

INSIGHTS INTO POTENTIAL PCV EFFECTS ON PNEUMOCOCCAL LINEAGES USING WHOLE GENOME SEQUENCING.

Rama Kandasamy^{1,2}, Rebecca Gladstone³, John Lees³, Lesley McGee⁴, Robert F. Breiman⁵, Keith P. Klugman⁶, Jukka Corander^{3,7,8}, Andrew J. Pollard^{1,2}, Stephen D. Bentley³
¹Oxford Vaccine Group, Department of Paediatrics, University of Oxford, Oxford, OX3 7LE, United Kingdom. ²NIHR Oxford Biomedical Research Centre, Oxford, OX3 7LE, United Kingdom. ³Pathogen Genomics, Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, United Kingdom. ⁴Respiratory Diseases Branch, Centers for Disease Control and Prevention, Atlanta Georgia, United States of America. ⁵Emory Global Health Institute, Emory University, Atlanta, Georgia, United States of America. ⁶Rollins School of Public Health, Emory University, USA. ⁷Department of Biostatistics, University of Oslo, Oslo, Norway. ⁸Infection Genomics. ⁸Helsinki Institute for Information Technology, Department of Mathematics and Statistics, University of Helsinki, Helsinki, Finland.
 On behalf of all partners of the Pneumococcal Sequencing Consortium, <http://www.pneumogen.net/gps/>

BACKGROUND

Using the Global Pneumococcal Sequencing Project database of pneumococcal whole genome sequences collected prior to any pneumococcal conjugate vaccine (PCV) introduction we aimed to identify how pneumococcal lineages might be affected by PCV introduction.

METHODS

DNA from 7596 pneumococcal isolates, comprising 2512 carriage, 4815 invasive, and 269 isolates from other sources, underwent whole-genome-sequencing on the Wellcome Trust Sanger Institutes core sequencing pipeline.

Lineages were defined by clustering genomes on sequence similarity using hierBAPS. Simpson's index of serotype diversity (a value between 0 and 1) for each lineage was calculated, with higher values indicating greater diversity.

RESULTS

45 pneumococcal lineages were identified.

Two monophyletic lineages which were represented entirely by serotype 1 and serotype 5 isolates respectively (Simpson's index of 0) were observed, which would thus be unlikely to escape negative selection following vaccine introduction.

Contrastingly however, three lineages which were comprised of 65/324 (20.1%), 54/157 (34.4%), and 51/183 (27.9%) serotypes included in the PCV13 vaccine, and had Simpson's indices of 0.83, 0.88, and 0.92 respectively, may be more likely to escape selection as a result of PCV introduction.

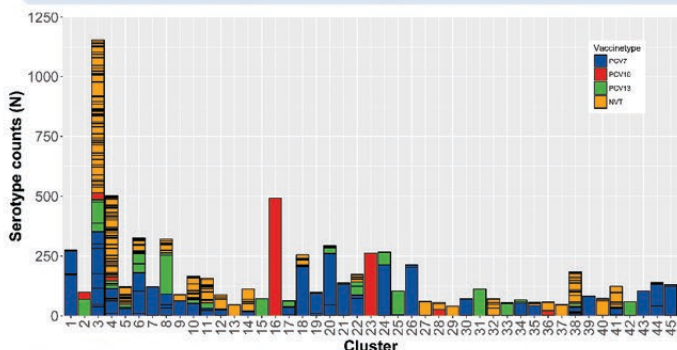


Figure 1. Proportion of each global pre-PCV pneumococcal population clusters covered by PCV7 (serotypes 4, 6B, 9V, 14, 18C, 19F, and 23F), the additional PCV10 (serotypes 1, 5, and 7F), and the additional PCV13 serotypes (serotypes 3, 6B, and 19A). 7596 pneumococcal isolates underwent whole-genome sequencing, were mapped to the *Streptococcus pneumoniae* ATCC 700669 genome, and SNPs identified. SNPs across whole-genomes were then used to inform population clusters using the hierBAPS clustering tool.

Table1. Characteristics of each identified population cluster.

Primary cluster	Number of isolates	Number of secondary clusters	Number of different STs	Number of different serotypes	Simpson's index of serotype diversity
1	276	3	34	6	0.50
2	99	7	24	2	0.42
3	1155	15	508	72	0.96
4	502	13	215	59	0.96
5	122	6	34	17	0.85
6	325	8	87	24	0.83
7	121	6	21	5	0.08
8	320	9	87	14	0.7
9	89	6	24	4	0.44
10	166	6	35	17	0.84
11	157	6	37	15	0.88
12	88	6	26	8	0.68
13	46	6	16	1	0
14	112	5	28	9	0.75
15	71	4	6	1	0
16	492	8	34	1	0
17	64	4	16	5	0.62
18	255	4	13	54	0.34
19	99	3	22	6	0.29
20	293	8	100	8	0.43
21	137	6	44	4	0.08
22	174	5	44	12	0.76
23	261	6	12	1	0
24	266	7	59	5	0.33
25	104	5	20	3	0.09
26	213	7	29	4	0.09
27	60	6	15	3	0.1
28	55	5	8	4	0.58
29	41	3	8	1	0
30	71	6	18	3	0.06
31	112	7	12	2	0.02
32	71	5	19	7	0.73
33	55	1	11	4	0.17
34	66	2	10	4	0.38
35	57	2	22	6	0.43
36	58	1	6	4	0.52
37	48	1	4	2	0.12
38	183	1	60	27	0.92
39	83	1	6	2	0.02
40	73	2	12	7	0.27
41	124	1	40	11	0.8
42	59	1	7	2	0.03
43	103	1	12	1	0
44	140	1	15	7	0.49
45	130	1	24	4	0.12

CONCLUSIONS

Using whole-genome sequencing of pneumococci collected prior to PCV introduction, it is apparent that a number of lineages which contain PCV serotypes also contain non-vaccine types and may be more capable of progressively escaping selective pressure of PCV introduction in settings where these strains exist.

Further correlation of these data with post-PCV pneumococcal populations may be used to further inform PCV effect on pneumococcal lineages.

The findings and conclusions in this report are those of the authors and do not necessarily represent the official position of the Centers for Disease Control and Prevention.