

# Indian population genomic datasets describing the pre-vaccine evolutionary epidemiology of *Streptococcus pneumoniae* using a Whole Genome Sequencing approach

Geetha N<sup>1</sup>, Vandana G<sup>1</sup>, Avid H<sup>1</sup>, Feroze G<sup>1</sup>, Ravikumar KL<sup>1</sup>, Stephen D. Bentley<sup>2</sup>, Rebecca A. Gladstone<sup>2</sup>, Stephanie W. Lo<sup>2</sup>, Robert F. Breiman<sup>3</sup>, Lesley McGee<sup>4</sup> and the Global Pneumococcal Sequencing Consortium

<sup>1</sup>Central Research laboratory, KIMS H&RC, Bangalore; <sup>2</sup>Infection Genomics, The Wellcome Trust Sanger Institute, <sup>3</sup>Emory University, Atlanta, USA, <sup>4</sup>Respiratory Diseases Branch, Centers for Disease Control and Prevention, Atlanta, USA

## BACKGROUND AND AIMS

Pneumococcal infections cause a high death toll in India. The recently rolled out (May, 2017) PCV in the national infant immunization schedule is expected to reduce disease burden. Here we present a population genomic study of the *S. pneumoniae* strains collected across India before vaccine introduction

## METHODS

- i. 386 isolates of *S. pneumoniae* strains collected across India from 2009-2017 were sequenced
- ii. Isolates were sequenced on Illumina HiSeq 2500 platform as a part of GPS study
- iii. The Sanger Institute's bioinformatics pipeline was used to determine population genetic structure, phylogenomics, serotype prevalence and resistotypes

## RESULTS

- i. Population genetic structure analysis reveals high level of heterogeneity and 22 genetically distinct sequence clusters.
- ii. 130 known and 97 novel STs. 46 isolates belonging to 11 (1, 2, 3, 14, 15, 19, 25, 26, 27, 32 and 38) PMEN clones were detected.

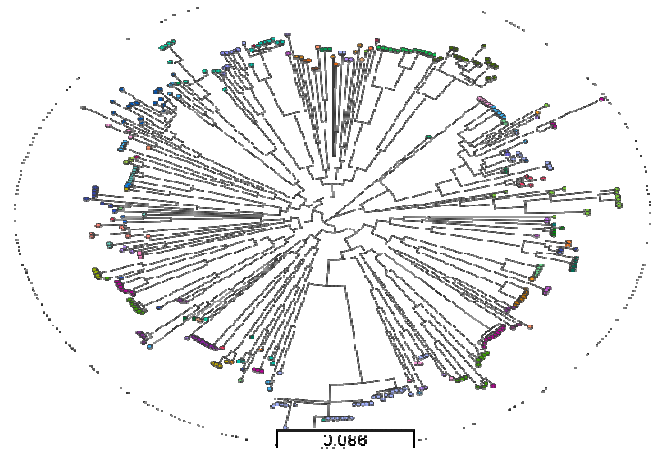


Fig 1: Population Snapshot of the isolates

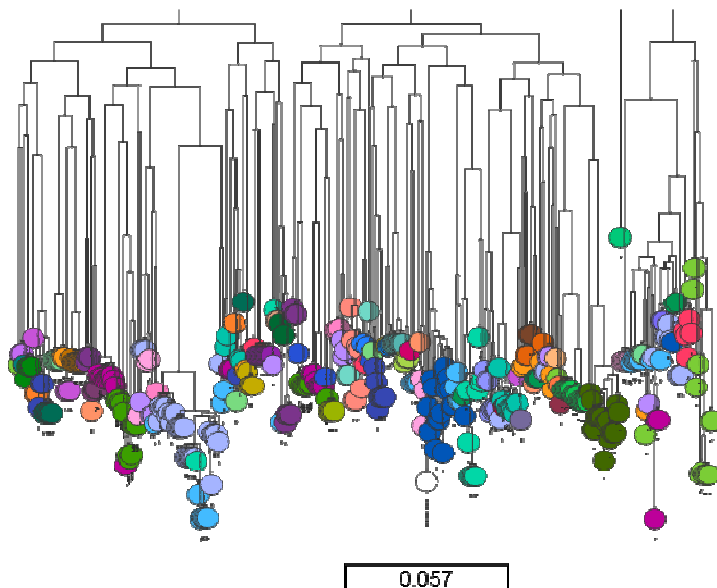


Fig 2: Population Snapshot of Serotype distribution of the isolates

- iii. Serotype 19F was the most common serotype, followed by 19A, 6A, 6B, 14 and 1

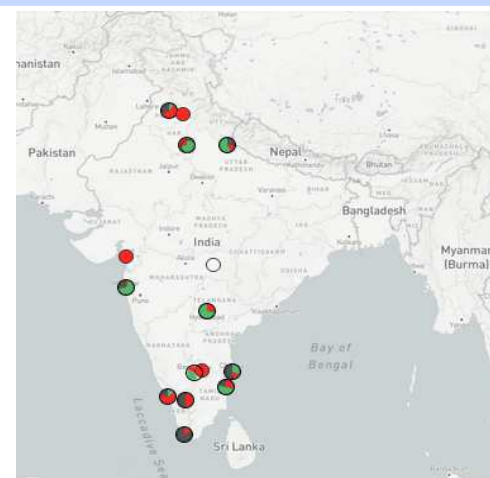


Fig 3: Geographic distribution of Penicillin resistance across India

- iv. 72 (18.7%) isolates were multi drug resistant.
- v. Resistance to Penicillin was observed in 12% of isolates

## CONCLUSION:

High throughput genome sequencing in this study elucidates STs, genome evolution, and population level characterization of circulating pneumococci in India during the pre-vaccine era

