RESEARCH ARTICLE

Proteomic variation and diversity in clinical *Streptococcus pneumoniae* isolates from invasive and non-invasive sites

Mustapha Bittaye*, Phil Cash, Ken Forbes

School of Medicine, Medical Sciences & Nutrition, University of Aberdeen, Aberdeen, United Kingdom

* r02mb14@abdn.ac.uk, mustaphabittaye@yahoo.com

Abstract

Streptococcus pneumoniae is responsible for a variety of invasive and non-invasive human infections. There are over 90 serotypes of *S. pneumoniae* differing in their ability to adapt to the different niches within the host. Two-dimensional gel electrophoresis was used to discriminate clinical *S. pneumoniae* isolates recovered from either blood cultures (invasive site isolates) or other sites, including sputum, tracheal aspirate, ear, eye and skin swabs (non-invasive site isolates). Global protein expression profiles for five invasive site and six non-invasive site isolates representing five different serotypes (serotypes 4, 6, 9, 14 and 23) were obtained for each isolate and combined into a single data set using Progenesis Same-Spots™ software. One-hundred and eighty six protein spots (39% of the protein spots in the dataset) differed significantly (ANOVA, $p < 0.05$) in abundance between the invasive site (101 upregulated protein spots) and non-invasive site (85 upregulated protein spots) isolates. Correlations between the bacterial proteomes and their sites of isolation were determined by Principal Component Analysis (PCA) using the significantly different protein spots. Out of the 186 variable protein spots, 105 exhibited a serotype-associated pattern of variability. The expression of the remaining 81 protein spots was concluded to be uniquely linked to the site of bacterial isolation. Mass spectrometry was used to identify selected protein spots that showed either constant or differential abundance levels. The identified proteins had a diverse range of functions including, capsule biogenesis, DNA repair, protein deglycation, translation, stress response and virulence as well as amino acid, carbohydrate, lipid and nucleotide metabolism. These findings provide insight on the proteins that contribute towards the adaptation of the bacteria to different sites within the host.

Introduction

Streptococcus pneumoniae (also known as the pneumococcus) is a significant human pathogenic bacterium that remains a leading cause of morbidity and mortality worldwide. The pneumococcus is the principal bacterial cause of community-acquired pneumonia and acute otitis media as well as being associated with several other non-invasive infections including sinusitis and conjunctivitis [1]. The human nasopharynx is the natural ecological niche of the

bacterium, where it exists asymptomatically in a carrier state. Migration of the bacterium across the respiratory epithelium into the blood system and eventually the cerebro-spinal fluid (CSF) leads to the development of severe life-threatening, invasive pneumococcal diseases (IPDs) such as sepsis and meningitis [2]. Worldwide, pneumococcal infections are responsible for the deaths of 1.6 million people annually [3]. Young children, particularly in low-income countries, the elderly, the economically deprived as well as those with HIV/AIDS bear a disproportionate share of the disease burden. The highest mortality due to pneumococcal infection occurs in patients who develop pneumococcal sepsis and meningitis [4]. Immunization and antimicrobial therapy are the cornerstones for the prevention and management of pneumococcal infections. However, the lack of a universal vaccine against the many circulating pneumococcal serotypes, as well as the widespread prevalence of antibiotic resistance, present specific challenges for disease control and management [5]. These obstacles are compounded further by a lack of rapid, sensitive, and specific diagnostic tests for IPD.

S. pneumoniae is a genetically diverse species that exists in many different capsular and sequence types. The ability of the bacterium to cause either invasive disease or localised infections varies markedly between individual isolates. This pathogenic variability is influenced by both the bacterial capsular serotype and genetic background [6,7]. Moreover, the bacterial site of isolation has also been shown to influence the virulence behaviour displayed by the bacterium [8,9]. Comparative studies on the available genome sequences of *S. pneumoniae* isolates identified core genes that are common to all isolates as well as accessory regions comprising variable genes that correlate with virulence phenotype. Both the core genome and regions of diversity encode virulence determinants whose expression varies according to the bacterial serotype and sites of infection [6,7]. This dynamic nature of the pneumococcal expression of virulence factors in different host niches has been investigated in recent years using comparative transcriptomics. Both Ogunniyi *et al.* [8] and Orihuela *et al.* [9] reported niche-specific differences in the transcriptional profiles of pneumococci recovered from the nasopharynx, lung and blood following intranasal infection of mice. However, these transcript profiling studies have been carried out on commonly used laboratory-adapted *S. pneumoniae* strains, in mice, and cannot consider the expression levels of the related proteins and their post-translational modifications.

The comparison of pathogenic microorganisms at the proteomic level provides complementary data to those obtained using nucleic acid based methods. The availability of complete genome sequences for many virulent and avirulent strains of *S. pneumoniae* provides the basis for in depth investigations of the pneumococcal proteome, particularly for studies on pneumococcal pathogenesis [5]. To date, several proteomic studies on *S. pneumoniae* have been reported that focus primarily on identifying potential virulence determinants expressed by laboratory type strains of pneumococcus grown *in vitro* under conditions that mimic specific host environments [10–13]. However, less attention has been paid towards comparative proteomic analyses of clinical *S. pneumoniae* isolates recovered from different host niches in order to identify markers correlating with specific-disease phenotypes as well as obtaining information on pathogenic differences arising from strain variation.

In the following study, a global comparative proteomics approach using two-dimensional gel electrophoresis (2-DE) based protein profiling was employed to differentiate invasive and non-invasive site clinical *S. pneumoniae* isolates. The combination of 2-DE coupled with mass spectrometry (MS) for protein separation and identification respectively, have been extensively employed to study the pathogenesis for a number of important human bacterial pathogens [14–17]. In this study, we compared invasive site pneumococcal isolates recovered from blood cultures with non-invasive site isolates recovered from various mucosal surfaces (sputum, tracheal aspirate, ear, eye and skin swabs), in order to investigate the proteome variation and

diversity associated with the pneumococcal isolates from the two infection sites. Despite the extensive heterogeneity found for the pneumococcal isolates, it was possible to discriminate the invasive and non-invasive site isolates on the basis of their 2-dimensional (2-D) protein profiles. Moreover, the proteins that displayed relatively constant abundance levels across the pneumococcal isolates could in the long run be investigated for inclusion in the pipeline of potential vaccine markers. Thus, the pan proteomic analysis of multiple pneumococcal isolates recovered from invasive and non-invasive sites of infection is a valuable approach for identifying proteins involved in pneumococcal colonisation and virulence.

Results

Growth of *S. pneumoniae* in BHI-FCS broth

To optimise the bacterial culture conditions for the analysis of the pneumococcal cellular proteins, representative isolates were grown in BHI-FCS broth and assayed for bacterial cell density and viability as described in the Materials and methods section. The growth curves for two invasive site isolates (B03 and B162) and two non-invasive site isolates (E40 and E65) are shown in Fig 1. Under the experimental conditions used, the four pneumococcal isolates showed similar growth characteristics with the lag-phase for each isolate lasting at least 2 h (Fig 1A), which was followed by an increase in the bacterial cell density. All of the isolates reached their maximum cell density at 8 h post-inoculation followed by a decline in cell density thereafter. The cell viability followed a similar pattern as cell density, although no clear lag phase was identifiable (Fig 1B). Based on these data, the bacterial cells were routinely collected once the inoculated cultures had reached an OD650 of 1.0 to ensure that sufficient biomass (approximately 3×10^7 CFU/mL) was obtained for subsequent proteomic analysis.

Comparison of the pneumococcal 2-D protein profiles

To identify those pneumococcal proteins that differed significantly in abundance levels between the invasive and non-invasive site isolates, the isolates were classified into two groups according to their site of isolation ("invasive site" or "non-invasive site") (Table 1). The cellular proteins for each of the pneumococcal isolates shown in Table 1 were prepared and analysed for 3 biological replicates; representative 2-D protein profiles for each isolate are presented in S1 Fig. Following a manual review of the spot detection data, a total of 483 protein spots was detected for the 33 matched 2-D protein profiles comprising triplicate 2-D gels for each of the 11 clinical pneumococcal isolates. The majority of the detected protein spots had isoelectric points (*pI*) between 4.5 and 6 and molecular weights (*Mr*) range between 20 kDa to 55 kDa. Principal Component Analysis (PCA) was used to determine the degree of variation between the pneumococcal isolates and to identify those proteins that discriminated the two groups.

PCA failed to differentiate the invasive and non-invasive site pneumococcal isolates when all of the 483 detected protein spots were used for the analysis (data not shown). Based on the study of Smith *et al.* [15] PCA using those protein spots (186 protein spots) which differed significantly (ANOVA, p<0.05) in abundance between the two groups of pneumococcal isolates clearly discriminated the two groups of isolates (Fig 2A). The PCA showed that the expression profiles of 101 and 85 protein spots defined the invasive and non-invasive *S. pneumoniae* isolate groups respectively. These discriminatory protein spots (101 and 85 in the invasive and non-invasive groups respectively) showed high abundance levels in the respective groups. Based on the clustering of the 2-D protein profiles, those pneumococcal isolates within the invasive site group appeared to be more heterogeneous at the proteomic level than those isolates within the non-invasive site group (Fig 2A).
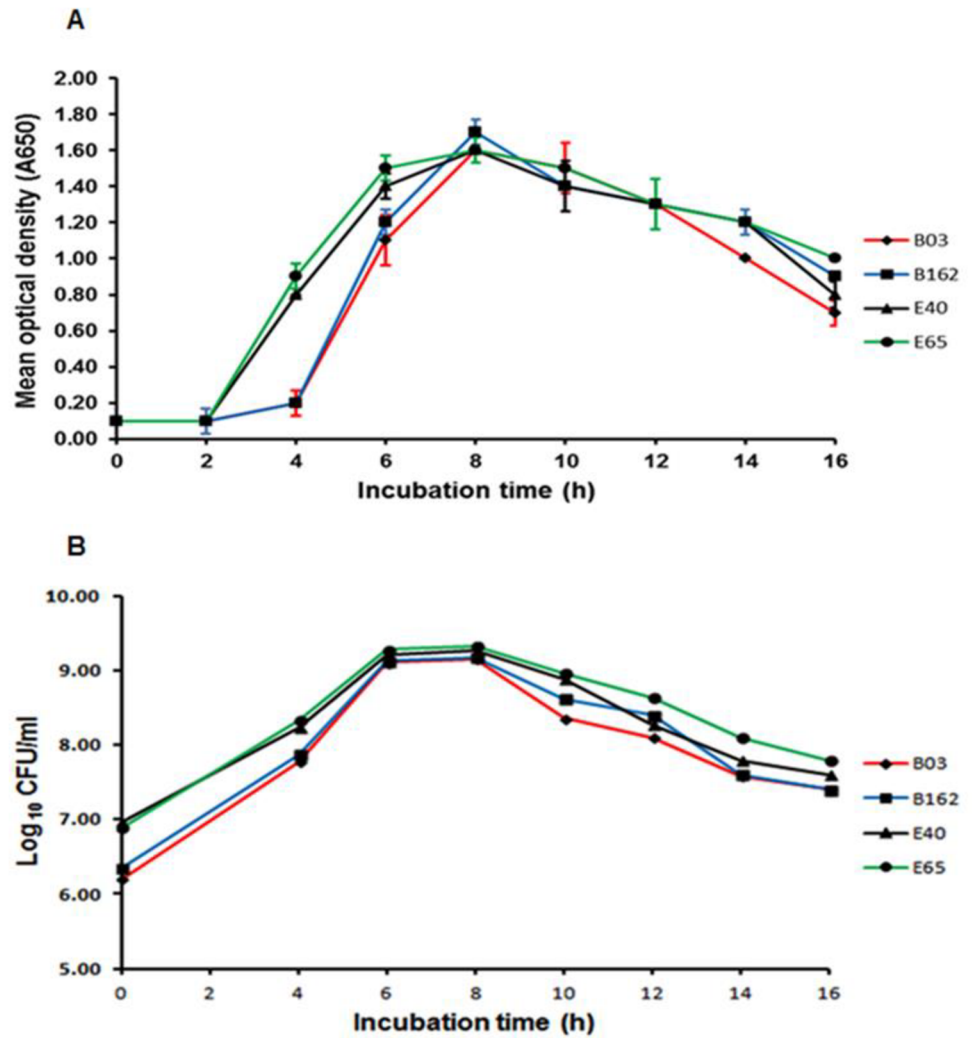
Fig 1. Characteristic growth and viability curves of representative *S. pneumoniae* isolates in BHI-FCS broth. (A) Measurement of the bacterial cell density **(B)** Measurement of bacterial viability.

https://doi.org/10.1371/journal.pone.0179075.g001

Table 1. Clinical *S. pneumoniae* isolates used in the study.

Isolate Category	Isolate ID	Serotype	Site of Isolation
Invasive	B03	14	Blood culture
	B40	4	Blood culture
	B154	6	Blood culture
	B162	9	Blood culture
	B269	4	Blood culture
Non-invasive	E22	23	Eye swab
	E35	14	Sputum
	E40	14	Skin swab
	E47	23	Sputum
	E65	9	Ear swab
	E107	6	Tracheal aspirate

https://doi.org/10.1371/journal.pone.0179075.t001
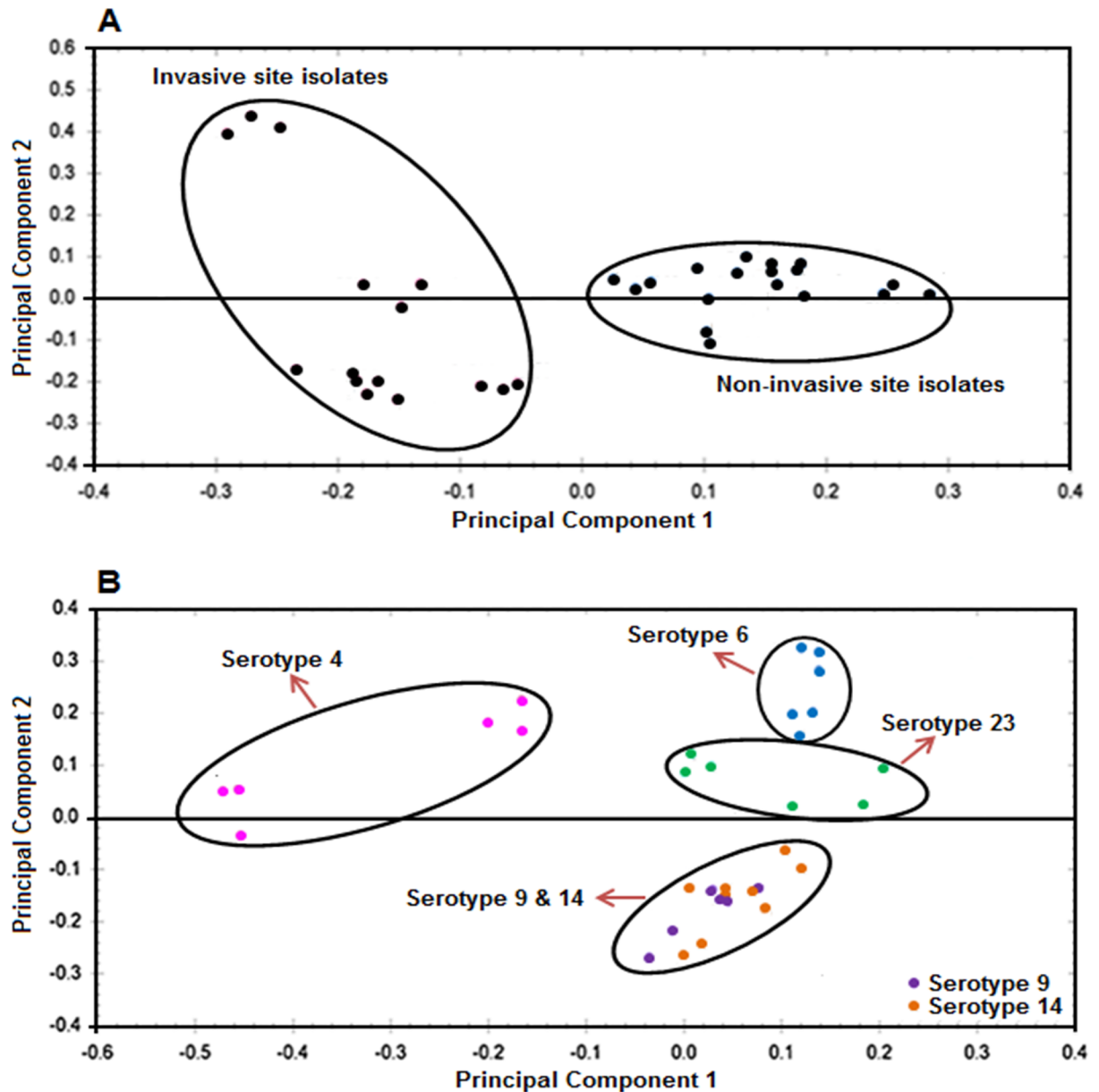
Fig 2. Principal component analysis of the 2-D gel data. The PCA graphs displayed the clustering of the bacterial isolates grouped according to their **(A)** sites of isolation and **(B)** capsular serotype. The PCA was plotted using those protein spots that differed significantly in abundance levels (ANOVA, $p < 0.05$) between isolates in the classified groups.

https://doi.org/10.1371/journal.pone.0179075.g002

The eleven pneumococcal isolates analysed in this study represented five different serotypes (Table 1). Since capsular serotype plays a key role in pneumococcal pathogenesis [18], it was possible that the capsular serotypes might contribute towards the discrete groups identified above by PCA (Fig 2A). To test this, the 2-D gel images were grouped according to the pneumococcal serotypes, producing five groups. Two hundred and sixty six protein spots were identified that showed significant differences (ANOVA, $p < 0.05$) in abundance between the five groups and PCA discriminated the isolates on the basis of serotype (Fig 2B). The pneumococcal isolates of serotypes 4, 6 and 23 formed separate, unique clusters, whereas serotypes 9

Functional category	Spot #	Gene Locus	Protein description	Gene name	Expected pI/Mr	Observed pI/Mr	Peptides matched	Peptide sequences	Sequence coverage (%)	Mascot score	ANOVA (p)	Highest mean	Fold change#	Serotype Variability	Subcellular localization
Capsule biogenesis	1319	SP_0349	Tyrosine-protein kinase CpsD	cpsD	8.96/24887	6.49/26980	3	3	16	195	<0.05	Invasive	1.20	NO	Membrane
Carbohydrate metabolism	1054	SP_0027	Ribose-phosphate pyrophosphokinase 1	prs1	5.68/35451	6.06/37640	15	10	35	799	<0.001	Non-invasive	0.67	YES	Cytoplasmic
	1639	SP_1192	Galactose-6-phosphate isomerase subunit LacB*	lacB	5.47/18958	5.82/17306	8	6	52	380	<0.01	Non-invasive	0.70	YES	Cytoplasmic
	1752	SP_1193	Galactose-6-phosphate isomerase subunit LacA	lacA	5.35/15226	5.63/14383	27	4	28	360	<0.001	Non-invasive	0.59	YES	Cytoplasmic
	2161	SP_0896	ATP-dependent 6-phosphofructokinase	pfkA	5.51/35174	5.80/35609	24	14	52	967	<0.05	Invasive	1.30	YES	Cytoplasmic
	2182	SP_1192	Galactose-6-phosphate isomerase subunit LacB*	lacB	5.47/18958	5.59/18636	5	5	33	348	<0.001	Non-invasive	0.42	YES	Cytoplasmic
	2198	SP_1655	Phosphoglyceromutase	gpmA	5.14/26051	5.29/27484	16	8	49	577	<0.001	Invasive	2.20	YES	Cytoplasmic
Lipid metabolism	1118	SP_0419	Enoyl-acyl Carrier Protein Reductase	fabk	5.31/34177	5.29/36273	12	9	43	549	<0.01	Invasive	1.30	NO	Cytoplasmic
Nucleotide metabolism	786	SP_2228	Inosine-5'-monophosphate dehydrogenase	guaB	5.18/52445	5.33/52514	16	12	34	896	<0.05	Invasive	0.77	NO	Cytoplasmic
	903	SP_0019	Adenylosuccinate synthase	purA	5.48/47570	5.76/46997	20	15	45	1063	<0.001	Non-invasive	0.67	NO	Cytoplasmic
	1188	SP_1458	Thioredoxin reductase	trxB	4.76/33147	4.56/33686	5	5	18	353	<0.001	Invasive	1.60	NO	Cytoplasmic
	1259	SP_0944	Uridylate kinase	pyrH	5.36/26703	5.37/26980	17	9	53	680	<0.01	Invasive	1.20	NO	Cytoplasmic
	1320	SP_0231	Adenylate kinase	adk	4.96/23706	5.03/24145	4	4	28	242	<0.001	Invasive	1.50	NO	Cytoplasmic
	2167	SP_0012	Hypoxanthine-guanine phosphoribosyltransferase	hpt	5.26/20212	5.46/20442	6	5	29	282	<0.05	Invasive	1.10	YES	Cytoplasmic
	2186	SP_1445	GMP synthase	guaA	4.93/57472	5.03/57603	6	5	17	447	<0.001	Invasive	0.32	NO	Cytoplasmic
Other metabolism	890	SP_0784	Glutathione reductase*	gor	5.82/48790	6.32/48769	6	4	12	341	<0.001	Invasive	3.40	NO	Cytoplasmic
	894	SP_0784	Glutathione reductase*	gor	5.82/48790	6.49/47875	8	5	22	413	<0.001	Non-invasive	0.48	NO	Cytoplasmic
	944	SP_0715	L-lactate oxidase*	lctO-2	5.77/41518	5.93/44460	20	10	36	737	<0.001	Non-invasive	0.50	YES	Cytoplasmic
	2159	SP_0715	L-lactate oxidase*	lctO-2	5.77/41518	6.15/43645	20	9	35	744	<0.001	Invasive	4.00	NO	Cytoplasmic
Protein deglycation	1480	SP_0804	Putative 4-methyl-5(B-hydroxyethyl)-thiazole monophosphate biosynthesis protein		4.64/20164	4.73/20067	5	4	42	260	<0.001	Invasive	1.80	YES	Cytoplasmic
Protein/DNA repair	751	SP_0341	Hypothetical protein		5.51/55065	5.82/55511	17	14	40	910	<0.05	Invasive	0.83	NO	Cytoplasmic
	2170	SP_1540	Single-stranded DNA-binding protein*	ssb	5.41/17351	6.23/19338	14	8	53	653	<0.001	Invasive	1.80	NO	Cytoplasmic
	2175	SP_1540	Single-stranded DNA-binding protein*	ssb	5.41/17351	5.72/18983	12	6	54	481	<0.00	Non-invasive	0.71	NO	Cytoplasmic
Translation	716	SP_2078	Arginine--tRNA ligase	argS	5.05/63472	5.16/63186	19	17	36	1182	<0.01	Invasive	1.50	YES	Cytoplasmic
	1323	SP_0435	Elongation factor P	efp	4.86/20600	4.82/24145	4	4	29	306	<0.001	Invasive	1.30	YES	Cytoplasmic
	1546	SP_2206	Ribosomal subunit interface protein	yfiA	5.08/21099	5.16/20824	10	8	43	561	<0.001	Non-invasive	0.67	YES	Cytoplasmic
	1654	SP_1355	50S ribosomal protein L10	rplJ	5.08/17479	5.07/17307	28	10	65	806	0.973	Conserved	1.00	Conserved	Cytoplasmic
	1795	SP_1541	30S ribosomal protein S6	rpsF	5.15/11153	4.78/11954	4	3	44	227	<0.05	Invasive	1.20	NO	Cytoplasmic
	2189	SP_0945	Ribosome-recycling factor	frr	6.02/20657	5.97/20824	2	2	10	106	<0.001	Non-invasive	0.59	NO	Cytoplasmic
Virulence/Stress	673	SP_0517	Chaperone protein DnaK	dnaK	4.63/64842	4.47/66792	33	17	31	1163	0.137	Conserved	1.20	Conserved	Cytoplasmic
	684	SP_0730	Pyruvate oxidase	spxB	5.06/65266	5.16/65568	30	17	39	1321	0.303	Conserved	1.20	Conserved	Membrane
	1635	SP_2012	Glyceraldehyde-3-phosphate dehydrogenase	gap	5.29/35856	6.45/18636	6	3	8	193	<0.01	Invasive	1.50	NO	Cytoplasmic
	1675	SP_2012	Glyceraldehyde-3-phosphate dehydrogenase	gap	5.29/35856	6.45/17307	7	5	16	285	<0.05	Invasive	1.20	NO	Cytoplasmic
	1681	SP_1128	Enolase	eno	4.69/47103	4.82/15489	10	7	26	602	0.976	Conserved	1.00	Conserved	Cytoplasmic
	2187	SP_1999	Catabolite control protein A	ccpA	5.47/37073	5.82/42060	18	12	41	862	<0.05	Non-invasive	0.83	YES	Cytoplasmic
Unknown	1197	SP_1478	Oxidoreductase, aldo/keto reductase family		5.10/31403	4.90/32463	9	6	34	395	<0.01	Non-invasive	0.83	YES	Cytoplasmic
	1241	SP_1536	Conserved hypothetical protein		5.92/28367	6.10/28519	6	5	23	377	<0.05	Invasive	1.20	YES	Cytoplasmic
	1420	SP_0622	Nitroreductase family protein		6.34/22687	6.36/20824	5	5	33	298	<0.01	Non-invasive	0.67	YES	Cytoplasmic
	1701	SP_1465	Hypothetical protein		6.06/16904	6.45/16678	1	1	12	73	<0.001	Invasive	1.90	NO	Extracellular
	2160	SP_0627	Conserved hypothetical protein		5.25/38041	6.17/42845	11	10	37	763	<0.001	Invasive	3.80	NO	Cytoplasmic

The functional categorisation and subcellular localisation prediction of the proteins were carried out as described in the Materials and Methods Section. The gene loci and gene names of the proteins are based on the *S. pneumoniae* TIGR4 genome annotation and the UniProt database respectively.

*Protein isoforms showing differential expression levels between the invasive and non-invasive site isolates.

#Fold change calculated using the invasive group as a reference.

Fig 3. *S. pneumoniae* proteins showing constant and differential expression between the invasive and non-invasive site isolates.

and 14 formed a fourth cluster within the analysis (Fig 2B). The pneumococcal isolates collected from the two general host niches indicated in Table 1 were distributed across each of the four clusters identified in Fig 2B.

Within the complete dataset of 483 protein spots, a greater proportion of the protein spots differed in abundance among the serotypes (266, 55%) than between the sites of isolation (186, 38.5%). Eighty-one of the proteins that exhibited significant differences between the sites of isolation were not significantly different among the serotypes. These 81 proteins spots, which uniquely discriminated the pneumococcal isolates by site of isolation (i.e. that conferred no serotype variability), were of particular interest since they might play a role in the colonisation and invasion of specific niches within the host during infection. Nineteen of these protein spots were identified by LC-MS (Fig 3); 12 of the protein spots showed increased abundances in the invasive site isolates and 7 protein spots showed the highest abundance in the non-invasive site isolates (Fig 3).

Identification of *S. pneumoniae* proteins

Forty selected protein spots that were either constant or variable in abundance between the invasive and non-invasive site isolates were successfully identified by liquid chromatography-mass spectrometry (LC-MS) against known *S. pneumoniae* proteins, based on significant MASCOT scores. The identified protein spots represented 32 unique pneumococcal proteins (Figs 3 and 4). Thirty-six of the identified protein spots were significantly different (ANOVA,
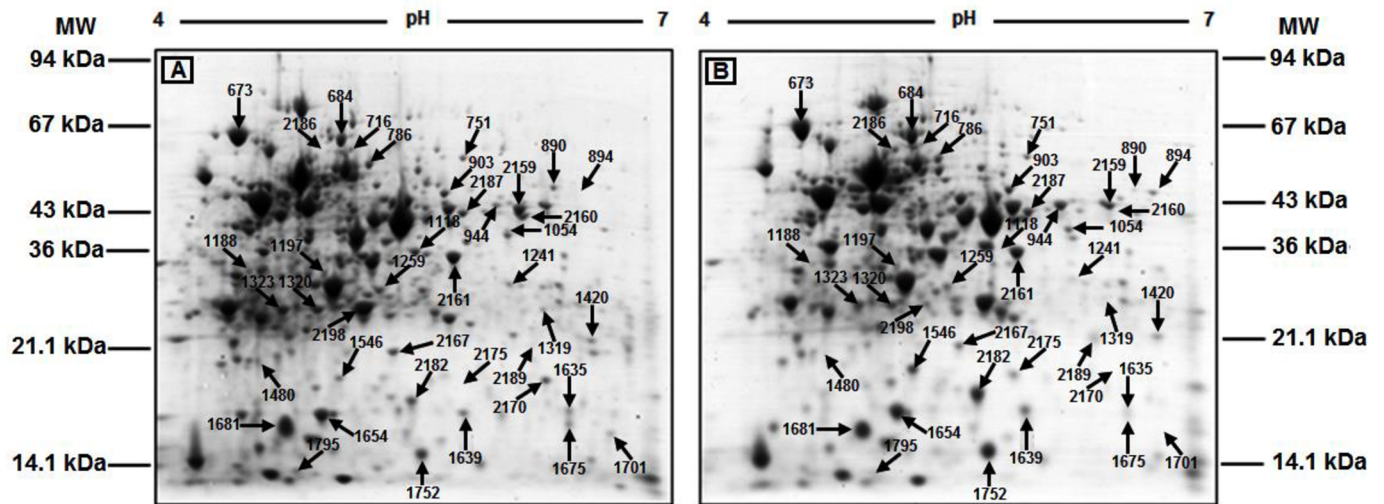
Fig 4. Annotated 2-DE protein profiles of an invasive site and non-invasive site pneumococcal isolate. Displayed are representative 2-DE protein profiles of the pneumococcal total cellular proteins extracted from an **(A)** invasive site isolate (B03) and a **(B)** non-invasive site isolate (E35) showing the pH range and the location of the molecular weight marker proteins. Protein spots, identified by LC-MS, are indicated using the Progenesis SameSpots™ assigned spot numbers.

p<0.05) in abundance between isolates for the two sites whereas 4 proteins (DnaK, SpxB, RplJ and Eno) showed no significant differences in abundance. Fold differences between the highest and lowest mean normalised volumes for the differentially expressed protein spots identified using the invasive group as a reference ranged between 0.25 to 4.0 (Fig 3). The normalised volumes of the protein spots for the identified proteins across the full series of 2-D gels are presented in the form of a heat map diagram in Fig 5. Twenty of the differentially expressed proteins identified were predominantly abundant in the invasive site isolates and 16 proteins showed the highest abundance in isolates from the non-invasive sites (Fig 3).

The migration of proteins in 2-DE gels can be altered by protein post-translational modifications (such as phosphorylation, glycosylation, acetylation and proteolysis) as well as degradation products and covalent reaction intermediates. These processes in addition to alternative splicing of mRNA and single-point mutations could also result in the generation of protein isoforms, in which protein products of the same gene have different observed M_r and pI. Four of the differentially expressed identified proteins were each present as two protein spots exhibiting different migrations consistent with them being isoforms of the same protein, which included Gor (Spots 890 and 894), LacB (spot 1639, 2182), LctO-2 (spots 944, 2159) and Ssb (spots 2170 and 2175) (Fig 3). Two protein spots (spots 1635 and 1675) with increased abundances in the invasive group were identified as Gap. However, the proteins observed Mr and pI values were significantly different from the corresponding values for Gap (Fig 3). Inspection of the identified peptides for Gap (S1 Table) indicates that they are derived from the N-terminus of the protein. This would be consistent with the protein spots identified as Gap being proteolytic degradation products of the intact protein. However, this was not the case for the protein spot identified as Eno, since the identified peptides were distributed throughout the intact Eno protein sequence (Fig 3, S1 Table). Consequently, the reason for the migration of the protein at an observed Mr of 15 kDa rather than the predicted 47 kDa is unknown.

The proteins were assigned to a number of functional groups, which included capsule biogenesis, DNA repair, protein deglycation, translation, stress response, virulence as well as

NCBI Gene Locus	Protein description	Uniprot Gene name	Differential Expression
SP_0012	Hypoxanthine-guanine phosphoribosyltransferase	hpt	Up-regulated in invasive isolates
SP_0231	Adenylate kinase	adk	
SP_0349	Tyrosine-protein kinase CpsD	cpsD	
SP_0419	Enoyl-acyl Carrier Protein Reductase	fabk	
SP_0435	Elongation factor P	efp	
SP_0627	Conserved hypothetical protein		
SP_0715	L-lactate oxidase*	lctO-2	
SP_0784	Glutathione reductase*	gor	
SP_0804	Putative 4-methyl-5(B-hydroxyethyl)-thiazole monophosphate biosynthesis protein		
SP_0896	ATP-dependent 6-phosphofructokinase	pfkA	
SP_0944	Uridylate kinase	pyrH	
SP_1458	Thioredoxin reductase	trxB	
SP_1465	Hypothetical protein		
SP_1536	Conserved hypothetical protein		
SP_1540	Single-stranded DNA-binding protein*	ssb	
SP_1541	30S ribosomal protein S6	rpsF	
SP_1655	2,3-bisphosphoglycerate-dependent phosphoglycerate mutase	gpmA	
SP_2012	Glyceraldehyde-3-phosphate dehydrogenase	gap	
SP_2012	Glyceraldehyde-3-phosphate dehydrogenase	gap	
SP_2078	Arginine--tRNA ligase	argS	
SP_0019	Adenylosuccinate synthase	purA	Up-regulated in non-invasive isolates
SP_0027	Ribose-phosphate pyrophosphokinase 1	prs1	
SP_0341	Hypothetical protein		
SP_0622	Nitroreductase family protein		
SP_0715	L-lactate oxidase*	lctO-2	
SP_0784	Glutathione reductase*	gor	
SP_0945	Ribosome-recycling factor	frr	
SP_1192	Galactose-6-phosphate isomerase subunit LacB*	lacB	
SP_1192	Galactose-6-phosphate isomerase subunit LacB*	lacB	
SP_1193	Galactose-6-phosphate isomerase subunit LacA	lacA	
SP_1445	GMP synthase	guaA	
SP_1478	Oxidoreductase, aldo/keto reductase family		
SP_1540	Single-stranded DNA-binding protein*	ssb	
SP_1999	Catabolite control protein A	ccpA	
SP_2206	Ribosomal subunit interface protein	yfiA	
SP_2228	Inosine-5'-monophosphate dehydrogenase	guaB	
SP_0517	Chaperone protein DnaK	dnaK	Conserved
SP_0730	Pyruvate oxidase	spxB	
SP_1128	Enolase	eno	
SP_1355	50S ribosomal protein L10	rplJ	

*Protein isoforms showing differential expression levels between the invasive and non-invasive site isolates

Fig 5. A heat map representation of the 40 protein spots identified by LC-MS. The identified protein spots showing conserved and differential abundance levels between the invasive and non-invasive site isolates were plotted on the heat map using the normalised spot volumes of the replicate gels of each isolate in the invasive and non-invasive group. The gene locus numbers are based on the publicly available *S. pneumoniae* TIGR4 genome annotation.

https://doi.org/10.1371/journal.pone.0179075.g005

amino acid, carbohydrate, lipid, and nucleotide metabolism. Those proteins with unknown functions including both hypothetical and conserved hypothetical proteins represented 13% of the identified differentially expressed proteins. Proteins predicted by PSORTb to be localised in the pneumococcal cell envelope (i.e. membrane and extracellular proteins) represented 8% of the identified differentially expressed proteins and all had significantly (ANOVA, $p < 0.05$) higher abundances in the invasive site isolates compared to the non-invasive site isolates (Fig 3).

Discussion

Proteomics is used extensively to characterize and compare the proteomes of pathogenic bacteria in order to understand infection mechanisms. In the present study, the comparison of clinical pneumococcal isolates by 2-DE protein expression profiling, has identified proteins associated with pneumococcal strain variability and site of infection. The selection of the bacterial isolates for analysis was primarily based on the bacterial sites of isolation, however, the serotypes (serotypes 4, 6, 9, 14 and 23) of the selected isolates are also clinically significant apropos of their disease potential, widespread geographical prevalence and antibiotic resistance [19,20]. The growth characteristics of the bacteria in BHI-FCS broth was studied to ensure standard *in vitro* growth conditions were used to avoid previously documented growth

effects on proteome [13,21] and to ensure sufficient material were obtained for the proteomic analysis.

This report demonstrates that clinical *S. pneumoniae* isolates recovered from different sites within the host show characteristic cellular proteomes. Similar observations have been reported from proteomic analyses of *E. coli* and *H. pylori* [15–17]. Proteomic discrimination of invasive and non-invasive site isolates was due to the differential expression of 186 protein spots: 101 invasive site and 85 non-invasive site protein spots. Eighty one of these protein spots that were unique in discriminating the isolates by site of isolation provide an opportunity for determining the key activities of the bacteria that define the bacterial adaptation in blood and non-invasive sites of infection. Moreover, the clustering of the 2-D protein profiles revealed a greater heterogeneity among isolates in the invasive group compared to their non-invasive counterparts. This variability could be attributed to the differential adaptation of the bacteria to the challenging environment of the blood probably due to differences in the bacterial serotypes. Growing evidence suggests that pneumococci with different serotypes differ in their invasiveness [22,23]. Thus, the proteomic data presented in the current study could serve as a basis for defining the proteins that contribute to this variability.

Analysis of the bacterial protein profiles by PCA, according to the bacterial serotypes revealed the relatedness of the bacterial isolates as well as those proteins that displayed a serotype-specific pattern of expression. These proteomic data indicated a clustering of the isolates belonging to serotypes 9 and 14. This observation agrees with previous studies that used multilocus sequence typing to demonstrate that some pneumococcal serotype 14 isolates are variants of the 9V clone, especially on the basis of their antibiotic resistance profiles [24–26]. This suggests that the discrimination of the pneumococcal isolates by PCA may also reflect other phenotypic features of the bacteria such as antibiotic resistance.

Moreover, a higher number of proteins differed among the serotypes analysed compared to the isolation sites, indicating the high degree of diversity that exist among pneumococcal isolates [27,28]. Out of the 186 differentially expressed protein spots that discriminated the invasive and non-invasive site isolates, 105 proteins also differed among the pneumococcal isolates on the basis of their serotypes. These strain variable proteins can provide useful insights into defining the pathogenic variation that exists among clinical *S. pneumoniae* isolates.

Forty protein spots corresponding to 32 unique proteins were identified by LC-MS and functionally classified using the KEGG database. Metabolic proteins accounted for 50% of the differentially expressed proteins identified and the majority of these were involved in carbohydrate and nucleotide metabolisms. The ability of bacteria to respond to the diverse environments within the host is essential for their survival and metabolic proteins, particularly those involved in stress response and carbohydrate utilization, are central to pathogen infectivity [29–32].

One of the metabolic proteins identified as being upregulated in the invasive isolates was thioredoxin reductase (spot 1188). This enzyme is an ubiquitous antioxidant that plays an important role in the bacterial defence against extracellular oxidative stress and protects proteins from degradation or inactivation by reactive oxygen species (ROS) [33]. The enzyme acts by directly reducing oxidised cysteine or providing reducing equivalents to methionine sulfoxide reductase (Msr) enzymes, which reduce oxidized methionines [34,35]. Thioredoxin reductase enzymes have been extensively studied in *E. coli* where they play an anti-inflammatory role by suppressing neutrophil chemotaxis [36]. In the present study, there was a 1.6-fold increase in thioredoxin reductase abundance in the invasive compared to the non-invasive site isolates and the increased levels of ROS produced by host defences, such as the lactoperoxidase system, may account for the high levels of this protein in invasive site pneumococci.

The carbohydrate metabolic proteins, LacA (spot 1752) and LacB (spots 1639, 2182) showed 1.4- to 2.4-fold increase in abundance in the non-invasive site isolates. Glucose levels in the blood are generally higher than at other host sites infected by *S. pneumoniae* [31]. To thrive in low glucose host niches, pneumococci have developed metabolic mechanisms for pursuing alternative energy sources. Both LacA and LacB are involved in the tagatose-6-phosphate pathway, which is important for the utilization of galactose and lactose [37]. The expression of both enzymes is regulated by a global protein regulator, catabolite control protein A (CcpA; spot 2187), through a regulatory process called carbon catabolite repression. CcpA, which represents an important virulence determinant and metabolic regulator of sugar in *S. pneumoniae* [38,39] was upregulated in the non-invasive site isolates in this study. This protein was also detected with a M_r and pI that differed significantly from the predicted values for M_r and pI, suggesting it might undergo post-translational modification. Iyer and colleagues implicated CcpA in pneumococcal fitness during mucosal tissue infection by showing that *ccpA* mutant bacteria were attenuated in nasopharyngeal colonisation and infection of the lungs in murine models [39].

The capsular polysaccharide (CPS) biosynthesis protein, tyrosine-protein kinase (CpsD), was increased in abundance in the invasive site isolates. The regulation of CPS production through phosphorylation of the tyrosine residues in CpsD is crucial for the virulence of *S. pneumoniae* in mice [40,41]. *S. pneumoniae* undergoes spontaneous phase variation in CPS expression between an opaque and transparent form. The opaque phase is associated with increased capsule production and predominates in the blood [42]. The CpsD protein identified in the present study has a pI of 6.49 which differed from the theoretical pI of 8.96, thus suggesting a modification of the protein potentially by phosphorylation [43]. Tyrosine phosphorylated CpsD acts in concert with other CPS proteins (CpsB and CpsC) to regulate the assembly, transport and attachment of CPS to the pneumococcal cell wall [40,41,44,45]. According to Morona *et al.*, [44,45] tyrosine phosphorylation of CpsD negatively regulates CPS production. In addition CpsD-deficient *S. pneumoniae* D39 bacteria produced a significant amount of CPS, but were unable to cause bacteremia in a murine model. However, Bender and colleagues [40] associated high levels of phosphorylated CpsD with increased CPS production in the *S. pneumoniae* D39 strain. Despite these contrasting findings on the precise role of tyrosine phosphorylation in the regulation of CPS production in *S. pneumoniae* D39 strain, the increased abundance of CpsD in the invasive site isolates in our study suggests a potential role in increasing CPS production. This proposal is consistent with the findings of the many studies that reported an association of increased CPS production with *S. pneumoniae* isolated from invasive human infections such as bacteraemia, sepsis and meningitis [42,46–48].

Among the proteins that showed no differences in abundance between the pneumococcal isolates were several proteins with well-documented roles in pneumococcal virulence and pathogenesis. These included the stress response proteins, DnaK and SpxB and the "moonlighting protein", Eno. Both Eno and SpxB can induce protective immune responses in mouse infection models and could serve as potential immunogens for inclusion in a pneumococcal protein-based vaccine [49]. Although the pneumococcal DnaK protein has been considered as a vaccine candidate [50] it induces only a weak immune response in mice [49]. The low antigenicity of DnaK was due to the lack of surface-exposed epitopes on live intact pneumococcal cells [51]. PSORTb analysis of the inferred protein amino acid sequence derived from the UniProt database indicated DnaK to be of cytoplasmic origin.

The data presented in this study have demonstrated the potential of proteomics to not only discriminate clinical *S. pneumoniae* isolates recovered from invasive and non-invasive sites in the host, but also to identify candidate proteins that may be implicated in these two different aspects of the biology of *S. pneumoniae*. However, establishing the relationship between

differential patterns of bacterial protein synthesis and the clinical phenotypes requires the inclusion of more pneumococcal isolates representing other serotypes and sites of isolation such as CSF, although isolates from these different niches may well have their own specific protein requirements. The proteomic analyses used in this study were biased towards the cytoplasmic proteins as evident by the small number of extracellular proteins identified. Given the significance of extracellular proteins in pneumococcal virulence and as potential vaccine antigens, their selective recovery and analysis will provide improved proteomic coverage to define further the clinical phenotypes. Further characterisation and confirmation of the proteins' variation will be required using complementary techniques (for example by Western blotting or qPCR), although these approaches have limitations when investigating variable clinically relevant bacterial isolates as discussed elsewhere [15]. These difficulties will be increased in studies of genetically diverse bacteria such as *S. pneumoniae*. Nevertheless, the data presented above demonstrate the capacity of proteomics to differentiate clinical *S. pneumoniae* isolates according to their sites of isolation and to identify some of the proteins involved, and will lead to a better understanding of this significant human pathogen as well as the identification of new diagnostic and therapeutic markers.

Materials and methods

Origin of *S. pneumoniae* isolates

The *S. pneumoniae* isolates used in this study were selected from a collection of clinical pneumococcal isolates available from Aberdeen Royal Infirmary, Scotland; information was available on the isolate serotypes and sites of isolation [52]. The *S. pneumoniae* isolates characterised in the following study were isolated from "invasive" (blood) and "non-invasive" (sputum, tracheal aspirate, ear, eye and skin swabs) sites and included 5 different capsular serotypes of pneumococcus (Table 1). The bacterial isolates were originally serotyped by the Scottish Pneumococcal Reference Laboratory, Stobhill Hospital, Glasgow, United Kingdom, by coagglutination [53]. All of the *S. pneumoniae* isolates were stored at -70˚C in microorganism preservation vials, PROTECT™ (Technical Service Consultants Ltd, Lancashire, UK) and recovered as described in the following section. The bacteria were grown on a blood agar plate together with a 5 μg Optochin disk to confirm their identity as *S. pneumoniae* by optochin sensitivity. The serotypes of the recovered isolates were confirmed by molecular serotyping as described in the Materials an methods section.

Growth of *S. pneumoniae* isolates

Viable bacteria were recovered from the storage medium by inoculation onto 5% Columbia Blood Agar plates (E & O Laboratories Ltd, Bonnybridge, Scotland), and incubated at 37˚C for 2–3 days in an anaerobic jar containing a CampyGen™ gas pack (Oxoid Ltd, Basingstoke, UK) for the generation of a microaerophilic atmosphere (5% Oxygen, 10% Carbon dioxide, and 85% Nitrogen). Between 20 and 30 *S. pneumoniae* colonies were passaged initially into 10 mL Brain Heart Infusion (BHI) broth (Oxoid, UK) supplemented with 5% (v/v) Foetal Calf Serum (FCS). The inoculated BHI-FCS broth medium was incubated stationary at 37˚C for 15 h in an anaerobic jar containing a CampyGen™ gas pack. In order to generate a sufficient number of bacterial cells for the analysis of *S. pneumoniae* cellular proteins, a subculture in BHI-FCS broth from the first bacterial passage was performed. Briefly, 3 mL of bacterial suspension from the 10 mL broth culture (OD650 of 1.2) was added into 20 mL fresh BHI-FCS broth to give a starting OD650 of 0.1 and incubated stationary at 37˚C under a microaerophilic atmosphere until an OD650 of 1.0 was reached.

Monitoring the growth of *S. pneumoniae* isolates *in vitro*

The bacterial growth and viability were determined as follows. The *S. pneumoniae* isolates were grown in 20 mL BHI-FCS broth at 37°C in two independent biological replicates as previously described. Optical density measurements at 650 nm wavelength (OD650) were taken at 2 h intervals between 0 h and 16 h post-inoculation to determine the bacterial cell density. Serial 10-fold dilutions of the growth media ranging from 10^{-1} to 10^{-7} were prepared at times post-inoculation shown in Fig 1B to determine the bacterial viability. Five microliters of each dilution were plated onto blood agar plates and colony counts determined after 15 h incubation of the agar plates at 37°C in an anaerobic jar containing a CampyGen™ gas pack.

Molecular capsular typing

The serotypes of the pneumococcal isolates used in this study were confirmed by PCR using primers published by the Centre for Disease Control (http://www.cdc.gov/streplab/pcr.html) shown in S2 Table. Bacterial DNA was extracted from one blood agar plate showing a confluent growth of bacteria by a boiling method. Briefly the bacterial cells were resuspended in 1 mL of molecular grade water and washed twice by centrifugation at 14000 x g for 10 min. The cell pellet was resuspended in 200 uL of molecular grade water and incubated in a boiling water bath for 10 min. The sample was cooled in ice for 5 min and then clarified at 14000 x g for 10 min. The extracted bacterial DNA was used in a duplex PCR reaction for the simultaneous detection of the serotype-specific genes and the pneumococcal capsular *wxg* gene, which was used as a positive control. Each 25 uL PCR reaction contained 12.5 uL of 2X MyTaq™ Red Mix (Bioline, UK), 500 nM of each primer and 25 ng bacterial DNA and was amplified using a GeneAmp™ PCR System 9700 thermal cycler (Applied Biosystems, USA) according to the following thermal cycling steps: 95°C for 5 min; 35 amplification cycles of 94°C for 30 s, 54°C for 1 min, and 72°C for 1 min; and a final extension step at 72°C for 7 min. The PCR products were separated on a 3% (w/v) Agarose LE gel (Sigma-Aldrich, UK) containing 10 uL of SYBR® Safe DNA gel stain (Thermo Fisher Scientific, UK) at 110 volts for 45 min and visualised using a UV illuminator.

Extraction of total bacterial cellular proteins and analysis by 2-DE

Whole cell protein extracts were prepared from three independent cultures of each pneumococcal isolate prepared as previously described. The bacterial cells were pelleted from the culture media by centrifugation at 2500 x g for 10 min and washed twice with sterile PBS. The cell pellets were resuspended in lysis buffer (7 M urea, 2 M thiourea, 4% (w/v) CHAPS, 0.3% (w/v) DTT, 1% (v/v) IPG buffer, pH gradient 4–7) (GE Healthcare, Buckinghamshire, UK) and incubated with lysozyme (10 μg/μL) and a nuclease cocktail containing DNase I (3 μg/μL) and RNase A (1.5 μg/μL) (Sigma-Aldrich, St. Louis, USA) for 30 min on ice. The enzyme treated cell suspensions were then frozen overnight at -20°C to enhance cell disruption. The treated cell suspensions were sonicated six times by continuous pulse at 90% power for 30 s each using an Ultrason 250 Sonicator (LabPlant, UK) fitted with a microprobe. To minimise protein degradation by proteolysis, the cell lysates were cooled for 30 s on ice between each pulse of sonication. Finally, the cell lysates were clarified at 16200 x g for 10 min and the soluble proteins in the supernatants concentrated using the ReadyPrep™ 2-D Cleanup Kit (Bio-Rad Laboratories Ltd, Hertfordshire, UK) according to the manufacturer's instructions. The protein concentrations were determined by RC DC Protein assay using BSA (Bio-Rad Laboratories, Hercules, USA) as the standard. The bacterial proteins (100 μg) were analysed by 2-DE using small format 2-DE gels according to the protocol described by Cash and Argo [54]. The proteins were separated on 7 cm IPG gel strips (Bio-Rad Laboratories Ltd, Hertfordshire, UK) with a pH

gradient of 4–7 in the first dimension and the second dimension separation was carried out on 12% single concentration polyacrylamide slab gels (Bio-Rad Laboratories Ltd, Hertfordshire, UK). The resolved proteins were detected by Colloidal Coomassie Blue G250 staining as described by Cash and Argo [54].

Imaging and analysis of the 2-DE protein profiles

The 2-DE protein profiles were scanned at 600 dpi using an Image Scanner III (GE Healthcare Ltd, UK) supported by Labscan software. The 16 bit gel image files were transferred to the Progenesis SameSpots™ software, version 4.5 (Nonlinear Dynamics, Newcastle, UK) for spot detection and analysis as a single data set according to the method described by Smith *et al*. [15]. Briefly, gel images representing three biological replicates for each bacterial isolate were reviewed and cropped as required. A reference gel image was selected for the data set of 2-DE gel images. The remaining gels were sequentially matched against the reference gel and spots on the 2-DE profiles were detected and transformed into normalised abundances using the inbuilt routines provided by the Progenesis SameSpots™ software. The detected spots were automatically assigned numbers by the software and these are used to refer to the spots in the text. Statistical differences in normalised spot abundance between the groups were detected by ANOVA ($p < 0.05$) using the software's statistical package. Further analysis of the data by Principal Component Analysis (PCA), using the inbuilt software routines was carried out as described by Smith *et al*. [15]. Fold-differences in the mean normalised protein spot volumes between the invasive and non-invasive site isolates were calculated relative to the values for the invasive group, which was used as a reference.

In-Gel protein digestion

Protein spots, selected according to the criteria described in the Results section, were excised as 1 mm diameter gel plugs from dried Coomassie Blue stained 2-DE gels and transferred into a 96 well reaction plate for in-gel digestion using a ProGest digestor robot (Genome solutions, Huntingdon, UK). The gel plugs were first equilibrated in water, destained with acetonitrile and then sequentially washed in 25mM ammonium bicarbonate (AMBIC), acetonitrile and 50 mM AMBIC. The gel pieces were finally dehydrated with acetonitrile and the cysteine residues of the proteins reductively alkylated by incubating for 10 mins at 60˚C in 30 µL reducing solution (10 mM dithiothreitol in 50 mM AMBIC), followed by incubation at room temperature (RT) for 15 min in 30 µL alkylating solution (55 mM iodoacetamide in 50 mM AMBIC). This was followed by washing and drying of the gel pieces with 50 mM AMBIC and acetonitrile respectively. The proteins within the gel pieces were digested with 15 µL trypsin (20 ng/µl in sequencing grade water, Promega, UK) for 8 h at 37˚C. The peptide digests were eluted in a mixture of 10% formic acid and acetonitrile and dried in a SpeedVac (Savant, USA). The dried peptides were reconstituted in 50 µL loading buffer (98% water, 2% acetonitrile and 0.1% formic acid) and dissolved at RT by shaking for 15 mins at 1800 rpm. The tryptic peptide digests were clarified at 18,800 x g for 5 min and the peptides analysed by LC-MS.

Protein identification by LC-MS

The tryptic peptide digests were analysed using an UltiMate 3000 ultra-high-pressure nano HPLC system (Dionex (UK) Ltd., Camberley, Surrey, UK) interfaced with a HCTultra PTM Discovery System (Bruker Daltonics Ltd., Coventry, UK). The peptides were separated on a PepSwift monolithic PS-DVB capillary column (200 µm i.d. x 5 cm; Dionex (UK) Ltd.) at a flow rate of 2.0 µL/min using a linear gradient of acetonitrile provided by solvent A (3% acetonitrile in 0.05% formic acid) and solvent B (80% acetonitrile in 0.04% formic acid). The

peptide separation was carried out using a gradient of 3–45% solvent B for 12 minutes, followed by a column wash in 90% solvent B for 1 minute and then equilibration in 3% solvent B for 5 minutes. MS/MS data, averaged from two spectra, with a scan range of 100–2200 m/z, were acquired in data-dependent Auto MS2 mode. Up to 3 precursor ions were selected from the MS scan (range 300–1500 m/z, averages 3) in each Auto MS2 cycle. Precursors were actively excluded after selection twice within a 1.0 min window, and singly-charged ions were also excluded. Peptide peaks were detected and deconvoluted automatically using Data Analysis Software (Bruker). Mass lists in the form of Mascot Generic Files were created automatically and used as the input for the MS/MS ion searches in the NCBIprot database. The search was performed by a local MASCOT server (version 2.2) using the Matrix Science web server (www.matrixscience.com) and the search parameters were as follows: Taxonomy—S. pneumoniae; Enzyme—Trypsin; Fixed modifications—Carbamidomethyl (C); Variable modification —Oxidation (M); Mass values—Monoisotopic; Peptide mass tolerance– 1.5 Da; Fragment mass tolerance– 0.5 Da; Max missed cleavages– 1; Peptide charge– 2+ and 3+; Instrument type —ESI-TRAP. The MASCOT significance threshold $p < 0.05$ was used for peptide identification and a false discovery rate (FDR) of less than 1% using a target decoy search was used for the peptide validation.

Bioinformatic analysis of identified proteins

The identified proteins were categorised into their respective biological functions using the Kyoto Encyclopedia of Genes and Genome (KEGG) database (http://www.genome.jp/kegg/pathway.html). The association of the proteins with the KEGG pathways was determined using *S. pneumoniae* TIGR4 as the reference organism. The subcellular localizations of the identified proteins were predicted using the web-based algorithm PSORTb 3.0 (http://www.psort.org/psortb/index.html). Since PSORTb cannot predict lipoprotein motifs, the subcellular localizations of the proteins predicted as unknown by PSORTb were further characterised using CELLO2GO (http://cello.life.nctu.edu.tw/cello2go/).

Supporting information

S1 Table. Peptides identified by LC-MS analysis of selected *S. pneumoniae* proteins. (PDF)

S2 Table. Primers used for the molecular capsular typing. (TIF)

S1 Fig. Representative global proteome profiles of invasive (B) and non-invasive (E) site isolates grown in BHI-FCS broth and resolved on small format 2-D gels. (TIF)

Acknowledgments

Author Contributions

Conceptualization: MB PC.

Funding acquisition: MB.

Investigation: MB.

Methodology: MB.

Project administration: MB.

Resources: PC KF.

Supervision: PC KF.

Writing – original draft: MB.

Writing – review & editing: MB PC KF.

References

1. van der Poll T, Opal SM. Pathogenesis, treatment, and prevention of pneumococcal pneumonia. The Lancet 2009 Oct 31; 374(9700):1543–1556. https://doi.org/10.1016/S0140-6736(09)61114-4 PMID: 19880020.

2. Mook-Kanamori BB, Geldhoff M, van der Poll T, van de Beek D. Pathogenesis and pathophysiology of pneumococcal meningitis. Clin Microbiol Rev 2011 Jul; 24(3):557–591. https://doi.org/10.1128/CMR.00008-11 PMID: 21734248.

3. Vila-Corcoles A, Ochoa-Gondar O. Preventing pneumococcal disease in the elderly: recent advances in vaccines and implications for clinical practice. Drugs Aging 2013 May; 30(5):263–276. https://doi.org/10.1007/s40266-013-0060-5 PMID: 23420119.

4. Royal College of Physicians of Ireland. Immunisation Advisory Committee. Immunisation guidelines for Ireland. 2008. [cited 2016 Nov 3]. http://health.gov.ie/wp-content/uploads/2014/03/immunisation_guidelines.pdf

5. Bittaye M, Cash P. Streptococcus pneumoniae proteomics: determinants of pathogenesis and vaccine development. Expert review of proteomics 2015; 12(6):607–621. https://doi.org/10.1586/14789450.2015.1108844 PMID: 26524107.

6. Obert C, Sublett J, Kaushal D, Hinojosa E, Barton T, Tuomanen EI, et al. Identification of a Candidate Streptococcus pneumoniae core genome and regions of diversity correlated with invasive pneumococcal disease. Infect Immun 2006 Aug; 74(8):4766–4777. https://doi.org/10.1128/IAI.00316-06 PMID: 16861665

7. Blomberg C, Dagerhamn J, Dahlberg S, Browall S, Fernebro J, Albiger B, et al. Pattern of accessory regions and invasive disease potential in Streptococcus pneumoniae. J Infect Dis 2009 Apr 1; 199 (7):1032–1042. https://doi.org/10.1086/597205 PMID: 19203261.

8. Ogunniyi AD, Giammarinaro P, Paton JC. The genes encoding virulence-associated proteins and the capsule of Streptococcus pneumoniae are upregulated and differentially expressed in vivo. Microbiology 2002 Jul; 148(Pt 7):2045–2053. https://doi.org/10.1099/00221287-148-7-2045 PMID: 12101293.

9. Orihuela CJ, Radin JN, Sublett JE, Gao G, Kaushal D, Tuomanen EI. Microarray analysis of pneumococcal gene expression during invasive disease. Infect Immun 2004 Oct; 72(10):5582–5596. https://doi.org/10.1128/IAI.72.10.5582-5596.2004 PMID: 15385455

10. Guo Z, Han J, Yang XY, Cao K, He K, Du G, et al. Proteomic analysis of the copper resistance of Streptococcus pneumoniae. Metallomics 2015 Mar; 7(3):448–454. https://doi.org/10.1039/c4mt00276h PMID: 25608595.

11. Sanchez CJ, Hurtgen BJ, Lizcano A, Shivshankar P, Cole GT, Orihuela CJ. Biofilm and planktonic pneumococci demonstrate disparate immunoreactivity to human convalescent sera. BMC Microbiol 2011 Nov 2; 11:245. https://doi.org/10.1186/1471-2180-11-245 PMID: 22047041.

12. Bae SM, Yeon SM, Kim TS, Lee KJ. The effect of protein expression of Streptococcus pneumoniae by blood. J Biochem Mol Biol 2006 Nov 30; 39(6):703–708. PMID: 17129405.

13. Lee MR, Bae SM, Kim TS, Lee KJ. Proteomic analysis of protein expression in Streptococcus pneumoniae in response to temperature shift. J Microbiol 2006 Aug; 44(4):375–382. PMID: 16953172.

14. Chaussee MA, Dmitriev AV, Callegari EA, Chaussee MS. Growth phase-associated changes in the transcriptome and proteome of Streptococcus pyogenes. Arch Microbiol 2008 Jan; 189(1):27–41. https://doi.org/10.1007/s00203-007-0290-1 PMID: 17665172.

15. Smith A, van Rooyen JP, Argo E, Cash P. Proteomic analysis of Escherichia coli associated with urinary tract infections. Proteomics 2011 Jun; 11(11):2283–2293. https://doi.org/10.1002/pmic.201000626 PMID: 21598392.

16. Fowsantear W, Argo E, Pattinson C, Cash P. Comparative proteomics of Helicobacter species: the discrimination of gastric and enterohepatic Helicobacter species. J Proteomics 2014 Jan 31; 97:245–255. https://doi.org/10.1016/j.jprot.2013.07.016 PMID: 23899588.

17. Park SA, Lee HW, Hong MH, Choi YW, Choe YH, Ahn BY, et al. Comparative proteomic analysis of Helicobacter pylori strains associated with iron deficiency anemia. Proteomics 2006 Feb; 6(4):1319–1328. https://doi.org/10.1002/pmic.200500293 PMID: 16404725.

18. McAllister LJ, Ogunniyi AD, Stroeher UH, Leach AJ, Paton JC. Contribution of serotype and genetic background to virulence of serotype 3 and serogroup 11 pneumococcal isolates. Infect Immun 2011 Dec; 79(12):4839–4849. https://doi.org/10.1128/IAI.05663-11 PMID: 21930754.

19. Isaacman DJ, McIntosh ED, Reinert RR. Burden of invasive pneumococcal disease and serotype distribution among Streptococcus pneumoniae isolates in young children in Europe: impact of the 7-valent pneumococcal conjugate vaccine and considerations for future conjugate vaccines. International Journal of Infectious Diseases 2010; 14(3):e209. https://doi.org/10.1016/j.ijid.2009.05.010 PMID: 19700359

20. Song JY, Nahm MH, Moseley MA. Clinical Implications of Pneumococcal Serotypes: Invasive Disease Potential, Clinical Presentations, and Antibiotic Resistance. Journal of Korean Medical Science 2013 Jan 1; 28(1):4–15. https://doi.org/10.3346/jkms.2013.28.1.4 PMID: 23341706

21. Lee KJ, Bae SM, Lee MR, Yeon SM, Lee YH, Kim KS. Proteomic analysis of growth phase-dependent proteins of Streptococcus pneumoniae. Proteomics 2006 Feb; 6(4):1274–1282. https://doi.org/10.1002/pmic.200500415 PMID: 16429463.

22. Brueggemann AB, Griffiths DT, Meats E, Peto T, Crook DW, Spratt BG. Clonal relationships between invasive and carriage Streptococcus pneumoniae and serotype- and clone-specific differences in invasive disease potential. J Infect Dis 2003 May 1; 187(9):1424–1432. https://doi.org/10.1086/374624 PMID: 12717624.

23. Hanage WP, Kaijalainen TH, Syrjänen RK, Auranen K, Leinonen M, Mäkelä PH, et al. Invasiveness of Serotypes and Clones of Streptococcus pneumoniae among Children in Finland. Infection and Immunity 2005 Jan 1; 73(1):431–435. https://doi.org/10.1128/IAI.73.1.431-435.2005 PMID: 15618181.

24. McEllistrem MC, Noller AC, Visweswaran S, Adams JM, Harrison LH. Serotype 14 variants of the France 9V(-3) clone from Baltimore, Maryland, can be differentiated by the cpsB gene. J Clin Microbiol 2004 Jan; 42(1):250–256. PMID: 14715761. https://doi.org/10.1128/JCM.42.1.250-256.2004

25. McGee L, McDougal L, Zhou J, Spratt BG, Tenover FC, George R, et al. Nomenclature of major antimicrobial-resistant clones of Streptococcus pneumoniae defined by the pneumococcal molecular epidemiology network. J Clin Microbiol 2001 Jul; 39(7):2565–2571. https://doi.org/10.1128/JCM.39.7.2565-2571.2001 PMID: 11427569.

26. Coffey TJ, Daniels M, Enright MC, Spratt BG. Serotype 14 variants of the Spanish penicillin-resistant serotype 9V clone of Streptococcus pneumoniae arose by large recombinational replacements of the cpsA-pbp1a region. Microbiology 1999 Aug; 145 (Pt 8)(Pt 8):2023–2031. https://doi.org/10.1099/13500872-145-8-2023 PMID: 10463168.

27. Donati C, Hiller NL, Tettelin H, Muzzi A, Croucher NJ, Angiuoli SV, et al. Structure and dynamics of the pan-genome of Streptococcus pneumoniae and closely related species. Genome biology 2010; 11(10): R107. https://doi.org/10.1186/gb-2010-11-10-r107 PMID: 21034474.

28. Hiller NL, Janto B, Hogg JS, Boissy R, Yu S, Powell E, et al. Comparative genomic analyses of seventeen Streptococcus pneumoniae strains: insights into the pneumococcal supragenome. J Bacteriol 2007 Nov; 189(22):8186–8195. https://doi.org/10.1128/JB.00690-07 PMID: 17675389.

29. Yamamoto T. Stress response of pathogenic bacteria—are stress proteins virulence factors? Nihon Saikingaku Zasshi 1996 Oct; 51(4):1025–1036. PMID: 8994347.

30. Rohmer L, Hocquet D, Miller SI. Are pathogenic bacteria just looking for food? Metabolism and microbial pathogenesis. Trends Microbiol 2011 Jul; 19(7):341–348. https://doi.org/10.1016/j.tim.2011.04.003 PMID: 21600774.

31. Shelburne SA, Davenport MT, Keith DB, Musser JM. The role of complex carbohydrate catabolism in the pathogenesis of invasive streptococci. Trends Microbiol 2008 Jul; 16(7):318–325. https://doi.org/10.1016/j.tim.2008.04.002 PMID: 18508271.

32. Polissi A, Pontiggia A, Feger G, Altieri M, Mottl H, Ferrari L, et al. Large-scale identification of virulence genes from Streptococcus pneumoniae. Infect Immun 1998 Dec; 66(12):5620–5629. PMID: 9826334.

33. Gennaris A, Collet JF. The 'captain of the men of death', Streptococcus pneumoniae, fights oxidative stress outside the 'city wall'. EMBO Mol Med 2013 Dec; 5(12):1798–1800. https://doi.org/10.1002/emmm.201303482 PMID: 24293314.

34. Holmgren A. Thioredoxin and glutaredoxin systems. J Biol Chem 1989 Aug 25; 264(24):13963–13966. PMID: 2668278.

35. Collet JF, Messens J. Structure, function, and mechanism of thioredoxin proteins. Antioxid Redox Signal 2010 Oct; 13(8):1205–1216. https://doi.org/10.1089/ars.2010.3114 PMID: 20136512.

36. Nakamura H, Herzenberg LA, Bai J, Araya S, Kondo N, Nishinaka Y, et al. Circulating thioredoxin suppresses lipopolysaccharide-induced neutrophil chemotaxis. Proc Natl Acad Sci U S A 2001 Dec 18; 98 (26):15143–15148. https://doi.org/10.1073/pnas.191498798 PMID: 11742067.

37. Afzal M, Shafeeq S, Kuipers OP. LacR is a repressor of lacABCD and LacT is an activator of lacTFEG, constituting the lac gene cluster in Streptococcus pneumoniae. Appl Environ Microbiol 2014 Sep; 80 (17):5349–5358. https://doi.org/10.1128/AEM.01370-14 PMID: 24951784.

38. Carvalho SM, Kloosterman TG, Kuipers OP, Neves AR. CcpA ensures optimal metabolic fitness of Streptococcus pneumoniae. PLoS One 2011; 6(10):e26707. https://doi.org/10.1371/journal.pone.0026707 PMID: 22039538.

39. Iyer R, Baliga NS, Camilli A. Catabolite control protein A (CcpA) contributes to virulence and regulation of sugar metabolism in Streptococcus pneumoniae. J Bacteriol 2005 Dec; 187(24):8340–8349. https://doi.org/10.1128/JB.187.24.8340-8349.2005 PMID: 16321938

40. Bender MH, Cartee RT, Yother J. Positive correlation between tyrosine phosphorylation of CpsD and capsular polysaccharide production in Streptococcus pneumoniae. J Bacteriol 2003 Oct; 185 (20):6057–6066. PMID: 14526017. https://doi.org/10.1128/JB.185.20.6057-6066.2003

41. Morona JK, Miller DC, Morona R, Paton JC. The effect that mutations in the conserved capsular polysaccharide biosynthesis genes cpsA, cpsB, and cpsD have on virulence of Streptococcus pneumoniae. J Infect Dis 2004 May 15; 189(10):1905–1913. https://doi.org/10.1086/383352 PMID: 15122528.

42. Weiser JN. Phase variation in colony opacity by Streptococcus pneumoniae. Microb Drug Resist 1998; 4(2):129–135. https://doi.org/10.1089/mdr.1998.4.129 PMID: 9651000.

43. Zhu K, Zhao J, Lubman DM, Miller FR, Barder TJ. Protein pI shifts due to posttranslational modifications in the separation and characterization of proteins. Anal Chem 2005 May 1; 77(9):2745–2755. https://doi.org/10.1021/ac048494w PMID: 15859589.

44. Morona JK, Paton JC, Miller DC, Morona R. Tyrosine phosphorylation of CpsD negatively regulates capsular polysaccharide biosynthesis in streptococcus pneumoniae. Mol Microbiol 2000 Mar; 35 (6):1431–1442. PMID: 10760144.

45. Morona JK, Morona R, Miller DC, Paton JC. Mutational analysis of the carboxy-terminal (YGX)4 repeat domain of CpsD, an autophosphorylating tyrosine kinase required for capsule biosynthesis in Streptococcus pneumoniae. J Bacteriol 2003 May; 185(10):3009–3019. PMID: 12730159. https://doi.org/10.1128/JB.185.10.3009-3019.2003

46. Andersson B, Eriksson B, Falsen E, Fogh A, Hanson LA, Nylen O, et al. Adhesion of Streptococcus pneumoniae to human pharyngeal epithelial cells in vitro: differences in adhesive capacity among strains isolated from subjects with otitis media, septicemia, or meningitis or from healthy carriers. Infect Immun 1981 Apr; 32(1):311–317. PMID: 7216490.

47. Serrano I, Melo-Cristino J, Ramirez M. Heterogeneity of pneumococcal phase variants in invasive human infections. BMC Microbiol 2006 Jul 26; 6:67. https://doi.org/10.1186/1471-2180-6-67 PMID: 16872486.

48. Kim JO, Weiser JN. Association of intrastrain phase variation in quantity of capsular polysaccharide and teichoic acid with the virulence of Streptococcus pneumoniae. J Infect Dis 1998 Feb; 177(2):368–377. PMID: 9466523.

49. Ling E, Feldman G, Portnoi M, Dagan R, Overweg K, Mulholland F, et al. Glycolytic enzymes associated with the cell surface of Streptococcus pneumoniae are antigenic in humans and elicit protective immune responses in the mouse. Clin Exp Immunol 2004 Nov; 138(2):290–298. https://doi.org/10.1111/j.1365-2249.2004.02628.x PMID: 15498039.

50. Kim SW, Choi IH, Kim SN, Kim YH, Pyo SN, Rhee DK. Molecular cloning, expression, and characterization of dnaK in Streptococcus pneumoniae. FEMS Microbiol Lett 1998 Apr 15; 161(2):217–224. PMID: 9570114.

51. Kolberg J, Hoiby EA, Aase A, Sletten K, Rodal G, Michaelsen TE, et al. Streptococcus pneumoniae heat shock protein 70 does not induce human antibody responses during infection. FEMS Immunol Med Microbiol 2000 Dec; 29(4):289–294. PMID: 11118910.

52. Amezaga MR, Carter PE, Cash P, McKenzie H. Molecular epidemiology of erythromycin resistance in Streptococcus pneumoniae isolates from blood and noninvasive sites. J Clin Microbiol 2002 Sep; 40 (9):3313–3318. PMID: 12202572. https://doi.org/10.1128/JCM.40.9.3313-3318.2002

53. Smart LE. Serotyping of Streptococcus pneumoniae strains by coagglutination. J Clin Pathol 1986 Mar; 39(3):328–331. PMID: 3958218.

54. Cash P, Argo E. Analysis of bacterial proteins by 2DE. Methods Mol Biol 2009; 519:131–144. https://doi.org/10.1007/978-1-59745-281-6_9 PMID: 19381581.