to a SARS-CoV-2 protein is not sufficient to guarantee that the two will elicit cross-reactive immunity or, for that matter, any immune response whatsoever, since that sequence may not be processed as an antigen by macrophages or presented to T cells. The full range of determinants of antigenicity are as yet unknown. Among the key factors seem to be the concentration of the antigen; how dissimilar it is from its host; where the antigen is expressed within a protein (e.g., whether it is freely accessible in a random loop or protected within a pleated beta sheet); how the antigen is presented to the immune system (e.g., by ingestion, inoculation, infection); and the inflammatory context in which the antigen is processed (e.g., in the presence of an adjuvant or bystander infection) [21]. In this context, it is notable that the rate of SARS-CoV-2

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 4 September 2020 doi:10.20944/preprints202007.0141.v3

matches to pneumococcal vaccine antigens is very high (14%) so that the probability of the immune system encountering a cross-reactive antigen is also reasonably high, as is the probability for CRM197-conjugated vaccines and rubella and measles, whereas the rates of matches to mycobacterial and pertussis antigens is very low (less than 0.1%) making the probability of encountering a cross-reactive antigen very low. However, the data presented here does not rule out that possibility and it is therefore striking that the rate of SARS-CoV-2 matches to pertussis and mycobacterial proteins is not significantly different than to the commensal and probiotic control bacteria *L. lactis, L. paracasei, C. leptum* and *E. coli* (TABLE 4). On the one hand, these data could be interpreted to mean that mycobacteria and pertussis are as unlikely as the commensal and probiotic bacteria to protect against SARS-CoV-2. On the other hand, if mycobacteria or pertussis bacteria are sufficient to induce protection against SARS-CoV-2, one might argue that these commensal and probiotic bacteria could do so as well; since everyone encounters them, protection against SARS-CoV-2 should then be universal. However, several factors undermine this latter conclusion. One is that the mode of presentation of vaccines to the host immune system is very different than that of commensal and probiotic bacterial antigens. Vaccines either actively infect the host (e.g., polio or influenza), or are (more often) inoculated at very high concentrations; in both cases the resulting tissue damage initiates an immune response. Probiotic and commensal microbes, in contrast, are retained (in healthy people) with the gut and do not cause tissue damage or initiate an active immune response [22]. Additionally, the immune system often develops tolerance for commensal and probiotic organisms and such organisms express large numbers of antigens that mimic host antigens, including T cell receptors and human leukocyte antigens, thereby camouflaging themselves from immune surveillance [23-27].. Thus, while any given commensal or probiotic microbe has some small probability of expressing antigens that could potentially protect against SARS-CoV-2 infection, their general inability to elicit active immunity militates against this being a likely scenario.

The concentration of antigen presented to the immune system is also a determinant of whether an active immune response results so that microbes expressing very large numbers of antigens in very small quantities are unlikely to elicit a strong immune response to most of them. The concentration of protein contaminants in pneumococcal vaccines is clearly sufficient to induce immunity. CRM197 is present in equal amounts to the capsular polysaccharides in the vaccines and is present because it is known to be highly antigenic. In Prevnar-13, for example, there are 30.4 µg of capsular polysaccharides and 34.0 µg of CRM197 for a total of 64.4 micrograms of antigen per dose [28]. Protein contaminants may make up an additional 3%, or 1.92 µg, of antigenic material according to WHO guidelines and confirmed by laboratory analysis [5-7]. This 1.92 µg of protein is virtually identical to the 2.2 µg of each of twelve of the capsular polysaccharides present (plus 4.4 µg of serotype 6) or the 2.3 micrograms of CRM197 conjugated to each polysaccharide type [28] and is therefore sufficient to induce an immune response, especially since PspA and PspC are strongly antigenic and cross-reactive. Pneumovax-23, in contrast, has 25 µg of each capsular polysaccharide, adding up to a total of 575 µg of antigen [29]. The three percent protein contamination allowed by WHO [5-7] could result in 17.25 µg of total PsaA, PspA and PspC per dose, which is certainly sufficient to induce immunity. For comparison, each 0.5-mL dose of Adacel®, a diphtheria-tetanus-pertussis vaccine (Sanofi Pasteur) contains only 2.5 µg detoxified pertussis toxoid, 5 µg FHA, 3 µg pertactin and 5 µg FIM acellular pertussis antigens [30].

In addition to being present in concentrations that could induce protective immunity, the pneumococcal-SARS-CoV-2 similarities reported here satisfy multiple criteria involving sequence identities and search parameters for predicting potential antigenic cross-reactivity so that it is possible that pneumococcal vaccination can protect individuals against SARS-CoV-2 disease. Evidence of protection against SARS-CoV-2 by T cells reactive to unidentified, cross-reactive microbes has been reported by several groups [31-33]. The studies report that 40- to 60% of people unexposed to SARS-CoV-2 had SARS-CoV-2-reactive CD4+ T cells. This cross-reactivity is proposed to result from prior exposure to coronaviruses that cause colds but this hypothesis has not

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 4 September 2020 doi:10.20944/preprints202007.0141.v3

yet been tested [33]. Moreover, the studies also report that this cross-reactive immunity is greatest in young people and least in older people, which is not consistent with cold virus exposures, nor is the fact that over 90% of people have T-cells reactive to cold viruses but few seem to be immune to SARS-CoV-2 [33]. Such waning immunity is, however, consistent with waning childhood vaccination immunity and particularly for vaccinations such as pneumococci that are not universal. In light of the data presented here, it is therefore possible that at least some proportion of individuals with cross-reactive immunity developed it through exposure to pneumococcal vaccinations. Such cross-reactivity would also explain the epidemiological observation that pneumococcal vaccination rates correlate inversely with rates of serious SARS-CoV-2 disease and death, but that vaccination rates with other commonly used vaccines (DTP, MMR, polio, meningitis, and BCG), do not [1].

One might ask whether the immunity conferred by pneumococcal (and perhaps other) vaccines is sufficient to prevent SARS-CoV-2 infection completely. The current study is incapable of addressing that question meaningfully but the fact that the vast majority of similarities between pneumococcal proteins and SARS-CoV-2 involve the replicase (TABLES 2 and 3) suggests that any protection would be reactive rather than preventative. The reason for this is that the replicase is not expressed until cells are infected so that pneumococcal-related immunity would mainly come into play only at that point. This factor might explain why many people seem to become infected with SARS-CoV-2 and remain infectious without themselves displaying symptoms of COVID-19 [34-38]. Indeed, increasing evidence indicates that the primary protection against SARS-CoV-2 is T-cell mediated rather than antibody-mediated [31-33], suggesting that control of the infection is at the level of cellular infection rather than against free virus. Nonetheless, it is notable that the next most prevalent set of SARS-CoV-2-pneumococcal protein similarities after the replicase involve the viral spike protein (TABLES 2 and 3), which is a major target for antibodies and which might, therefore mediate SARS-CoV-2 infectivity.

The observation that viral and bacterial proteins exhibit antigens similar enough to be cross-reactive may be surprising but it is not novel. Härkönen, et al. [39] found that rabbit antibodies to HSP65 of *Mycobacterium bovis* (from which BCG is derived) recognized capsid protein VP1 of coxsackievirus A9, VP1, and/or VP2 of coxsackievirus B4. Misko, et al. [40] demonstrated that Epstein-Barr virus mimicked a *Staphylococcus aureus* replication initiation protein and induced antibodies cross-reactive with it. Trama, et al., [41] and Williams, et al. [42] have documented antibodies against the gp41 protein of human immunodeficiency virus that cross-react with commensal bacteria in the human gut. Ross, et al. [43] reported that sera from chickens inoculated with infectious bursal disease viruses or infectious bursal disease vaccines cross-reacted with *Mycoplasma gallisepticum* and *Mycoplasma synoviae*. And Bordenave [44] found that antibodies against *Salmonella abortusequi* also recognized tobacco mosaic virus. In short, while the phenomenon may be rare – and, indeed, the data reported here suggests that such similarities may occur at a rate as high as 1/70 pairwise protein combinations or as low as 1/1000 -- bacterial antigens are known to occasionally induce antibodies that cross-react with viral antigens or vice versa. This observation is consistent with the fact that every possible sequence of five amino acids has been shown to appear randomly in the microbial proteome [45, 46]. Completely unrelated microbes should, therefore, have a small, but finite, probability of expressing identical antigens capable of inducing cross-reactive immune responses. The question becomes one of whether these antigens are ever encountered by the host and presented to the immune system in a way that initiates cross-reactive immunity.

The almost completely negative results reported here for antigenic mimicry between SARS-CoV-2 proteins and proteins from measles, mumps, diphtheria, pertussis and tetanus at E = 0.1 (TABLE 1), and the relatively low rate of similarities with poliovirus at E = 1.0 (TABLE 2), are consistent with the lack of association between these vaccines and SARS-CoV-2 rates of disease or death [1], although Pawlowski, et al. [2] found some protective effect from polio vaccination and

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 4 September 2020 doi:10.20944/preprints202007.0141.v3

the measles-mumps-rubella (MMR) combination vaccine. The current study would suggest that the rubella component of MMR is the major protective agent, though measles also exhibits some high-quality antigenic similarities to SARS-CoV-2. Indeed, Franklin, et al., [47] also report significant similarities between both rubella and measles proteins and SARS-CoV-2, and their key results were independently reproduced here in FIGURE 2. Additionally, Gold [48] has also proposed that the measles-mumps-rubella vaccine may confer protection against SARS-CoV-2. However, there are significantly fewer similarities between measles and rubella proteins and those of SARS-CoV-2 proteins (and none with mumps proteins) than there are with pneumococcal proteins making pneumococci a much higher probability source of protection. Moreover, epidemiological evidence does not support measles containing vaccines (which often include rubella) as protective against SARS-CoV-2, though using measles-containing vaccines as Root-Bernstein (2020) did, may hide important rubella-related protection since not all measles-containing vaccine include rubella and rubella vaccination can be performed independently from measles vaccination. The suggestion that polio vaccine be tested as a SARS-CoV-2 [49] is likewise not well-supported by either the data presented here, which found only one significant similarity between polio proteins and SARS-CoV-2 proteins at E = 0.1 and five at E = 1.0 (TABLES 1 and 2 and FIGURE 2) or by epidemiological data [1] though, once again, Pawlowski, et al. [2] found some protective effect in children.

The data presented here must be interpreted both probabilistically -- which is to say as a guide to whether any particular vaccine has a greater or lesser probability of providing antigens that are both cross-reactive and protective against SARS-CoV-2 infection or complications – and antigenically, which is a measure of how strong an immune response a sequence actually elicits. Using both criteria, pneumococcal vaccine antigens are the most probable candidates for providing such protection since there are many matches and the pneumococcal proteins are known to be highly antigenic. The rubella antigens the next most likely for the same reasons. However, we cannot know for certain until the appropriate immunological cross-reactivity studies are conducted to determine both whether antibodies against the vaccine antigens recognize SARS-CoV-2 antigens and protect against infection, and whether SARS-CoV-2 antibodies recognize the potentially cross-reactive antigens identified in FIGURES 1-4.

The criteria just described apply equally to considerations of whether there is cross-reactivity to BCG vaccine. Tuberculosis (BCG) vaccination has also been proposed to protect against SARS-CoV-2 [50]. While BCG vaccination was purported to be associated with SARS-CoV-2 protection in several epidemiological studies (reviewed in [51]) that result was not replicated in others (e.g., [1, 2, 52] and serious concerns about methodologies have called into question the association [51, 53]. The current study leads to the conclusion that BCG protection against SARS-CoV-2 is unlikely. While between 5 (E = 0.1) and 40 (E = 1.0) similarities were found between *M. tuberculosis* proteins and SARS-CoV-2 proteins, this number is insignificant in relation to the number of proteins expressed by *M. tuberculosis* and BCG (approximately 4000). This paucity of significant *M. tuberculosis similarities* (0.04%) as compared with the high incidence of pneumococcal similarities (11.6 -14 %) makes it probable that pneumococcal proteins will induce cross-reactive antibodies and extremely unlikely that any of the *M. tuberculosis* antigens will do so. Indeed, none of the *M. tuberculosis* proteins identified in FIGURE 3 are among the known dominant antigens expressed by either *M. tuberculosis* infection or BCG vaccination [54-58].

The question of whether pertussis antigens may protect against SARS-CoV-2 is more complicated than that for BCG. There appear to be no epidemiological studies associating pertussis vaccination with protection against SARS-CoV-2 infection or death and the one study that has looked for such an association found none [1]. However, while acellular pertussis vaccines have a very small number of sequences that are potentially cross-reactive with SARS-CoV-2 proteins, the whole cell vaccine, which is still available in some countries, has many matches specifically to the SARS-CoV-2 spike protein, which is a major target of neutralizing antibodies, than other vaccines

(TABLE 4). The difficulty is that with 3260 proteins in the whole cell vaccine, the probability that any of these potentially cross-reactive sequences are actually processed as major antigens inducing significant antibody responses is small, particularly compared to pneumococcal and rubella vaccines (TABLES 1 and 2). However, some of these proteins have been incorporated into the acellular pertussis vaccines and are known to be highly antigenic. Thus, total number of matches is probably a less useful predictor of antigenic cross-reactivity than whether the potentially cross-reactive proteins are known to be highly antigenic, as is the case with the pneumococcal and rubella proteins. Again, theory can be a guide here, but experiment will provide the final answers.

Finally, it must be mentioned that the correlations between pneumococcal vaccination (and perhaps other vaccinations) and decreased risk of SARS-CoV-2 cases and deaths may be due not to cross-reactivity between pneumococcal (or other vaccine) antigens and SARS-CoV-2 antigens but rather to protection against super-infection of SARS-CoV-2 by pneumococci and other bacteria. While it is common to attribute all of the symptoms of COVID-19 to SARS-CoV-2 infection, a rapidly expanding literature is demonstrating that, as with influenza [59, 60], serious COVID-19 cases are characterized by bacterial super-infections of which pneumococci, *Haemophilus influenzae* and Mycoplasmas are the most common [61-69]. For example, a recent study from China found that 60% of COVID-19 patients had streptococcal infections, about 55% *Klebsiella pneumoniae* infections and 40% had Hib [70]. Indeed, severe COVID-19 cases are characterized by elevated procalcitonin levels [68, 69] and by eosinopenia [66, 67], both of which are diagnostic for disseminated bacterial infections [71, 72]. If this is the case, then pneumococcal and Hib vaccination may not prevent SARS-CoV-2 infection but should decrease the probability of developing the complications associated with severe COVID-19 disease.

5. Conclusions

To conclude, there are many reasons to investigate whether pneumococcal, Hib, meningitis and rubella vaccination may protect against SARS-CoV-2 infection or complications. Epidemiologically, a strong inverse association of pneumococcal vaccinations with rates of SARS-CoV-2 rates of disease and death has been documented by two studies [1, 2]. The epidemiological association makes sense in terms of the particular proteins found in pneumococcal vaccines that are identified in this study as being potentially protective. These are CRM197, PspA, PsaA and PspC, all proteins known to be highly antigenic [73]. Since CRM197 is also found in Hib vaccines, which have also been associated with protection against SARS-CoV-2 [2], its cross-reactivity with SARS-CoV-2 proteins should be investigated. The other pneumococcal proteins (PspA, psaA and PspC) are under active investigation as more effective and broadly protective pneumococcal vaccine components to replace the polysaccharide-based vaccines [74-77]. Some of these vaccine candidates are already in human trials [77, 78]. Thus, it should be possible rapidly and readily to determine whether such pneumococcal protein-based vaccines can be effective mitigators of SARS-CoV-2 disease and these vaccines may provide needed protection until a SARS-CoV-2 vaccine is produced in sufficient quantities to be effective worldwide. And finally, rubella vaccination should also be investigated further since rubella proteins have the second highest rate of similarities to SARS-CoV-2 proteins in this study and rubella vaccination has been reported to have some protective efficacy against SARS-CoV-2 [2].

Because pneumococcal vaccination has the highest degree of protection in both studies that have compared it with other vaccines [1, 2], it seems logical to focus current efforts on this type of vaccination. Regardless of the efficacy of such pneumococcal vaccines in protecting against serious SARS-CoV-2 infection, increased use of pneumococcal vaccination should be urged because the world will be facing dual epidemic/pandemics this coming Fall and Winter and perhaps for many years hereafter, involving concurrent influenza and SARS-CoV-2 epidemic/pandemics. Increasing pneumococcal and Hib (which also contains CRM197) vaccination coverage has been demonstrated to be one of the most effective means to lower the incidence of pneumonias and intensive care unit cases following influenza infections [79, 80]. At a minimum, decreasing the rates of invasive pneumococcal and *Haemophilus influenzae* superinfections following influenza infections will free up badly needed resources, personnel and intensive care units for treating SARS-CoV-2 patients. Several nations have already adopted, or are considering, policies to increase pneumococcal vaccination coverage for just this reason [81-84]. If the current research is accurate, Hib should be added to this list and nations adopting these policies may also benefit in having fewer serious SARS-CoV-2 cases because of protection from cross-reactive antigens. This is a no-lose and possibly win-win situation.

Funding: This research received no external funding.

Acknowledgments: None

Conflicts of Interest: The author declares no conflict of interest.

References

1. Root-Bernstein, R. Age and location in severity of COVID-19 pathology: Do lactoferrin and pneumococcal vaccination explain low infant mortality and regional differences? *BioEssays*, **2020**. https://doi.org/10.1002/bies202000076. Preprint version: Why Infants Rarely Die of COVID-19 and Morbidity and Mortality Rates Vary by Location: Pneumococcal and Hib Vaccinations as Possible Means to Mitigate Future Pandemics. MDPI Preprints doi: 10.20944/preprints202004.0233.v2
2. Pawlowski, C.; Puranik, A.; Bandi, H.; Venkatakrishnan, A.J.; Agarwal, V.; Kennedy, R.; O'Horo, J.C.; Gores, G.J.; Williams, A.W.; Halamka, J.; Badley, A.D.; Soundarajan V. Exploratory analysis of immunization records highlights decreased SARS-CoV-2 rates in individuals with recent non-COVID-19 vaccinations. MedXriv preprint, **2020**. https://www.medrxiv.org/content/10.1101/2020.07.27.20161976v2; https://doi.org/10.1101/2020.07.27.20161976
3. Watanabe, Y.; Allen, J.D.; Wrapp, D.; McLellan, J.S.; Cripin, M. Site-specific glycan analysis of the SARS-CoV-2 spike. *Science.* **2020**; *eabb9983*. doi: 10.1126/science.abb9983
4. Shajahan, A.; Supekar, N.T.; Gleinich, A.S.; Azadi, P. Deducing the N- and O-glycosylation profile of the spike protein of novel coronavirus SARS-CoV-2. *Glycobiology*, **2020**; *cwaa042*, https://doi.org/10.1093/glycob/cwaa042
5. WHO. Recommendations to Assure the Quality, Safety and Efficacy of Pneumococcal Conjugate Vaccines. World Health Organization, Geneva: **2009**; http://www.who.int/biologicals/areas/vaccines/pneumo/Pneumo_final_23APRIL_2010.pdf?ua=1
6. Lee, C.; Cuh, H.J.; Park, M.; Kim, R.K.; Whan, Y.H.; Choi, S.K.; Baik, Y.O.; Park, S.S.; Lee, I. Quality improvement of capsular polysaccharide in *Streptococcus pneumoniae* by purification process optimization. *Front. Bioeng. Biotechnol.* **2020**; *8;* 39. doi: 10.3389/fbioe.2020.00039
7. Morais, V.; Dee, V.; Suárez, N. Purification of capsular polysaccharides of *Streptococcus pneumoniae*: Traditional and new methods. *Front. Bioeng. Biotechnol.*, **2018**; https://doi.org/10.3389/fbioe.2018.00145
8. Yu, X.; Sun, Y.; Frasch, C.; Concepcion, N.; Nahm, M.H. Pneumococcal capsular polysaccharide preparations may contain non-C-polysaccharide contaminants that are immunogenic. *Clin. Diagn. Lab. Immunol.* **1999** *Jul; 6(4);* 519–524.
9. Yu, J.; Briles, D.E.; England, J.A.; Hollingshead, S.K.; Glezen, W.P.; Nahm, M.H. Immunogenic protein contaminants in pneumococcal vaccines. *JID* **2003** *187;* 1019-1023.
10. Brooks-Walter, A.; Briles, D. E.; Hollingshead, S. K. The PspC gene of *Streptococcus pneumoniae* encodes a polymorphic protein, PspC, which elicits cross-reactive antibodies to PspA and provides immunity to pneumococcal bacteremia. *Infect. Immun.* **1999**; *67*; 6533–6542.
11. Ogunniyi, A.D.; Woodrow, M.C.; Poolman, J.T.; Paton. J.C. Protection against *Streptococcus pneumoniae* elicited by immunization with Pneumolysin and CbpA. *Infect. Immun.* **2001**; *69(10)*; 5997–6003. doi: 10.1128/IAI.69.10.5997-6003.2001
12. Möginger, U.; Resemann, A.; Martin, C.E.; Parameswarappa, S.; Govindan, S.; Wamhoff, E.C.; Broecker, F.; Suckau, D.; Pereira, C.L.; Anish, C.; Seeberger, P.H.; Kolarich, D. Cross Reactive Material 197 glycoconjugate vaccines contain privileged conjugation sites. *Sci. Rep.* **2016**; *6*; 20488. doi: 10.1038/srep20488.
13. Rudensky, A.Yu.; Preston-Hurlburt, P.; Hong, S.C.; Barlow, A.. Sequence analysis of peptides bound to MHC class II molecules. *Nature.* **1991**; *353(6345);* 622-7.

14. Hemmer, B.; Kondo, T.; Gran, B.; Pinilla, C.; Cortese, I.; Pascal, J.; Tzou, A.; McFarland, H.F.; Houghten, R.; Martin, R. Minimal peptide length requirements for CD4+ T cell clones—implications for molecular mimicry and T cell survival. *Intl. Immun.* **2000**, *12(3)*,375–383, https://doi.org/10.1093/intimm/12.3.375
15. Ekeruche-Makinde, J.; Miles, J.J.; van den Berg, H.A.; Skowera, A.; Cole, D.K.; Dolton, G.; Schauenbur,g A.J.A.; Tan, M.P.; Pentier, J.M.; Llewellyn-Lacey, S.; Miles, K.M.; Bulek, A.M.; Clement, M.; Williams, T.; Trimby, A.; Bailey, M.; Rizkallah, P.; Rossjohn, J., Peakman, M.; Price, D.A,; Burrows, S.R.; Sewell, A.K.; Wooldridge, L. Peptide length determines the outcome of TCR/peptide-MHCI engagement. *Blood.* **2013**; *121(7);* 1112–1123. doi: 10.1182/blood-2012-06-437202
16. Cunningham, M.W.; McCormack, J.M.; Fenderson, P.G.; Ho, M.K., et al., Human and murine antibodies cross-reactive with streptococcal M protein and myosin recognize the sequence GLN-LYS-SER-LYS-GLN in M protein. *J. Immunol.* **1989**; *143(8);* 2677-83.
17. Kanduc, D. Quantifying the possible cross-reactivity risk of an HPV16 vaccine. *J. Exp. Ther. Oncol.* **2009**; *8(1);* 65-76.
18. Root-Bernstein, R. Autoreactive T-cell receptor (Vbeta/D/Jbeta) sequences in diabetes are homologous to insulin, glucagon, the insulin receptor, and the glucagon receptor. *J. Mol. Recognit.* **2009**;.*22(3);* 177-187.
19. Root-Bernstein, R. Rethinking molecular mimicry in rheumatic heart disease and autoimmune myocarditis: Laminin, collagen IV, CAR, and B1AR as initial targets of disease. *Front. Ped. Rheumatol.* **2014**; *2;* 85. doi: 10.3389/fped.2014.00085.
20. Root-Bernstein, R.; Podufaly, A. Autoreactive T-cell receptor (Vbeta/D/Jbeta) sequences in diabetes recognize insulin, the insulin receptor, and each other, and are targets of insulin antibodies. Open *Autoimmunity J*. **2012**; *4*; 10-22, DOI: 10.2174/1876894601204010010
21. Root-Bernstein, R. How to make a non-antigenic protein (auto) antigenic: Molecular complementarity alters antigen processing and activates adaptive-innate immunity synergy. *Anticancer Agents Med. Chem.* **2015**; *15(10)*; 1242-59. doi: 10.2174/1871520615666150716105057.
22. Takiishi, T.; Fenero, C.I.M.; Câmara, N.O.S. Intestinal barrier and gut microbiota: Shaping our immune responses throughout life. *Tissue Barriers.* **2017**; *5(4)*; e1373208. doi:10.1080/21688370.2017.1373208
23. Damian, R.T. Molecular mimicry in biological adaptation In Nickol, B.B., editor, *Host-Parasite Interfaces*. New York: Academic Press; **1979**; 103–26.
24. De Groot, A.S.; Moise, L.; Liu, R.; Gutierrez, A.H.; Tassone, R.; Bailey-Kellogg, C.; Martin, W. Immune camouflage: Relevance to vaccines and human immunology. *Hum. Vaccin. Immunother.* **2014**; *10*; 3570–3575. doi: 10.4161/hv.36134.
25. Moise, L.; Beseme, S.; Tassone, R.; Liu, R.; Kibria, F.; Terry, F.; Martin, W.; de Groot, A.S. T cell epitope redundancy: Cross-conservation of the TCR face between pathogens and self and its implications for vaccines and autoimmunity. *Expert Rev. Vaccines.* **2016**; *15*; 607–617. doi: 10.1586/14760584.2016.1123098
26. Root-Bernstein, R. Autoimmunity and the microbiome: T-cell receptor mimicry of "self" and microbial antigens mediates self tolerance in holobionts: The concepts of "holoimmunity" (TcR-mediated tolerance for the holobiont) and "holoautoimmunity" (loss of tolerance for the holobiont) are introduced. *Bioessays.* **2016**; *38(11)*; 1068-1083. doi: 10.1002/bies.201600083
27. Root-Bernstein, R. Human immunodeficiency virus proteins mimic human T cell receptors inducing cross-reactive antibodies. *Int. J. Mol. Sci.* **2017**; *18(10)*; 2091. doi: 10.3390/ijms18102091.
28. FDA. Package insert – Prevnar-13. **2017**. https://www.fda.gov/files/vaccines%2C%20blood%20%26%20biologics/published/Package-Insert------Prevnar-13.pdf Accessed 28 June 2020.
29. FDA. Package insert – PNEUMOVAX 23. **1983**. https://www.fda.gov/media/80547/download Accessed 28 June 2020.
30. CDC. About Diphtheria, Tetanus, and Pertussis Vaccines. **2020**. https://www.cdc.gov/vaccines/vpd/dtap-tdap-td/hcp/about-vaccine.html Accessed 28 June 2020.
31. Grifoni, A.; Weiskopf, D.; Ramirez, S.I.; Mateus, J.; Dan, J.M.; Moderbacher, C.R.; Rawlings, S.A.; Sutherland, A.; Premkumar, L.; Jadi, R.S.; Marrama, D.; de Silva. A.M.; Frazier, A.; Carlin, A.F.; Greenbaum, J.A.; Peters, B.; Krammer, F.; Smith, D.M.; Crotty, S.; Sette, A. Targets of T cell responses to SARS-CoV-2 coronavirus in humans with COVID-19 disease and unexposed individuals. *Cell.* **2020**; *181(7);* 1489-1501.e15. doi:10.1016/j.cell.2020.05.015
32. Mateus, J.; Grifoni, A.; Tarke, A.; Sidney, J.; Ramirez, S.I.; Dan, JM.; Burger, Z.C.; Rawlings, S.A.; Smith, D.M.; Phillips, E.; Mallal, S.; Lammers, M.; Rubiro, P.; Quiambao, L.; Sutherland,, A.; Yu E.D.; da Silva

Antunes, R.; Greenbaum, J.; Frazier, A.; Markmann, A.J.; Premkumar, L.; de Silva, A.; Peters, B.; Crotty, S.; Sette, A.; Weiskopf, D. Selective and cross-reactive SARS-CoV-2 T cell epitopes in unexposed humans. *Science.* **2020***: eabd3871*. doi: 10.1126/science.abd3871.

33. Sette, A.; Crotty, S. Pre-existing immunity to SARS-CoV-2: the knowns and unknowns. *Nat. Rev. Immunol.* **2020**; *20(8);* 457-458. doi: 10.1038/s41577-020-0389-z.
34. Merckx, J.; Labrecque, J.A.; Kaufman, J.S. Transmission of SARS-CoV-2 by Children. *Dtsch, Arztebl, Int.* **2020**; *117(33-34);* 553-560. doi: 10.3238/arztebl.2020.0553.
35. Zhang, W.; Cheng, W.; Luo, L.; Ma, Y.; Xu, C.; Qin, P.; Zhang, Z. Secondary transmission of coronavirus disease from presymptomatic persons, China. *Emerg. Infect. Dis.* **2020;** *26(8)*; 1924-1926. doi: 10.3201/eid2608.201142. Epub 2020 May 26. PMID: 32453686 Free PMC article.
36. Kam, K.Q.; Yung, C.F.; Cui, L.; Tzer Pin Lin, R.; Mak, T.M.; Maiwald, M.; Li, J.; Chong, C.Y.; Nadua, K.; Tan, N.W.H.; Thoon, K.C. A well infant with coronavirus disease 2019 with high viral load. *Clin Infect. Dis.* **2020**; *71(15);* 847-849. doi: 10.1093/cid/ciaa201.
37. Qian, G.; Yang, N.; Ma, A.H.Y.; Wang, L.; Li, G.; Chen, X; Chen, X. COVID-19 transmission within a family cluster by presymptomatic carriers in China. *Clin. Infect. Dis.* **2020**; *71(15)*; 861-862. doi: 10.1093/cid/ciaa316.
38. Ye, F.; Xu, S.; Rong, Z.; Xu. R.; Liu, X.; Deng, P.; Liu, H.; Xu, X. Delivery of infection from asymptomatic carriers of COVID-19 in a familial cluster. *Int. J. Infect. Dis.* **2020**; *94*; 133-138. doi: 10.1016/j.ijid.2020.03.042.
39. Härkönen, T.; Puolakkainen, M.; Sarvas, M.; Airaksinen, U.; Hovi, T.; Roivainen, M. Picornavirus proteins share antigenic determinants with heat shock proteins 60/65. *J. Med. Virol***. 2000**; *62(3)*; 383-91. doi: 10.1002/1096-9071(200011)62:3<383::aid-jmv11>3.0.co;2-#.
40. Misko, I.S.; Cross, S.M.; Khanna, R.; Elliott, S.L.; Schmidt, C.; Pye, S.J.; Silins, .SL. Crossreactive recognition of viral, self, and bacterial peptide ligands by human class I-restricted cytotoxic T lymphocyte clonotypes: implications for molecular mimicry in autoimmune disease. *Proc. Natl. Acad. Sci.. U S. A.* **1999**; *96(5)*; 2279-84. doi: 10.1073/pnas.96.5.2279.
41. Trama, A.M.; Moody, M.A.; Alam, S.M.; Jaeger, F.H.; Lockwood, B.; Park,s R.; Lloyd, K.E.; Stolarchuk, C.; Scearce, R.; Foulger, A.; Marshall, D.J.; Whitesides, J.F.; Jeffries, T.L., Jr.; Wiehe, K.; Morris, L.; Lambson, B.; Soderber,g K.; Hwang, K.K.; Tomaras, G.D.; Vandergrift, N.; Jackson, K.J.,L.; Roskin K.M.; Boyd, S.D.; Kepler, T.B.; Liao, H.X.; Haynes, B.F. HIV-1 envelope gp41 antibodies can originate from terminal ileum B cells that share cross-reactivity with commensal bacteria. *Cell Host Microbe.* **2014**; *16(2)*; 215-226. doi: 10.1016/j.chom.2014.07.003.
42. Williams, W.B.; Liao, H.X.; Moody, M.A.; Kepler, T.B.; Alam, S.M.; Gao, F.; Wiehe,, K.; Trama A.M.; Jone,s K.; Zhang, R.; Song, H.; Marshall, D.J.; Whitesides, J.F.; Sawatzki, K.; Hua, A.; Liu, P.; Tay, M.Z.; Seaton, K.E.; Shen, X.; Foulger, A.; Lloyd, K.E.; Parks, R.; Pollara, J.; Ferrari, G.; Yu, J.S.; Vandergrift, N.; Montefiori, D.C.; Sobieszczyk, M.E.; Hammer, S.; Karuna, S.; Gilbert, P.; Grove, D.; Grunenberg, N.; McElrath, M.J.; Mascola, J.R.; Koup, R.A.; Corey, L.; Nabel, G.J.; Morgan, C.; Churchyard, G.; Maenza, J.; Keefer, M.; Graham, B.S.; Baden, L.R.; Tomaras, G.D.; Haynes, B.F. HIV-1 VACCINES. Diversion of HIV-1 vaccine-induced immunity by gp41-microbiota cross-reactive antibodies. *Science.* **2015**; *349(6249):* aab1253. doi: 10.1126/science.aab1253.
43. Ross, T.; Slavik, M.; Bayyari, G.; Skeeles, J. Elimination of mycoplasmal plate agglutination cross-reactions in sera from chickens inoculated with infectious bursal disease viruses. *Avian Dis* . **1990**; *34(3)*; 663-7.
44. Bordenave, G. L'idiotypie comparée des anticorps de lapins différents contre salmonella abortus-equi et contre le virus de la mosaique du tabac. observation d'une réactivité croisée entre certains idiotypes d'anticorps contre ces deux matériels antigéniques [Comparison of idiotypes of rabbit antibodies against Salmonella abortusequi and tobacco mosaic virus. Study of cross reactions between antibody idiotypes against these 2 antigenic materials]. *Eur. J. Immunol.* **1973**; *3(11);* 726-31. doi: 10.1002/eji.1830031114.
45. Root-Bernstein, R. Vaccination markers: Designing unique antigens to be added to vaccines to differentiate between natural infection and vaccination. *Vaccine.* **2005**; *23(17-18)*; 2057-2059. doi:10.1016/j.vaccine.2005.01.008
46. Root-Bernstein R. Positive vaccination markers. *Hum. Vaccin.* **2007**; *3(3)*; 104-105. doi:10.4161/hv.3.3.4138
47. Franklin, R.; Young, A.; Neumann, B.; Fernandez, R.; Joannides, A.; Reyahi, A.; Modis, Y. Homologous protein domains in SARS-CoV-2 and measles, mumps and rubella viruses: preliminary evidence that MMR vaccine might provide protection against COVID-19. *MedRxiv* **2020**. doi: https://doi.org/10.1101/2020.04.10.20053207

48. Gold, J.E. MMR vaccine appears to confer strong protection from COVID-19: Few deaths from SARS-CoV-2 in highly vaccinated populations. *ResearchGate Preprint*, **2020**, DOI: 10.13140/RG.2.2.32128.25607
49. Chumakov, K., Gallo, R. Could an old vaccine be a godsend for new coronavirus? Using the oral polio vaccine could prevent or reduce the spread of COVID-19 to immunized individuals. *USA Today*. **21 April 2020**. https://www.usatoday.com/story/opinion/2020/04/21/oral-polio-vaccine-has-potential-treat-coronavirus-column/5162859002/ Accessed 21 APRIL 2020
50. Netea, M.G.; Giamarellos-Bourboulis, E.J.; Dominguez-Andres, J.; Curtis, N.; van Crevel, R.; van d Veerdonk, F.L.; Bonten, M. Trained immunity: A tool for reducing susceptibility to and the severity of SARS-CoV-2 Infection. *Cell.* **2020***; 181(5)*; 969-977.doi: 10.1016/j.cell.2020.04.042.
51. Riccò, M.; Gualerzi, G.; Ranzieri, S.; Bragazzi, N.L. Stop playing with data: there is no sound evidence that Bacille Calmette-Guérin may avoid SARS-CoV-2 infection (for now). *Acta Biomed.* **2020***; 91(2)*; 207-213. doi: 10.23750/abm.v91i2.9700.
52. Hamiel, U.; Kozer, E.; Youngster, I. SARS-CoV-2 Rates in BCG-vaccinated and unvaccinated young adults. *JAMA.* **2020***; 323(22)*; 2340-2341. doi: 10.1001/jama.2020.8189.
53. Pereira, M.; Paixão, E.; Trajman,, A.; de Souza, R.A.; da Natividade, M.S.; Pescarini, J.M.; Pereira, S.M.; Barreto, F.R.; Ximenes, R.; Dalcomo, M.; Ichihara, M.Y.; Nunes, C.; Barral-Netto, M.; Barreto, M.L. The need for fast-track, high-quality and low-cost studies about the role of the BCG vaccine in the fight against COVID-19. Version 2. *Respir. Res.* **2020***; 21(1)*; 178. doi: 10.1186/s12931-020-01439-4.
54. De Bruyn, J.; Bosmans, R.; Turneer, M.; Weckx, M.; Nyabenda, J.; Van Vooren, J.P.; Falmagne, P.; Wiker, H.G.; Harboe, M. Purification, partial characterization, and identification of a skin-reactive protein antigen of *Mycobacterium bovis* BCG. *Infect. Immun.* **1987***; 55(1);* 245-52.
55. Harboe, M.; Wiker, H.G.; Nagai, S. Protein antigens of mycobacteria studied by quantitative immunologic techniques. *Clin. Infect. Dis.* **1992***; 14(1)*; 313-9. doi: 10.1093/clinids/14.1.313.
56. Romain, F.; Laqueyrerie, A.; Militzer, P.; Pescher, P.; Cavarot, P.; Lagranderie, M.; Auregan, G.; Gheorghiu, M.; Marchal, G. Identification of a *Mycobacterium bovis* BCG 45/47-kilodalton antigen complex, an immunodominant target for antibody response after immunization with living bacteria. *Infect. Immun.* **1993***; 61(2)*; 742–750.
57. Aguilo, N.; Gonzalo-Asensio, J.; Alvarez-Arguedas, S.; Marinova, D.; Gomez, A.B.; Uranga, S.; Spallek, R.; Singh, M.; Audran, R.; Spertini, F.; Martin,, C. Reactogenicity to major tuberculosis antigens absent in BCG is linked to improved protection against Mycobacterium tuberculosis. *Nat. Commun.* **2017***; 8*; 16085. https://doi.org/10.1038/ncomms16085
58. Mustafa, A.S.; Skeiky, Y.A.; AL-Attiha, R.; Alderson, M.R.; Hewinson, R.G.; Vordermeier, H.M. Immunogenicity of *Mycobacterium tuberculosis* antigens in *Mycobacterium bovis* BCG-vaccinated and *M. bovis*-infected cattle. *Infect. Immun***. 2006***; 74(8)*; 4566–4572. doi: 10.1128/IAI.01660-05
59. Morens, D.M.; Taubenberger, J.K.; Fauci, A.S. Predominant role of bacterial pneumonia as a cause of death in pandemic influenza: Implications for pandemic influenza preparedness. *J. Infect. Dis.* **2008***; 198*; 962.
60. Root-Bernstein, R.S.; Podufaly, A.; Aimone, F. Antigenic complementarity between influenza a virus and *Haemophilus influenzae* may drive lethal co-infection such as that seen in 1918-19. *J. Virol. Antivir. Res.* **2013**, *2*; 1 doi: 10.4172/2324-8955.1000104.
61. Cucchiari, D.; Pericàs, J.M.; Riera, J.; Gumucio, R.; Nicolás, D.; Hospital Clínic 4H Team. Pneumococcal superinfection in COVID-19 patients: A series of 5 cases. *Med. Clin. (Barc.).* **2020**; S0025-7753(20)30349-3. doi: 10.1016/j.medcli.2020.05.022.
62. Clancy, C.J.; Nguyen, M.H. COVID-19, superinfections and antimicrobial development: What can we expect? *Clin. Infect. Dis*. **2020**, doi:10.1093/cid/ciaa524.
63. Rawson, T.M.; Moore, L.S.P.; Zhu, N.; Ranganathan, N. Bacterial and fungal co-infection in individuals with coronavirus: A rapid review to support COVID-19 antimicrobial prescribing. *Clin. Infect. Dis.* **2020**; pii: ciaa530. doi: 10.1093/cid/ciaa530.
64. Xia, W.; Shao, J.; Guo, Y.; Peng, X.; Li, Z.; Hu, D. Clinical and CT features in pediatric patients with COVID-19 infection: Different points from adults. *Pediatr. Pulmonol*. **2020;** doi: 10.1002/ppul.24718.
65. Wang, L.; He, W.; Yu, X.; Hu, D.; Bao, M.; Liu, H.; Zhou, J.; Jiang H. Coronavirus disease 2019 in elderly patients: characteristics and prognostic factors based on 4-week follow-up. *J. Infect*. **2020**, pii: S0163-4453(20)30146-8. doi: 10.1016/j.jinf.2020.03.019
66. Zhang, J.J.; Dong, X.; Cao, Y.Y.; Yuan, Y.D. Clinical characteristics of 140 patients infected with SARS-CoV-2 in Wuhan, China. *Allergy*. **2020**, doi: 10.1111/all.14238.

67. Liu, G.S.; Li, H.; Zhao, S.C.; Lu, R.J.; Niu, P.H.; Tan, W.J. Viral and bacterial etiology of acute febrile respiratory syndrome among patients in Qinghai, China. *Biomed. Environ. Sci.* **2019;** *32(6);* 438-445. doi: 10.3967/bes2019.058.
68. Guan, W.; Ni, Z.; Hu, Y.; Liang, W.; Ou, C.; He, J.; et al. Clinical characteristics of coronavirus disease 2019 in China. *N. Engl. J. Med.* **2020**; *382(18)*; 1708–20. https://doi.org/10.1056/NEJMoa2002032;
69. Zhou, F.; Yu, T.; Du, R.; Fan, G,; Liu, Y.; Liu, Z.; et al. Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: A retrospective cohort study. *The Lancet* **2020**; *395*; 1054–62. https://doi.org/10.1016/S0140- 6736(20)30566-3
70. Zhu. X.; Ge, Y.; Wu T, et al. Co-infection with respiratory pathogens among COVID-2019 cases. Virus Res. 2020;285:198005. doi:10.1016/j.virusres.2020.198005
71. Lavoignet, C.E.; Le Borgne, P.; Chabrier, S.; Bidoire, J.; Slimani, H.; Chevrolet-Lavoignet, J.; Lefebvre, F.; Jebri, R.; Sengler, L.; Bilbault, P.; and the CREMS network. White blood cell count and eosinopenia as valuable tools for the diagnosis of bacterial infections in the ED. Eur. *J. Clin. Microbiol. Infect. Dis.* **2019**; *38(8)*; 1523-1532. doi: 10.1007/s10096-019-03583-2.
72. Debray, A.; Nathanson, S.; Moulin, F.; Salomon, J.; Davido, B. Eosinopenia as a marker of diagnosis and prognostic to distinguish bacterial from aseptic meningitis in pediatrics. *Eur. J. Clin. Microbiol. Infect. Dis.* **2019**; *38(10)*; 1821-1827. doi: 10.1007/s10096-019-03614-y.
73. van de Garde, M.D.B.; van Westen, E.; Poelen, M.C.M.; Rots, N.Y.; van Els, C.A.C.M. Prediction and validation of immunogenic domains of pneumococcal proteins recognized by human CD4+ T cells. *Infect. Immun*. **2019**; 87(6):e00098-19. doi: 10.1128/IAI.00098-19
74. Briles, D. E.; Hollingshead, S.; Brooks-Walter, A.; Nabors, G. S.; Ferguson, L.; Schilling, M.; Gravenstein, S.; Braun, P.; King, J.; Swift, A. The potential to use PspA and other pneumococcal proteins to elicit protection against pneumococcal infection. *Vaccine.* **2000**; *18*; 1707–1711
75. Ferreira, D.; Darrieux, M.; Débora, S.A.; Leite, L.; Ferreira, J.; Ho, P.; Miyaji, E.; Oliveira, M. Characterization of protective mucosal and systemic immune responses elicited by pneumococcal surface protein PspA and PspC nasal vaccines against a respiratory pneumococcal challenge in mice. *Clin. Vacc. Immuno.l CVI.* **2009**; *16*; 636-45. 10.1128/CVI.00395-08.
76. Schachern, P.A.; Tsuprun, V.; Ferrieri, P.; Briles, D.E.; Goetz, S.; Cureoglu, S.; Paparella, M.M.; Juhn, S. Pneumococcal PspA and PspC proteins: Potential vaccine candidates for experimental otitis media. *Int. J. Pediatr. Otorhinolaryngol.* **2014**; *78(9);* 1517–1521. doi: 10.1016/j.ijporl.2014.06.024
77. Lagousi, T.; Basdeki, P.; Routsias, J.; Spoulou, V. Novel protein-based pneumococcal vaccines: Assessing the use of distinct protein fragments instead of full-length proteins as vaccine antigens. *Vaccines (Basel)* **2019**; *7(1)*; 9. doi: 10.3390/vaccines7010009
78. Masomian, M.; Ahmad, Z.; Gew, L.T.; Poh, C.L. Development of next generation *Streptococcus pneumoniae* vaccines conferring broad protection. *Vaccines (Basel).* **2020**; *8(1)*; 132. doi: 10.3390/vaccines8010132.
79. Fedson, D.S.; Nicolas-Spony, L.; Klemets, P.; van der Linden, M.; Marques, A.; Salleras, L.; Samson, S.I. Pneumococcal polysaccharide vaccination for adults: New perspectives for Europe. *Expert Rev. Vaccines***.** **2011**; *10(8)*; 1143-67. doi: 10.1586/erv.11.99.
80. Mahamat, A.; Daurès, J.P.; de Wzieres, B. Additive preventive effect of influenza and pneumococcal vaccines in the elderly: Results of a large cohort study. *Hum. Vaccin. Immunother.* **2013**; *9(1)*; 128-35. doi: 10.4161/hv.22550.
81. Choi, Y.H.; Miller, E. Potential impact of Covid-19 response measures on invasive pneumococcal disease in England and Wales. *MedRxiv*. **2020;** doi: https://doi.org/10.1101/2020.06.01.20119057
82. National Institute for Communicable Diseases (South Africa). Pneumococcal conjugate vaccine use in the light of the COVID-19 pandemic. **2020**. https://www.nicd.ac.za/diseases-a-z-index/covid-19/advice-for-the-public/pneumococcal-conjugate-vaccine-use-in-the-light-of-the-covid-19-pandemic/ Accessed 10 July 2020
83. Statens Serum Institut. Selected risk groups are offered free pneumococcal vaccination. **2020**. https://www.sst.dk/da/Nyheder/2020/Udvalgte-risikogrupper-faar-tilbud-om-gratis-vaccination-mod-pneumokokker Accessed 7 April 2020
84. New Zealand. Policy on pneumococcal vaccination and COVID. **2020**. https://www.nzdoctor.co.nz/article/news/pneumococcal-vax-potential-option-improve-outcomes-compromised-patients-face-covid-19 Accessed 19 Aug 2020