



# Genetic stability of pneumococcal isolates during 35 days of human experimental carriage



R.A. Gladstone<sup>a,1</sup>, J.F. Gritzfeld<sup>b,1</sup>, P. Coupland<sup>c</sup>, S.B. Gordon<sup>b</sup>, S.D. Bentley<sup>a,\*</sup>

<sup>a</sup> Pathogen Genomics, Wellcome Trust Sanger Institute, Cambridge, UK

<sup>b</sup> Respiratory Infection Group, Liverpool School of Tropical Medicine, Liverpool, UK

<sup>c</sup> Sequencing R&D Group, Wellcome Trust Sanger Institute, Cambridge, UK

## ARTICLE INFO

### Article history:

Received 22 October 2014

Received in revised form 31 March 2015

Accepted 11 May 2015

Available online 23 May 2015

### Keywords:

*Streptococcus pneumoniae*

Pneumococcal vaccines

Vaccine efficacy against carriage (VEcol)

Experimental human pneumococcal carriage

## ABSTRACT

**Background:** Pneumococcal carriage is a reservoir for transmission and a precursor to pneumococcal disease. The experimental human pneumococcal carriage model provides a useful tool to aid vaccine licensure through the measurement of vaccine efficacy against carriage (VEcol). Documentation of the genetic stability of the experimental human pneumococcal carriage model is important to further strengthen confidence in its safety and conclusions, enabling it to further facilitate vaccine licensure through providing evidence of VEcol.

**Methods:** 229 isolates were sequenced from 10 volunteers in whom experimental human pneumococcal carriage was established, sampled over a period of 35 days. Multiple isolates from within a single volunteer at a single time provided a deep resolution for detecting variation. HiSeq data from the isolates were mapped against a PacBio reference of the inoculum to call variable sites.

**Results:** The observed variation between experimental carriage isolates was minimal with the maximum SNP distance between any isolate and the reference being 3 SNPs.

**Conclusion:** The low-level variation described provides evidence for the stability of the experimental human pneumococcal carriage model over 35 days, which can be reliably and confidently used to measure VEcol and aid future progression of pneumococcal vaccination.

© 2015 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

## 1. Introduction

Pneumococcal conjugate vaccines (PCV) have been introduced across the globe since 2000 to combat pneumococcal disease. The indirect effect of pneumococcal vaccination, through reduction of vaccine-type carriage, has been a key contributor to conjugate vaccine cost effectiveness [1–3]. However, carriage, which is necessary for both person-to-person transmission and spread within the body, is not part of the current pneumococcal vaccine licensure process [4]. Therefore, the impact of vaccination on nasopharyngeal carriage is of growing interest and has been presented as a vaccine implementation and licensing aid [5]. Experimental human pneumococcal carriage offers a model to measure vaccine efficacy against carriage (VEcol) in a small, controlled population where carriage can be monitored weekly, either as presence/absence or by density and duration [6]. The model involves inoculating the

nasopharynx of healthy adult volunteers with a known quantity of serotype 6B. Serotype 6B was selected for the experimental human pneumococcal carriage model as it was a PCV vaccine type, had an odds ratio of <1 for assessing invasive disease potential and had previously been used to investigate the immunizing effect of a carriage episode [7–10]. It is important to understand the genomic stability of the pneumococcal inoculum to further assess the safety of the experimental carriage model and consider any potential impact this could have on measuring VEcol. To this end, 229 experimental carriage isolates recovered from multiple volunteers and time-points, were sequenced to document the number of variable sites.

## 2. Methods

### 2.1. Recruitment and ethics

Healthy volunteers were enrolled with informed consent to an Experimental Human Pneumococcal Carriage trial [7]. All participants were non-smoking adults aged 18–60 who had no close contact with at risk individuals, including young children and the elderly. Ethical approval was obtained from the National Health

\* Corresponding author. Tel.: +44 1223 494942.

E-mail address: [sdb@sanger.ac.uk](mailto:sdb@sanger.ac.uk) (S.D. Bentley).

<sup>1</sup> These authors contributed equally.

**Table 1**  
Number of isolates per volunteer and time-point.

Volunteer	Days since inoculation (SNPs observed)						Total
	2	7	14	21	28	35	
362/172	10	10	10	10	7		47
382/193	8	8	8	7	8		39
381/192	8		2	8	8		26
391/203	8	8	8	1	1		26
368/178		2	10		3	8	23
366/176	1	10*	10				21
364/174	10	10					20
380/191		8	8				16
384/196	1			8			9
376/186	2						2

\*collected on day 8

\* Collected on day 8.

Service Research Ethics Committee (11/NW/0592) and the study was sponsored by the Royal Liverpool and Broadgreen University Hospitals Trust.

## 2.2. Experimental carriage

Pneumococcal stock preparation, inoculation and nasal wash were done as previously described [7]. Nasal washes were performed one week prior to pneumococcal challenge to determine natural carriage status. Twenty-four volunteers that were negative for natural pneumococcal carriage were inoculated with a serotype 6B strain, multi locus sequence type ST138, and nasal washes were taken on days 2, 7, 14, 21, 28 and 35 post-inoculation [11]. Carriage positive volunteers  $n = 10$  received amoxicillin 500 mg T.D.S. for 3 days after day 28 in order to ensure 6B colonisation clearance. One volunteer (368/178, Table 1) did not have antibiotics until after day 35.

In carriage positive volunteers, up to ten individual pneumococcal colonies (where available) were taken from a blood agar plate following overnight incubation, added to 1 ml of Todd Hewitt broth containing 0.5% yeast extract and incubated at 37 °C, 5% CO<sub>2</sub> for 7 h. The broth was then centrifuged at 17,000 × *g* for 3 min and the pellet was resuspended in 200 μl sterile PBS and 40 μl of 10 mg/ml of RNase A (Qiagen) and stored at –20 °C until extraction. DNA was extracted from the isolates using a modification of the QIAamp DNA Mini Kit (Qiagen). Briefly, the sample was added to 100 μl of TE buffer containing 0.04 g/ml lysozyme and 75 U/ml mutanolysin (Sigma) and incubated for 1 h at 37 °C. Following incubation, proteinase K was added and the sample was incubated for 30 min at 56 °C. Next, Buffer AL was added, the sample was incubated at room temperature for 10 min and then 260 μl of ethanol was added. All subsequent steps were as outlined in the kit protocol.

## 2.3. Sequencing

Sequencing was performed at the Wellcome Trust Sanger Institute, on the Illumina HiSeq2000 platform. The inoculum was

sequenced using the PacBio platform to provide a reference genome. Pacbio sequencing of the inoculum provided 335,112 filtered sub-reads. De novo assembly was performed with the Hierarchical Genome Assembly Process HGAP.2, using PacBio's smrtanalysis version 2.1. The assembly was blasted against itself and inspected in Artemis Comparison Tool (ACT), 3 small spurious contigs covered elsewhere in the assembly were removed and the reads mapped back against the new assembly to create a final consensus fasta sequence using the PacBio SMRT Portal version v2.2.0 build 133,335. The resultant assembly had 6 contigs with a N50 of 1.15 Mb (accession CTRR01000001–CTRR01000006). Supplementary Table 1 includes the ENA accession numbers for the sequence data.

Illumina reads were mapped to the PacBio assembly of the inoculum to identify single nucleotide polymorphisms (SNPs) using SMALT [12] with GATK indel realignment [13] and filtered with a minimum base call quality of 50 and minimum mapping quality of 30 as previously described [14]. The SNP calling was replicated with reduced minimum base call quality of 30 and minimum mapping quality of 20. This was to assess the potential numbers of false negatives and indicate whether variation was missed used the stringent SNP calling, both approaches gave identical results. SNP alignments were used to construct a maximum likelihood phylogenetic tree with 100 bootstrap replicates of RAxML v7.0.4 [15] and combined with temporal data to detect root to tip divergence in Path-O-Gen [16].

## 3. Results

Of the 24 challenged volunteers, 10 were carriage positive for serotype 6B ST138. From these 10 volunteers, 229 isolates from multiple time-points were sequenced (Table 1).

### 3.1. Sequence variation

A total of 37 positions in the reference sequence contained a SNP in at least one isolate (variable sites) and 232 SNPs were identified within the 229 isolates. No SNPs were detected for 59/229 isolates. The greatest number of SNP differences from the reference sequence for a single isolate was 3 ( $n = 8/229$ ). No more than 11 variable sites were observed for all isolates taken from a single volunteer. Seven or less variable sites were observed within any volunteers time-point. Between 1 and 3 variable sites could be observed at the primary time-point of 2 days for all volunteers with the exception of 376/186. For 376/186 only two isolates identical to the reference could be retrieved at day 2 and no isolates at later time-points. Furthermore the variation observed within the most comprehensive set of isolates from volunteer 362/172 was only detected on day 2. All remaining 362/172 isolates ( $n = 37$ ) were identical to the reference sequence.

One shared SNP was detected in multiple isolates from two separate volunteers 380/191 (16/16 isolates) and 384/196 (4/9 isolates). The gene annotation for this SNP site in the reference was DNA replication initiation control protein YabA. The C to A transversion resulted in a non-synonymous change from Threonine to Asparagine.

Horizontal exchange of genetic material (homologous recombination), introducing multiple SNPs within a localised area of the genome was not detected in this dataset with an average distance of 55,694 base pairs (bp) and a minimum of 1319 bp between SNPs. Full SNP profiles can be found in Supplementary Table 1.

The correlation between sampling time and different genetic distances from the root to the tips of the phylogenetic tree was assessed using Path-O-Gen root to tip phylogenetic divergence. No

trend for increasing diversity over time was observed for the whole dataset ( $n = 229$ ) or for each volunteer's isolates (data not shown).

#### 4. Discussion

Mutation and recombination contribute independently to evolution of the pneumococcal genome. The mutation rate has been previously reported to equate to  $\sim 3.14$  SNPs per pneumococcal genome (2 megabases) each year or 0.26 per month, for the Pneumococcal Molecular Epidemiology Network clone 1 (PMEN1) and has been found to be consistent across the species [17,18]. This has subsequently been reported to be representative of a number of pneumococcal lineages that include serotype 6B [17]. However, an estimated 88% of SNPs in the PMEN1 lineage were due to recombination events [18]. Micro and macro recombination events have each been estimated to only occur once every 17 years [19], rare events outside the timescale of experimental human pneumococcal carriage. In addition multiple serotype colonisation, which could provide the opportunity for pneumococcal recombination events was not detected in the volunteers. It is therefore not unexpected that recombination events were not detected in this study.

A report of the number of SNPs observed between seven independent natural pneumococcal carriage isolates sharing an ST that also belonged to serogroup 6 (6C) was previously observed to range from 67 through to 128 SNPs [20]. This is considerably more variation than that observed in this experimental carriage, especially considering the smaller sample size and aggressive filtering of SNPs in the 6C study. Two isolates from the 6C study that were collected from siblings and likely represent a transmission event were separated by only 8 SNPs. This is more comparable to the number of variable sites seen between individuals in experimental carriage but multiple sampling from an individual may have revealed further within host diversity. Calculated mutation rates are a description of the number of changes that occur and go to fixation and remain in the population, but many more mutations may occur that could be observed by deep sampling within a host that do not become fixed. One study of within host diversity for natural *Staphylococcus aureus* carriage reported between six and 40 SNPs between isolates from a host at a single time-point [21]. Low-level sequence variation within the host for pneumococcal carriage would likely also be a common observation if deep sampling of hosts was performed routinely.

#### 5. Limitations

Whilst the number of volunteers closely followed was small, the depth of multiple isolate sequencing provided an insight into the variation within a volunteer at a single time-point and over time. However, we were not able to retrieve equal numbers of pneumococcal colonies from all individuals, which may have been a factor of bacterial load, and therefore may have missed further within host variation. It was important to look at isolate sequence variation in multiple hosts to assess any potential differences between hosts but this is limited by the isolates available for some volunteers.

The low level sequence variation observed, the presence of variants at the time-point 2 days after inoculation, and the observation of a shared variable site in isolates from two volunteers could represent low-level un-sampled diversity contained within the original inoculum. The variation observed at the 2-day time-point for volunteer 362/172 could not be observed in all later time-points and may even represent the diversity stabilising over time or a bottleneck of the carriage state.

The timescale of this study is short however to examine the impact of a vaccine on carriage acquisition, carriage only needs to be detected once, right after challenge.

#### 6. Conclusion

In this sequencing study of serial isolates collected from an experimental human pneumococcal carriage model we found low rates of sequence variation between the isolates retrieved and the original inoculum. Recombination events were not detected and there was no indication for evolution within the host with no observed increasing diversity over time.

Carriage data could be used to measure the long-term impact of vaccination and estimate post-vaccine changes in invasive disease incidence [22]. Because carriage is relatively easy to measure and is more common than disease, it would require a much smaller sample size in a vaccine trial, making it a very useful endpoint. The low-level sequence variation observed within experimental carriage over 35 days provides a stable model that can be reliably used to aid future progression of pneumococcal vaccination.

#### Conflict of interest statement

The authors declare they have no conflicts of interest directly pertaining to this manuscript. Commercial links are as follows: SG laboratory has received consultancy fees from Pfizer, and research funding from GSK and Grifols. RG received PhD studentship for Pfizer. RG employed for one year on a GSK funded research project in 2012. SB has consulted for Merck.

#### Acknowledgements

The authors would like to thank Prof. Peter Hermans (Radboud University Medical Centre) for donating the 6B pneumococcal strain originally from Prof. Birgitta Henriques-Normark (Karolinska Institutet), and Dr. Daniela M. Ferreira for helpful advice and discussion. They also thank the Clinical Research team: Dr. Andrea M. Collins, Angela D. Wright (Research Nurse, LCRN Local Comprehensive Research Network), Carole A. Hancock (Research Nurse), and David Shaw (Research Nurse), the Clinical Research Unit at the Royal Liverpool University Hospital, and all the volunteers who participated in the study. The study was co-sponsored by the Royal Liverpool and Broadgreen University Hospitals NHS trust and the Liverpool School of Tropical Medicine. The EHPC study was funded by the Bill and Melinda Gates Foundation (OPP1035281). Sequencing was funded by The Wellcome Trust Sanger Institute, (098051).

#### Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.vaccine.2015.05.021>

#### References

- [1] Tocheva AS, Jefferies JM, Rubery H, Bennett J, Afimeke G, Garland J, et al. Declining serotype coverage of new pneumococcal conjugate vaccines relating to the carriage of *Streptococcus pneumoniae* in young children. *Vaccine* 2011;29:4400–4.
- [2] Melegaro A, Edmunds WJ. Cost-effectiveness analysis of pneumococcal conjugate vaccination in England and Wales. *Vaccine* 2004;22:4203–14.
- [3] Miller E, Andrews NJ, Waight PA, Slack MP, George RC. Herd immunity and serotype replacement 4 years after seven-valent pneumococcal conjugate vaccination in England and Wales: an observational cohort study. *Lancet Infect Dis* 2011;11:760–8.
- [4] Simell B, Auranen K, Kayhty H, Goldblatt D, Dagan R, O'Brien KL. The fundamental link between pneumococcal carriage and disease. *Expert Rev Vaccines* 2012;11:841–55.
- [5] Goldblatt D, Ramakrishnan M, O'Brien K. Using the impact of pneumococcal vaccines on nasopharyngeal carriage to aid licensing and vaccine implementation: a PneumoCarr meeting report March 27–28, 2012, Geneva. *Vaccine* 2013;32:146–52.
- [6] Gritzfeld JF, Cremers AJ, Ferwerda G, Ferreira DM, Kadioglu A, Hermans PW, et al. Density and duration of experimental human pneumococcal carriage. *Clin Microbiol Infect* 2014;20(12):O1145–51 (The Official Publication of the European Society of Clinical Microbiology and Infectious Diseases).

- [7] Gritzfeld JF, Wright AD, Collins AM, Pennington SH, Wright AK, Kadioglu A, et al. Experimental human pneumococcal carriage. *J Visualized Exp: JoVE* 2013. Issue 72.
- [8] Ferreira DM, Neill DR, Bangert M, Gritzfeld JF, Green N, Wright AK, et al. Controlled human infection and rechallenge with *Streptococcus pneumoniae* reveals the protective efficacy of carriage in healthy adults. *Am J Respir Crit Care Med* 2013;187:855–64.
- [9] Brueggemann AB, Griffiths DT, Meats E, Peto TE, 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;187:1424–32.
- [10] McCool TL, Cate TR, Moy G, Weiser JN. The immune response to pneumococcal proteins during experimental human carriage. *J Exp Med* 2002;195:359–65.
- [11] Browall S, Norman M, Tangrot J, Galanis I, Sjöström K, Dagerhamn J, et al. Intracolonial variations among *Streptococcus pneumoniae* isolates influence the likelihood of invasive disease in children. *J Infect Dis* 2014;209:377–88.
- [12] Wellcome Trust Sanger Institute. SMALT; 2014. (<http://www.sanger.ac.uk/resources/software/smalt/>) (accessed 01/07/2014).
- [13] DePristo MA, Banks E, Poplin R, Garimella KV, Maguire JR, Hartl C, et al. A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat Genet* 2011;43:491–8.
- [14] Harris SR, Feil EJ, Holden MT, Quail MA, Nickerson EK, Chantratita N, et al. Evolution of MRSA during hospital transmission and intercontinental spread. *Science* 2010;327:469–74.
- [15] Stamatakis A. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* 2006;22:2688–90.
- [16] Molecular evolution phylogenetics and epidemiology. Path-O-Gen v1.4. (<http://tree.bio.ed.ac.uk/software/pathogen/>) (accessed 01/07/2014).
- [17] Croucher NJ, Finkelstein JA, Pelton SI, Mitchell PK, Lee GM, Parkhill J, et al. Population genomics of post-vaccine changes in pneumococcal epidemiology. *Nat Genet* 2013;45:656–63.
- [18] Croucher NJ, Harris SR, Fraser C, Quail MA, Burton J, van der Linden M, et al. Rapid pneumococcal evolution in response to clinical interventions. *Science* 2011;331:430–4.
- [19] Mostowy R, Croucher NJ, Hanage WP, Harris SR, Bentley S, Fraser C. Heterogeneity in the frequency and characteristics of homologous recombination in pneumococcal evolution. *PLoS Genet* 2014;10:e1004300.
- [20] Loman NJ, Gladstone RA, Constantinidou C, Tocheva AS, Jefferies JM, Faust SN, et al. Clonal expansion within pneumococcal serotype 6C after use of seven-valent vaccine. *PLoS ONE* 2013;8:e64731.
- [21] Golubchik T, Batty EM, Miller RR, Farr H, Young BC, Larner-Svensson H, et al. Within-host evolution of *Staphylococcus aureus* during asymptomatic carriage. *PLoS ONE* 2013;8:e61319.
- [22] Weinberger DM, Bruden DT, Grant LR, Lipsitch M, O'Brien KL, Pelton SI, et al. Using pneumococcal carriage data to monitor postvaccination changes in invasive disease. *Am J Epidemiol* 2013;178:1488–95.