Journal or P	OR	IGINAL RESEARCH PAPER	Microbiology	
PARIPET	ONE FOR <i>STRI</i> NEG	ASSAY FOR ALL: A MICROARRAY TECHNOLOGY IDENTIFICATION AND SEROTYPING OF EPTOCOCCUS PNEUMONIAE IN CULTURE ATIVE QMPCR POSITIVE SERUM SPECIMENS	KEY WORDS: Streptococcus pneumoniae, Microarray, homologous serotypes, serum	
Avid Hussain Shaik	Avid HussainMSc,Central Research Laboratory, Kempegowda Institute of Medical ScShaikNear BDA Complex, Banashankari Stage II, Bengaluru, Karnataka 560070			
Kadahalli Lingegowda Ravikumar*MBBS, MD Central Research Laboratory, Kempegowda Institute of N Sciences, Near BDA Complex, Banashankari Stage II, Bengaluru, Karr 560070.*Corresponding Author			egowda Institute of Medical age II, Bengaluru, Karnataka	
Seraturing of Streptococcus preumoniae is of importance for disease management and vaccine studies. Current seraturing				

ABSTRACT

Serotyping of *Streptococcus pneumoniae* is of importance for disease management and vaccine studies. Current serotyping methods require culturing which is often negative and cannot differentiate homologous serotypes. Serum provides a convenient and widely available source of pathogen detection, but the presence of host DNA hinder the detection of *S. pneumoniae*. This study aimed to evaluate a microarray method to detect and serotype *S. pneumoniae* from serum specimens. Unique serotype-specific probes were identified and added to the custom pneumococcal array and evaluated with 32 homologous reference strains, 8 Non- pneumococcal isolates and 454 serum samples. With a combination of Microbial DNA enrichment, the custom pneumococcal array identified serotype of homologous reference strains and serum samples accurately with an additional advantage of multiple serotypes and other respiratory bacterial pathogens detection in spiked samples. Our results signify the ability of microarray technology to identify and detect *S. pneumoniae* serotype from serum specimens.

INTRODUCTION:

Pneumonia is the leading cause of deaths worldwide, accounting for 15% of all deaths of children under 5 years old. Pneumonia killed an estimated 935000 children under the age of five in 2013.WHO estimated the high burden of pneumonia in Asia with most cases in India (43 million).¹ The proportion of pneumococcal pneumonia among all severe pneumonia was 15.8% whereas pneumococcal pneumonia deaths among all-cause pneumonia deaths were 29.5%.^{2,3}

Identification and serotyping of pneumococcus have been an integral part of both pneumococcal research and clinical settings. All currently available vaccines against *Streptococcus pneumoniae* are based on selections of the over 90 different serotypes, which underlines the importance of serotyping for surveillance and vaccine efficacy monitoring. Precise serotyping of pneumococci is essential to better understand the pathogenicity and trends of drug resistance. It closely monitors the emergence of non-vaccine strains, replacement serotypes, and new serovars.⁴

The conventional culture-based and PCR methods such as Quellung, Pulse Field Gel Electrophoresis (PFGE), Restriction fragment length polymorphism (RFLP), sequential multiplex PCR, have significant limitations, like low capacity, may fail if prior antibiotics administered, cannot quantify serotype abundance levels, and have a limited number of targets.^{5,6,7} Modern molecular methods such as quantitative multiplex PCR (qmPCR), Sequencing and Microarray developed based on capsular polysaccharide (CPS) and glycosyltransferase (GT) genes have given new life to solve the problem associated with pneumococcal detection. The limitation of these methods was their inability to distinguish homologous serotypes (serotypes which share sequence similarity with few bases pairs difference).^{9,10,11,12,13,14}

A major limitation to our understanding of pneumococcal serotype distribution is that most clinical cases, remain culture negative due to the widespread use of antibiotics prior to admission to the hospital. Conventional serotyping assays cannot be performed directly on clinical samples since the growth of *S. pneumoniae* on culture media is required. Serum, on the other hand, provides a convenient source of true disease burden and pathogen genetic information. A barrier to the approach of detecting pneumococcal infection from serum is the overwhelming ratio of human to pathogen DNA in samples with low pathogen abundance, which is typical of most clinical specimens.^{15,16} In our previous study, we successfully identified S. *pneumoniae* serogroup/ type from serum samples using our custom pneumococcal chip by enriching microbial DNA through

Microbial DNA enrichment method.¹⁷ The custom array detected the *S. pneumoniae* serogroup accurately but not the serotype of homologous strains and serum specimens. In Present study, we redesigned our custom array to identify homologous serotypes of *S. pneumoniae* from Serum specimens. Additional probes were included to detect other potential bacterial pathogens and quantify their relative abundance levels in spiked specimens.

MATERIALS AND METHODS: REFERENCE STRAINS

A total of 32 *S. pneumoniae* homologous reference strains and 8 Non-pneumococcal isolates such as *Streptococcus oralis*, *Moraxella catarrhalis*, *Neisseria meningitidis*, *Haemophilus influenzae*, *Pseudomonas aeruginosa*, *Klebsiella pneumoniae*, *Staphylococcus aureus*, and *Mycoplasma pneumoniae* were sourced from Central Research Laboratory, KIMS Hospital. Pneumococcal reference Strains stored in skimmed milk, tryptone, glucose, and glycerol (STGG) media at –80 °C, were cultured on 5% sheep blood agar (Chromogen, Hyderabad) for 18–24 hrs at 37°C with 5% Co₂.

Serum Specimens

Culture negative qmPCR positive (N=408) and culture positive qmPCR positive (N=48) *S. pneumoniae* serum samples from children <5 years with IPD were sourced from Central Research Laboratory, KIMS Hospital, collected during 2014 to 2015 from India.

Ethics Statement

The study was conducted according to the guidelines and declaration of Helsinki. Written informed consent was obtained from each patient or legal guardian as applicable.

DNA extraction

Genomic DNA was extracted from reference strains, nonpneumococcal isolates and serum specimens using QiaAmp DNA mini kit protocol (Qiagen, Germany) following manufacturer's instructions. The extracted DNA was quantified using Nanodrop (Thermo fisher scientific, USA) and stored at -20§C until further analysis.

Custom Array Designing

In our previous study, the custom pneumococcal array was designed using suredesign software (Agilent, USA) and evaluated with 90 reference strains and 104 serum samples. The custom array comprises of 15167 oligonucleotide probes with multiple probes per serotype (16 - 575 probes) targeting 90 *S. pneumoniae* serotypes.¹⁷ The custom array also contained species specific

probes for Streptococcus pyogenes, Streptococcus mitis, Streptococcus oralis, Moraxella catarrhalis, Neisseria meningitidis, Haemophilus influenzae, Pseudomonas aeruginosa, Klebsiella pneumoniae, Staphylococcus aureus, Legionella pneumophila, Chlamydophila pneumoniae and Mycoplasma pneumoniae, which reside in human nasopharynx along with S. Pneumonia.¹⁷

Designing of Unique Serotype-specific probes

In the present study, the homologous strain sequences were subjected to MUSCLE (multiple sequence comparison by log expectation) alignment software and identified unique sequences. (Table-1) MUSCLE is a public domain_(https://www.ebi.ac.uk/Tools/ msa/muscle/), multiple sequence alignment software for protein and nucleotide sequences.²⁰ The unique sequences of Homologous groups/ types were further subjected to suredesign software (Agilent, USA) to design unique serotype specific probes and were added to custom pneumococcal array.

Microarray analysis of homologous reference strains and Serum samples

In the first step, the serum DNA (N=456) was subjected for NEBNext Microbiome DNA Enrichment Kit (New England BioLabs) as described in Shaik AH et al 2019¹⁷ and separated microbial DNA from host DNA. The enriched microbial DNA (N=456) (Figure-1) from serum and un-enriched DNA from homologous reference strains (N=32) (Table-2) was further subjected for microarray protocol as described in Shaik AH et al 2019 without any modifications. The samples were blinded throughout microarray experiment and data analysis.

...

The redesigned custom array was validated for detection of multiple serotypes by pooling reference strains and spiking other commensals with different serotypes in 1:1 ratio and subjected for Agilent microarray protocol described previously. (Table-3)

Relative abundance levels quantification

Multiple serotype detection:

The Image acquisition was performed by Agilent DNA scanner using Agilent Scan control software v.9.0. The data processing and signal intensity extraction were analysed using Agilent Feature extraction software v.3.0. Probes showing signal above 98 percentile or always above median were removed as a part of probe selection. Positive probes with 100x fold change with respect to the median of the negative probe distribution were removed for noise reduction. The signal intensities were determined by subtracting the local background values (as estimated by the Agilent platform) from the per-sample median. Each probe's three technical replicates were then averaged and a robust average (obtained removing values above the 95th percentile and below the 5th percentile) of all probes for the same serotype was considered. Species relative abundances were then estimated by relativizing this species average with respect to the sum of all averages.

RESULTS:

Design probes to address homologous strains of *S. Pneumoniae*

A set of unique probes were designed and added to custom pneumococcal Array (Table-1)

Homologous	Unique	Unique probe sequence	
Serotype	Di edorq		
6A and 6B	6A	GAACATGAAAGATGAAGCAATCAATGTACTGGATAAAGTTATGAAAAAGATTTTTAAAAAAGTTTTGCCCC TTTTTTGTCCTCTAAATACAAAAATAGCCCT	
9A and 9V	9V	TTAATTTGAAGGGGAATGCCGTGATTGGCTAAGATTTGTTCGAGNGACATGATAGTAAGCATTCAAGTTCT CTTGTTTATCAAAATAAGCGCCTAGGATA	
10F and 10C	10F probe1	GAGCAGGACAAAAGAGCCTCGTAAAAGGTATTGCAACTTGGTAATACCTTTTTGAGGTGCTTTTTGATAT	
10F and 10C	10F probe	TCTAAATCAGGTTCGAGCGCTAAAGCTGAGAAGACATTCTCATACATTTATGAAACGGATCCTGACAACC	
	2	TAAACTACTIGACAACTGGTAAGGCTGCGA	
22A and 22F	22F	GCGGCGACTACTTTAGCTGCATGCTCTGGATCAGGTTCAAGCACTAAAGGTGAGAAGACATTCTCATACA TTTATGAGACAGACCCTGATAACCTCAACTATTTG	
32A and 32F	32A	GCGGCGACTACTTTAGCTGCATGCTCTGGATCAGGTTCAAGCACTAAAGGTGAGAAGACATTCTCATACA TTTATGAGACAGACCCTGATAACCTCAACTATTTG	
32A and 32F	32F	TGTCCAATGAAGAGCAAGACTTGACAGTAGAAGGAAAAGTCAAATCTGTCTTGATTGA	
33A, 33B, 33F	33B	AGGTGTGGTTCTTTTTTCGAGTGTAGCCCATAGCTTTGAGCGCATAGTGGATGGTAGTTGGATGACAGCC AAAGTCAGAAGCTATTTCAGTCAAATAAGCGTCTGGATTGTCAGTAAGATAGTTTTTAAGTCTATCTCTATC AACCTTTCTTG	
33A, 33B, 33F	33A	TGAGTTAACGGGACTTTATTGGGCATGGAAGAATCTTGATTATAATTACTTAGGTTTAGTTCATTACCGT	
12A,12F,44,46	12A	TAAATGGTATTACGTGAGATTTGGAAAACGTGTGATGCTTCTGTTATACTACCTGTTCGCTCACAATAAGAG	
		AGAACIIIIIIACGAAAAICIAIIGAAIAIGCCAIAAAAAGAIIAIACCACAIIGIGIACIAI	
12A,12F,44,46	12A 46	TAATAGACTTGTTCGGTAACTGTAAAAAGTGTTATACTATTTTTATGGAAACAGTATACGACAAAGCACAAA AACTTAACTCAAAAAACTTCAAACTATTGATTGGTGTCAAAAAGGAAACCTTTCAACTCATGCTAGAACAC CTGAATTCAGCCTATCAGATTCAGCAC	
12A,12F,44,46	12A 44 46	GCGGCGACTACTTTAGCTGCATGCTCTGGATCAGGTTCAAGCACTAAAGGTGAGAAGACATTCTCATACA TTTATGAGACAGACCCTGATAACCTCAACTATTTGACAACTGCTAAGGCTGCGACAGCAAATATTACCAGT AACGTGG	
12A,12F,44,46	12F	AAGCCTTTCGAACTTTGGAAGATGTCATGAATCAACTCCAAGATGTTATACAAGGACTGG	
11A, 11D, 18F	11A	TTTTCCGTAGCGATTTGAAGATTATAAGCAGCCTTGAGTTGACCATTTTTCATATGATCTTCCTTC	
	probe1	TAAAAGTGGCATCTGGATCGGTTTTGGAAAAACTGTTAGGCCCTTCAACTGTCTCCTGATAGTTCTCATATT TTTCAGCACGTACTGAAAAATC	
11A, 11D, 18F	18F probe1	ATTGTTAGAAATCGATTTGACTGTCCTGATCGATTTGTCATGTTCTTATTTCATTTTACTATATTTTGGTTCGC GGGAAGTCTACTAAGATACTTA	
11A, 11D, 18F	18F probe2	AGCCTTATTATCCAGTTAATAATGATCGTAATAATCATTTGTATAAATCATATAAAAAACTTGCTGATGAGCA AGGGAATGTTATCTTTGGTGGCCGCTTAGGACACTATCGTTATTACGATATGCACCAAGTAATTGGAG	
11A, 11D, 18F	18F11A	AAAATGATGAAAAGTTCAAAACTATTTGCCCTTGCGGGCGTGACATTATTG	
11A, 11D, 18F	11A probe2	GAGTCTGGAACAAAATAGTTCTCGACCACAAAAAAGCTAGAGATTTCCAATTGTGG	
11A, 11B, 11C, 11D	11D	AGTGGATTGAAACTAGAATAGTGCACCTCTGCTTCTAAAACATTGTTAGAAATCGANNNTTTGACT	
11A, 11B, 11C, 11D	11B	TAGTGCAGGCCTGACTTTTGCATCTGCTTTGCTTTTAGCTGCTGCGACCAATCAGGTTCAGATACAAAAAC TTACTCATCAACCTTTAGTGGAAATCCAACTACATTTAACTATCTAT	

www.worldwidejournals.com

Volume-8 | Issue-5 | May-2019 | PRINT ISSN No. 2250 - 1991

11A, 11B, 11C, 11D	11C	AGTAAACTCTTCATAATCCTTTTTATTTTTATGAAGATATTGTTTGAAAGATGTGAGTTTCCACGGATGGGTT TGTGGAGGGATATACTTGCGTCTTTCTTTTTTGTTCTGGTTCTTGTTCAAAGTTTTTCGAATAGAGTTCATGAT TTAGTAGCTCCTTTGTGTGATAGATTTTGTCA
9N and 9L	9N	GTTGTTTTCAAAAATCGTNNNAAGTCCCTTTTGTTAGTTAGTGTTTTGCCTATATTTTCACACTATTTATT
9N and 9L	9L	GCTGTTTTCAAAAATCGTAGAATGTTATTTTTGATANNNAGTGTTTTCGCTTATTGTTTAACCCTATTTGTTCAA ACAGATATTTATATTATTCAAAGAA
15B and 15C	15B	ATTATAAGATTTGTGATTTTTAATAGGATTTCAGATTTAATAGGCTATCTAGCAACACCGTTGTGGTATTTATT
15B and 15C	15C	ATTATAAGATTTGTGATTTTTAATAGGATTTCAGATTTAATAGGCTATCTAGCAACACCGTTGTGGTATTTATT

Table:1 Unique probe ID's and sequences of homologous serotypes

Homologous Reference strains results:

The redesigned custom pneumococcal array identified the homologous reference strains to their respective serotype accurately (Table-2).

S.NoArray noSpecimenDyeHomologo us SerotypeDetected serotype11_1Reference StrainCy36A6A21_1Reference StrainCy515C15C31_2Reference StrainCy522A22A51_3Reference StrainCy522F22F51_3Reference StrainCy532A32A61_4Reference StrainCy532A32A71_4Reference StrainCy532F32F71_4Reference StrainCy532F32F81_4Reference StrainCy532A33A92_1Reference StrainCy533A33A102_1Reference StrainCy532F32F112_2Reference StrainCy533A33A122_3Reference StrainCy533B33B152_4Reference StrainCy534646171_1Reference StrainCy59V9V191_2Reference StrainCy59L9L211_3Reference StrainCy59L9L211_3Reference StrainCy59L9L211_3Reference StrainCy59L9L211_3Reference StrainCy59L9L211_3Reference StrainCy59L </th <th></th> <th></th> <th></th> <th></th> <th></th> <th></th>						
Image: Serotype serotype serotype 1 1_1 Reference Strain Cy3 6A 6A 2 1_1 Reference Strain Cy5 15C 15C 3 1_2 Reference Strain Cy5 22A 22A 5 1_3 Reference Strain Cy5 22F 22F 7 1_4 Reference Strain Cy5 32A 32A 9 2_1 Reference Strain Cy5 32F 32F 11 2_2 Reference Strain Cy5 33A 33A 10 2_1 Reference Strain Cy5 32F 32F 11 2_2 Reference Strain Cy5 33A 33A 13 2_3 Reference Strain Cy5 33B 33B 15 2_4 Reference Strain Cy5 33B 33B 15 2_4 Reference Strain Cy5 9A 9A 18 1_1 Refe	S.No	Array no	Specimen	Dye	Homologo	Detected
1 1_1 Reference Strain Cy3 6A 6A 2 1_1 Reference Strain Cy5 15C 15C 3 1_2 Reference Strain Cy3 6B 6B 4 1_2 Reference Strain Cy3 6B 6B 4 1_2 Reference Strain Cy3 10C 10C 6 1_3 Reference Strain Cy3 10C 10C 6 1_3 Reference Strain Cy3 10F 10F 8 1_4 Reference Strain Cy3 32A 32A 9 2_1 Reference Strain Cy3 31A 11A 10 2_1 Reference Strain Cy3 32F 32F 11 2_2 Reference Strain Cy3 33A 33A 13 2_3 Reference Strain Cy3 33A 33A 13 2_3 Reference Strain Cy3 35B 15B 16 2_4 Reference Strain Cy5 33B 33B <td< td=""><td></td><td></td><td></td><td></td><td>us Serotyp</td><td>eserotype</td></td<>					us Serotyp	eserotype
2 1_1 Reference Strain $Cy5$ $15C$ $15C$ 3 1_2 Reference Strain $Cy3$ $6B$ $6B$ 4 1_2 Reference Strain $Cy5$ $22A$ $22A$ 5 1_3 Reference Strain $Cy5$ $22A$ $22A$ 5 1_3 Reference Strain $Cy5$ $22F$ $22F$ 7 1_4 Reference Strain $Cy5$ $32A$ $32A$ 9 2_1 Reference Strain $Cy5$ $32A$ $32A$ 9 2_1 Reference Strain $Cy5$ $32F$ $32F$ 11 2_2 Reference Strain $Cy5$ $33A$ $33A$ 12 2_2 Reference Strain $Cy5$ $33B$ $33B$ 12 2_2 Reference Strain $Cy5$ $33A$ $33A$ 13 2_3 Reference Strain $Cy5$ $33B$ $33B$ 15 2_4 Reference Strain $Cy5$ 46 46 17 1_1 <t< td=""><td>1</td><td>1_1</td><td>Reference Strain</td><td>СуЗ</td><td>6A</td><td>6A</td></t<>	1	1_1	Reference Strain	СуЗ	6A	6A
3 1_2 Reference Strain Cy3 6B 6B 4 1_2 Reference Strain Cy5 22A 22A 5 1_3 Reference Strain Cy3 10C 10C 6 1_3 Reference Strain Cy3 10C 10C 6 1_3 Reference Strain Cy3 10F 10F 6 1_4 Reference Strain Cy3 32A 32A 9 2_1 Reference Strain Cy3 31A 11A 10 2_1 Reference Strain Cy3 32F 32F 11 2_2 Reference Strain Cy3 31B 11B 11B 12 2_2 Reference Strain Cy3 32A 33A 13 2_3 Reference Strain Cy3 31B 11B 14 2_3 Reference Strain Cy3 32A 12A 14 2_3 Reference Strain Cy3 35B 15B 15 2_4 Reference Strain Cy5 9A<	2	1_1	Reference Strain	Cy5	15C	15C
4 1_2 Reference Strain Cy5 $22A$ $22A$ 5 1_3 Reference Strain Cy3 $10C$ $10C$ 6 1_3 Reference Strain Cy5 $22F$ $22F$ 7 1_4 Reference Strain Cy3 $10F$ $10F$ 8 1_4 Reference Strain Cy3 $32A$ $32A$ 9 2_1 Reference Strain Cy3 $31A$ $11A$ 10 2_1 Reference Strain Cy3 $32A$ $32A$ 9 2_1 Reference Strain Cy3 $31A$ $11A$ 10 2_1 Reference Strain Cy3 $32F$ $32F$ 11 2_2 Reference Strain Cy3 $33A$ $33A$ 12 2_2 Reference Strain Cy3 $32A$ $33A$ 13 2_3 Reference Strain Cy3 $32A$ $33A$ 14 2_3 Reference Strain Cy5 $33B$ $33B$ 15 2_4 Reference Strain <td>3</td> <td>1_2</td> <td>Reference Strain</td> <td>Cy3</td> <td>6B</td> <td>6B</td>	3	1_2	Reference Strain	Cy3	6B	6B
5 1_3 Reference Strain Cy3 10C 10C 6 1_3 Reference Strain Cy5 22F 22F 7 1_4 Reference Strain Cy3 10F 10F 8 1_4 Reference Strain Cy3 10F 10F 8 1_4 Reference Strain Cy3 32A 32A 9 2_1 Reference Strain Cy3 11A 11A 10 2_1 Reference Strain Cy3 32F 32F 11 2_2 Reference Strain Cy3 31B 11B 11B 12 2_2 Reference Strain Cy3 32A 33A 13 2_3 Reference Strain Cy3 32A 12A 14 2_3 Reference Strain Cy3 33B 33B 15 2_4 Reference Strain Cy3 15B 15B 16 2_4 Reference Strain Cy3 9A 9A 18 1_1 Reference Strain Cy5 9U 9V	4	1_2	Reference Strain	Cy5	22A	22A
6 1_3 Reference Strain Cy5 22F 22F 7 1_4 Reference Strain Cy3 10F 10F 8 1_4 Reference Strain Cy3 10F 10F 8 1_4 Reference Strain Cy3 32A 32A 9 2_1 Reference Strain Cy3 11A 11A 10 2_1 Reference Strain Cy3 32F 32F 11 2_2 Reference Strain Cy3 31B 11B 12 2_2 Reference Strain Cy3 32A 33A 13 2_3 Reference Strain Cy3 32A 12A 14 2_3 Reference Strain Cy3 33B 33B 15 2_4 Reference Strain Cy3 35A 35B 16 2_4 Reference Strain Cy3 9A 9A 18 1_1 Reference Strain Cy5 9V 9V 19 1_2 Reference Strain Cy5 9L 9L	5	1_3	Reference Strain	Cy3	10C	10C
7 1_4 Reference Strain Cy3 10F 10F 8 1_4 Reference Strain Cy5 32A 32A 9 2_1 Reference Strain Cy3 11A 11A 10 2_1 Reference Strain Cy3 11A 11A 10 2_1 Reference Strain Cy3 32F 32F 11 2_2 Reference Strain Cy3 11B 11B 12 2_2 Reference Strain Cy3 32A 33A 13 2_3 Reference Strain Cy3 12A 12A 14 2_3 Reference Strain Cy3 33B 33B 15 2_4 Reference Strain Cy3 15B 15B 16 2_4 Reference Strain Cy3 9A 9A 18 1_1 Reference Strain Cy3 9N 9N 20 1_2 Reference Strain Cy5 9L 9L 21 1_3 Reference Strain Cy5 9L 9L	6	1_3	Reference Strain	Cy5	22F	22F
8 1_4 Reference Strain Cy5 32A 32A 9 2_1 Reference Strain Cy3 11A 11A 10 2_1 Reference Strain Cy3 32F 32F 11 2_2 Reference Strain Cy3 11B 11B 12 2_2 Reference Strain Cy3 33A 33A 13 2_3 Reference Strain Cy3 12A 12A 14 2_3 Reference Strain Cy3 33B 33B 15 2_4 Reference Strain Cy3 15B 15B 16 2_4 Reference Strain Cy3 9A 9A 18 1_1 Reference Strain Cy3 9A 9A 18 1_1 Reference Strain Cy3 9N 9N 20 1_2 Reference Strain Cy3 35A 35A 5. No Serotype Spiked serotype/organism Multiple seroty Microarray 1 19F 1 19F and 1 2 6B	7	1_4	Reference Strain	Cy3	10F	10F
9 2_1 Reference Strain Cy3 11A 11A 10 2_1 Reference Strain Cy5 32F 32F 11 2_2 Reference Strain Cy3 11B 11B 12 2_2 Reference Strain Cy5 33A 33A 13 2_3 Reference Strain Cy5 33B 33B 13 2_3 Reference Strain Cy5 33B 33B 14 2_3 Reference Strain Cy5 33B 33B 15 2_4 Reference Strain Cy5 46 46 17 1_1 Reference Strain Cy5 9A 9A 18 1_1 Reference Strain Cy5 9V 9V 19 1_2 Reference Strain Cy5 9L 9L 21 1_3 Reference Strain Cy5 9L 9L 21 1_3 Reference Strain Cy5 9L 9L 21 1_3 Reference Strain Cy5 9L 9L <td< td=""><td>8</td><td>1_4</td><td>Reference Strain</td><td>Cy5</td><td>32A</td><td>32A</td></td<>	8	1_4	Reference Strain	Cy5	32A	32A
10 2_1 Reference Strain Cy5 32F 32F 11 2_2 Reference Strain Cy3 11B 11B 12 2_2 Reference Strain Cy5 33A 33A 13 2_3 Reference Strain Cy5 33B 33B 14 2_3 Reference Strain Cy5 33B 33B 15 2_4 Reference Strain Cy5 46 46 17 1_1 Reference Strain Cy5 9A 9A 18 1_1 Reference Strain Cy5 9V 9V 19 1_2 Reference Strain Cy5 9L 9L 21 1_3 Reference Strain Cy5 9L 9L 21	9	2_1	Reference Strain	Cy3	11A	11A
11 2_2 Reference Strain $Cy3$ 11B 11B 12 2_2 Reference Strain $Cy5$ $33A$ $33A$ 13 2_3 Reference Strain $Cy3$ $12A$ $12A$ 14 2_3 Reference Strain $Cy3$ $12A$ $12A$ 14 2_3 Reference Strain $Cy5$ $33B$ $33B$ 15 2_4 Reference Strain $Cy3$ $15B$ $15B$ 16 2_4 Reference Strain $Cy5$ 46 46 17 1_1 Reference Strain $Cy5$ $9V$ $9V$ 18 1_1 Reference Strain $Cy5$ $9V$ $9V$ 19 1_2 Reference Strain $Cy5$ $9L$ $9L$ 20 1_2 Reference Strain $Cy5$ $9L$ $9L$ 21 1_3 Reference Strain $Cy3$ $35A$ $35A$ S. No Serotype Spiked serotype/organism Multiple seroty Microarray $19F$ and 1 $19F$	10	2_1	Reference Strain	Cy5	32F	32F
12 2_2 Reference Strain Cy5 33A 33A 13 2_3 Reference Strain Cy3 12A 12A 14 2_3 Reference Strain Cy5 33B 33B 15 2_4 Reference Strain Cy3 15B 15B 16 2_4 Reference Strain Cy5 46 46 17 1_1 Reference Strain Cy5 9A 9A 18 1_1 Reference Strain Cy5 9V 9V 19 1_2 Reference Strain Cy5 9L 9L 21 1_3 Reference Strain Cy3 35A 35A S. No Serotype Spiked serotype/organism Multiple seroty Microarray 1 19F 1 19F and 1 19F and 1 2 6B 13 6B and 13 33 3 19A 3 19A and 3 19A and 3	11	2_2	Reference Strain	Cy3	11B	11B
13 2_3 Reference Strain Cy3 12A 12A 14 2_3 Reference Strain Cy5 33B 33B 15 2_4 Reference Strain Cy3 15B 15B 16 2_4 Reference Strain Cy5 46 46 17 1_1 Reference Strain Cy5 9A 9A 18 1_1 Reference Strain Cy5 9V 9V 19 1_2 Reference Strain Cy5 9L 9L 20 1_2 Reference Strain Cy5 9L 9L 21 1_3 Reference Strain Cy3 35A 35A S. No Serotype Spiked serotype/organism Multiple seroty 1 19F 1 19F and 1 19F and 1 2 6B 13 6B and 13 3 3 19A 3 19A and 3 19A and 3	12	2_2	Reference Strain	Cy5	33A	33A
14 2_3 Reference Strain Cy5 33B 33B 15 2_4 Reference Strain Cy3 15B 15B 16 2_4 Reference Strain Cy5 46 46 17 1_1 Reference Strain Cy3 9A 9A 18 1_1 Reference Strain Cy5 9V 9V 19 1_2 Reference Strain Cy5 9L 9L 20 1_2 Reference Strain Cy5 9L 9L 21 1_3 Reference Strain Cy3 35A 35A S. No Serotype Spiked serotype/organism Multiple seroty Microarray 1 19F 1 19F and 1 2 6B 13 6B and 13 3 19A 3 19A and 3	13	2_3	Reference Strain	Cy3	12A	12A
15 2_4 Reference Strain Cy3 15B 15B 16 2_4 Reference Strain Cy5 46 46 17 1_1 Reference Strain Cy3 9A 9A 18 1_1 Reference Strain Cy5 9V 9V 19 1_2 Reference Strain Cy3 9N 9N 20 1_2 Reference Strain Cy5 9L 9L 21 1_3 Reference Strain Cy3 35A 35A S. No Serotype Spiked serotype/organism Multiple seroty Microarray 1 19F 1 19F and 1 2 6B 13 6B and 13 3 19A 3 19A and 3	14	2_3	Reference Strain	Cy5	33B	33B
16 2_4 Reference Strain Cy5 46 46 17 1_1 Reference Strain Cy3 9A 9A 18 1_1 Reference Strain Cy5 9V 9V 19 1_2 Reference Strain Cy3 9N 9N 20 1_2 Reference Strain Cy5 9L 9L 21 1_3 Reference Strain Cy3 35A 35A 5. No Serotype Spiked serotype/organism Multiple seroty Microarray 1 19F 1 19F and 1 2 6B 13 6B and 13 3 19A 3 19A and 3	15	2_4	Reference Strain	СуЗ	15B	15B
17 1_1 Reference Strain Cy3 9A 9A 18 1_1 Reference Strain Cy5 9V 9V 19 1_2 Reference Strain Cy3 9N 9N 20 1_2 Reference Strain Cy5 9L 9L 21 1_3 Reference Strain Cy3 35A 35A 5. No Serotype Spiked serotype/organism Multiple seroty Microarray 1 19F 1 19F and 1 2 6B 13 6B and 13 3 19A 3 19A and 3	16	2_4	Reference Strain	Cy5	46	46
18 1_1 Reference Strain Cy5 9V 9V 19 1_2 Reference Strain Cy3 9N 9N 20 1_2 Reference Strain Cy5 9L 9L 21 1_3 Reference Strain Cy3 35A 35A 5. No Serotype Spiked serotype/organism Multiple seroty Microarray 1 19F 1 19F and 1 2 6B 13 6B and 13 3 19A 3 19A and 3	17	1_1	Reference Strain	СуЗ	9A	9A
19 1_2 Reference Strain Cy3 9N 9N 20 1_2 Reference Strain Cy5 9L 9L 21 1_3 Reference Strain Cy3 35A 35A 5. No Serotype Spiked serotype/organism Multiple seroty Microarray 1 19F 1 19F and 1 2 6B 13 6B and 13 3 19A 3 19A and 3	18	1_1	Reference Strain	Cy5	9V	9V
20 1_2 Reference Strain Cy5 9L 9L 21 1_3 Reference Strain Cy3 35A 35A S. No Serotype Spiked serotype/organism Multiple seroty Microarray 1 19F 1 19F and 1 2 6B 13 6B and 13 3 19A 3 19A and 3	19	1_2	Reference Strain	СуЗ	9N 9N	
21 1_3 Reference Strain Cy3 35A 35A S. No Serotype Spiked serotype/organism Multiple seroty Microarray 1 19F 1 19F and 1 2 6B 13 6B and 13 3 19A 3 19A and 3	20	1_2	Reference Strain	Cy5	9L	9L
S. NoSerotypeSpiked serotype/organismMultiple seroty Microarray119F119F and 126B136B and 13319A319A and 3	21	1_3	Reference Strain	СуЗ	35A	35A
Microarray 1 19F 1 19F and 1 2 6B 13 6B and 13 3 19A 3 19A and 3	S. No Serotyp		e Spiked serotype/organism			Multiple serot
1 19F 1 19F and 1 2 6B 13 6B and 13 3 19A 3 19A and 3						Microarray
2 6B 13 6B and 13 3 19A 3 19A and 3	1	19F	1	1		19F and 1
3 19A 3 19A and 3	2	6B	13			6B and 13
	3	19A	3	3		

22	1_3	Reference Strain	Cy5	35B	35B
23	1_4	Reference Strain	СуЗ	35C	35C
24	1_4	Reference Strain	Cy5	44	44
25	2_1	Reference Strain	Cy5	11D	11D
26	2_1	Reference Strain	Cy3	18C	18C
27	2_2	Reference Strain	Cy5	33F	33F
28	2_2	Reference Strain	СуЗ	11C	11C
29	2_3	Reference Strain	Cy5	18F	18F
30	2_3	Reference Strain	Cy3	42	42
31	2_4	Reference Strain	Cy5	12F	12F
32	2_4	Reference Strain	Cy5	19F	19F

Table-2: custom pneumococcal array results- Homologous reference strains

Microarray results for serum samples:

Enriched serum sample processing with Agilent Microarray protocol: The serum samples (#456) hybridized on custom pneumococcal chip were shown strong signal and serotype was identified accurately. (Figure-1). There was no signal observed for species specific probes, indicating absence of other pathogens (listed in custom array designing section) in tested serum samples.

Multiple serotype carriage detection and relative abundance levels quantification

The redesigned custom pneumococcal array identified spiked organisms and the serotypes of pooled reference strains accurately. Also determined relative abundance levels of each pooled serotype. (Table:3).

S. No	Serotype	Spiked serotype/organism	Multiple serotype detection by Microarray	Serotype relative abundance levels quantification
1	19F	1	19F and 1	19F (50%), 1 (50%)
2	6B	13	6B and 13	6B (50%), 13 (50%)
3	19A	3	19A and 3	19A (50%), 3 (50%)
4	18C	5	18C and 5	18C (50%), 5 (50%)
5	23F	13	23F and 13	23F (50%), 13 (50%)
6	6A	3	6A and 3	6A (50%), 3 (50%)
7	1	14	1 and 14	1 (50%), 14 (50%)
8	19A	23F	19A and 23F	19A (50%), 23F (50%)
9	19F	Moraxella catarrhalis	19F and Moraxella catarrhalis	19F (50%), Moraxella catarrhalis (50%)
10	6B	Pseudomonas aeruginosa	6B and Pseudomonas aeruginosa	6B (50%), Pseudomonas aeruginosa (50%)
11	19A	Streptococcus oralis	19A and Streptococcus oralis	19A (50%), Streptococcus oralis (50%)
12	18C	Neisseria meningitidis	18C and Neisseria meningitidis	18C (50%), Neisseria meningitidis (50%)
13	23F	Klebsiella pneumoniae	23F and Klebsiella pneumoniae	23F (50%), Klebsiella pneumoniae (50%)
14	6A	Mycoplasma pneumoniae	6A and Mycoplasma pneumoniae	6A (50%), Mycoplasma pneumoniae (50%)
15	1	Haemophilus influenzae	1 and Haemophilus influenzae	1 (50%), Haemophilus influenzae (50%)
16	19A	Staphylococcus aureus	19A and Staphylococcus aureus	19A (50%), Staphylococcus aureus (50%)

Table-3: Multiple serotype carriage and relative abundance levels quantification results





Figure 1: S. pneumoniae serotype distribution in 456 serum samples

Figure-2: S. pneumoniae homologous Serotype distribution in 456 serum Samples

DISCUSSION

It is implicit in the literature that homologous serotypes/groups which are similar to each other in sequence but differ in few bases are difficult to discriminate using DNA-based assays.¹³ Out of 90 S. pneumoniae serotypes, 29 serotypes shared close homology. They were 6A/6B, 9A/9V/9N/9L, 10A/10B/10C/10F, 11A/11B/11C/11D/18F, 15B/15C, 18C/18F, 22A/22F, 32A/32F, 33A/33B/33F and 12A/12F/44/46. The majority of previously published nucleic acid serotyping methods do not discriminate such serotypes. $^{11,13},\,^{21\cdot28}$ In this study, these challenges were addressed by designing unique serotype-specific probes (Table-1) with the help of MUSCLE alignment software and added these probes to custom pneumococcal array. The modification resulted in accurate, identification of homologous serotypes among reference strains (N=32) and serum specimens (N=456) tested in the study. Of the 456 serum samples, Serotype 1 was the most predominant serotype followed by 6B, 14, 19F and 23F. (Figure-1). The Homologous serotypes 6A and 6B are most common and difficult to discriminate using DNA-based assays. Previously published nucleic acid serotyping methods failed to discriminate serotype 6A from 6B^{11,13,21-28}. In the current study, serotype 6A and 6B were distinguished by identifying a unique sequence of 23bp (GAACATGAAAGATGAAGCAATCAATGTACTGGATAAAGTTAT GAAAAAGA) which resulted in accurate identification of 6A (N=10) and 6B (N=43). The serotype 6B was predominant followed by 9V, 9N and 18C among all homologous serotypes observed in the study. (Figure-2). Serum samples belonging to homologous serogroups 9 (N=49), 10 (N=16), 11 (N=10), 12A/12F/44/46 (N=1), 15 (N=6), 18 (N=13), 22 (N=6), and 33A/33B/33F (N=3) were discriminated accurately to 9V (N=31), 9A (N=0), 9L (N=0), 9N (N=18), 10A (N=3), 10B (N=0), 10C (N=0) and 10F (N=13), 11A (N=10), 11B (N=0), 11C (N=0) and 11D (N=0), 12F (N=1), 15B (N=5), and 15C (N=1), 18C (N=13), 22F (N=6), 33B (N=2), and 33F (N=1) respectively. Our results are in concordance with the recent study based on PCRSeqtyping by Nagaraj et al.¹¹

Children are serially and simultaneously colonized in the nasopharynx by more than one S. pneumoniae serotypes. Studies about the impact of conjugate pneumococcal vaccines on NP colonization have shown a decrease in vaccine serotype (VT) colonization as well as an increase in nonvaccine serotype (NVT) colonization. The ability to distinguish and guantify between VT and NVT colonization is dependent on the sensitivity of the assay for detecting the carriage of multiple serotypes. Conventional quellung method does not reliably detect multiple serotypes if the second or third serotype is present in a much smaller proportion than the dominant type and is not practical for large pneumococcal vaccine field trials. Other methods, such as Immunoblot (IB) and mouse inoculation, are ooperator dependent, time-consuming, requires a culture step and is expensive. ^{29,30,31,32,33} In the present study, the redesigned custom pneumococcal array not only detected multiple serotypes from spiked samples and other bacterial pathogens (with 1:1 ratio) but also quantified relative abundance levels of each serotype in pooled samples (Table-3). However, future studies should explore other ratios perhaps 9:1 ratio to know the minor serotype and evaluate complex samples such as nasopharyngeal swabs and sputum. Studies from other groups have developed a low-density microarray which could detect limited numbers of bacterial pathogens.^{21,22}

The current method is much more affordable, in terms of perisolate or per-sample costs as two-colour hybridization approach was adopted for greater throughput (16 samples in single 8*15K array) than existing microarray methods.^{21,22} The Microarray assay described herein is a reliable assay for the identification, serotyping, multiple serotype detection and relative abundance quantification of 90 *S. pneumoniae* serotypes from culture and serum specimens.

REFERENCES

- UNICEF. The state of Asia-Pacific's children 2008 child survival, 2015. Available at:http://www.unicef.org/sapc08/report/report.php. Accessed 27 June 2018.
- Vashishtha VM. Current status of tuberculosis and acute respiratory infections in India: much more needs to be done! Indian Pediatr. 2010; 47:88
 Million Death Study Collaborators, Bassani DG, Kumar R, Awasthi S, Morris SK,
- Million Death Study Collaborators, Bassani DG, Kumar R, Awasthi S, Morris SK, Paul VK, et al. Causes of neonatal and child mortality in India: nationally

- representative mortality survey. Lancet. 2010; 376:1853-60.
 Farooqui H, Jit M, Heymann D L and Zodpey S. Burden of Severe Pneumonia, Pneumococcal Pneumonia and Pneumonia Deaths in Indian States: Modelling Based Estimates. PLoS ONE. 2015; 10: e0129191.
 Slotved HC, Kaltoft M, Skovsted IC, Kerrn MB, Espersen F. Simple, rapid latex
- Slotved HC, Kaltoft M, Skovsted IC, Kerrn MB, Espersen F. Simple, rapid latex agglutination test for serotyping of Pneumococci (Pneumotest-Latex). J Clin Microbiol 2004;42:18–22.
- Pai R, Gertz RE, Beall B. Sequential multiplex PCR approach for determining capsular serotypes of Streptococcus pneumoniae isolates. J Clin Microbiol 2006;44:124–31.
- 2006;44:124–31.
 Batt SL, Charalambous BM, McHughes TD, Martin S, Gillespie SH. Novel PCRrestriction fragment length polymorphism method for determining serotypes or serogroups of Streptococcus pneumoniae isolates. J Clin Microbiol 2005;43:56–61.
- Saha SK, Darmstadt GL, Baqui AH, Hossain B, Islam M, Foster D, et al. Identification of serotype in culture negative pneumococcal meningitis using sequential multiplex PCR: implication for surveillance and vaccine design. PLoS One 2008;3:e3576.
- Ganaie FA, Govindan V, Ravi Kumar KL. Standardisation and evaluation of a quantitative multiplex real-time PCR assay for the rapid identification of streptococcus pneumoniae. Pneumonia 2015;6:1.
- Raymond F, Boucher N, Allary R, Robitaille L, Lefebvre B, Tremblay C, et al. Serotyping of Streptococcus pneumoniae based on capsular genes polymorphisms. PLoS One 2013;8:e76197.
- Scott JR, Hinds J, Gould KA, Millar EV, Reid R, Santosham M et al. Nontypeable pneumococcal isolates among Navajo and White Mountain Apache communities: are these really a cause of invasive disease? J Infect Dis 206:73–80.
- 12. A Gervaix, J Corbeil, F Raymond. A new serotyping method of S. Pneumoniae using an automated microarray-based assay. BMC Proc 2011;5(Suppl 6):028.
- Bentley SD, Aanensen DM, Mavroidi A et al. Genetic analysis of the capsular biosynthetic locus from all 90 pneumococcal serotypes. PLoS Genet. 2006; 2: e31.
 Song JY, Eun BW, Nahm MH. Diagnosis of pneumocaccal pneumonia: current
- pitfalls and the way forward. Infect Chemother 2013;45:351–66. 15. Auburn S. Camping S. Clark TG. Diimde AA. Zongo I. Pinches R. et al. An effective
- Auburn S, Campino S, Clark TG, Djimde AA, Zongo I, Pinches R, et al. An effective method to purify Plasmodium falciparum DNA directly from clinical blood samples for whole genome high-throughput sequencing. PLoS One 2011;6:e22213.
- Feehery GR, Yigit E, Oyola SO, Langhorst EW, Schmidt VT, Stewart FJ, et al. A method for selectively enriching microbial DNA from contaminating vertebrate host DNA. PLoS One 2013;8:e76096.
- Shaik AH, Govindan V, Nagraj G, Ravikumar KL. Development of a microarraybased method for simultaneous detection and serotyping of Streptococcus pneumoniaefrom culture negative serum samples. J Appl Biol Biotech. DOI: 10.7324/JABB.2019.70103
- Vandana G, Feroze AG, Geetha N, Avid H, Kumar KR. Pan India distribution of pneumococcal serotypes (PIDOPS) causing invasive pneumococcal disease and pneumonia in children between 6 weeks and 5 years and their antimicrobial resistance—Phase I. Pedia Infect Dis 2016;8:47–51.
- Nagaraj G, Ganaie F, Govindan V, Ravikumar KL. Development of PCRSeqTyping—a novel molecular assay for typing of Streptococcus pneumonia. Pneumonia 2017;9:8.
- Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 2004;32:1792–97.
 Tomita Y, Okamoto A, Yamada K, Yagi T, Hasegawa Y, Ohta M et al. A new
- Tomita Y, Okamoto A, Yamada K, Yagi T, Hasegawa Y, Ohta M et al. A new microarray system to detect Streptococcus pneumoniae serotypes. J Biomed Biotechnol. 2011:352736.
- Wang Q, Wang M, Kong F, Gilbert GL, Cao B, Wang L. Development of a DNA microarray to identify the Streptococcus pneumoniae serotypes contained in the 23-valent pneumococcal polysaccharide vaccine and closely related serotypes. J Microbiol Methods. 2007;68:128–36.
- Batt SL, Charalambous BM, McHugh TD, Martin S, Gillespie SH. Novel PCRrestriction fragment length polymorphism method for determining serotypes or serogroups of Streptococcus pneumoniae isolates. J Clin Microbiol. 2005; 43:2656–66.
- Kong F, Gilbert GL. Using cpsA-cpsB sequence polymorphisms and serotype-/group-specific PCR to predict 51 Streptococcus pneumoniae capsular serotypes. J Med Microbiol. 2003; 52:1047–58.
- Kong F, Wang W, Tao J, Wang L, Wang Q, Sabananthan A et al. A molecularcapsular-type prediction system for 90 Streptococcus pneumoniae serotypes using partial cpsA-cpsB sequencing and wzy- or wzx-specific PCR. J Med Microbiol. 2005; 54:351–6.
- Pai R, Gertz RE, Beall B. Sequential multiplex PCR approach for determining capsular serotypes of Streptococcus pneumoniae isolates. JClin Microbiol.2006; 44:124–31.
- Pimenta FC, Roundtree A, Soysal A, Bakir M, du Plessis M, Wolter N et al. Sequential triplex real-time PCR assay for detecting 21 pneumococcal capsular serotypesthat account for a high global disease burden. J Clim Microbiol. 2013;51:647–52.
 Bello Gonzalez T, Rivera-Olivero IA, Sisco MC, Spadola E, Hermans PW, de Waard
- Bello Gonzalez T, Rivera-Olivero IA, Sisco MC, Spadola E, Hermans PW, de Waard JH. PCR deduction of invasive and colonizing pneumococcal serotypes from Venezuela: a critical appraisal. J Infect Dev Crites. 2014; 8:469–73.
 Melinda A, O'Brien K, Facklam R, Cynthia G, Schwartz B, George M. Immunoblot
- Melinda A, O'Brien K, Facklam R, Cynthia G, Schwartz B, George M. Immunoblot Method To Detect Streptococcus pneumoniae and Identify Multiple Serotypes from Nasopharyngeal Secretions. J.Clin. Microbiol. 2004, 42:1596.
 Mulholland K, Levine O, Nohynek H, Greenwood BM. Evaluation of vaccines for the
- Mulholland K, Levine O, Nohynek H, Greenwood BM. Evaluation of vaccines for the prevention of pneumonia in children in developing countries. Epidemiol. Rev. 1999;21:43-55.
- Huebner RE, Dagan R, Porath N, Wasas AD, Klugman KP. Lack of utility of serotyping multiple colonies for detection of simultaneous nasopharyngeal carriage of different pneumococcal serotypes. Pediatr. Infect. Dis. J. 2000;19:1017-20.
- Turner P, Hinds J, Turner C, Jankhot A, Gould K, Bentley SD et al. Improved Detection of Nasopharyngeal Cocolonization by Multiple Pneumococcal Serotypes by Use of Latex Agglutination or Molecular Serotyping by Microarray. J Clin Microbiol. 2011;49:1784–9.
- Kandasamy R, Gurung M, Thapa A, Ndimah S, Adhikari N, Murdoch DR et al. Multi-Serotype Pneumococcal Nasopharyngeal Carriage Prevalence in Vaccine Nadve Nepalese Children, Assessed Using Molecular Serotyping. PLoS One. 2015;10:e0114286