Prevalence and genetic diversity of pneumococcal serogroup 6 in Australia

F. Zhuo[1,2,†,‡], M. Xiao[3,‡], F. Kong[2], S. Oftadeh[2], F. Zhou[2], J. Zhang[1] and G. L. Gilbert[2]

1) *Department of Dermatology, Peking University People's Hospital, Beijing, China,* 2) *Centre for Infectious Diseases and Microbiology (CIDM), Institute of Clinical Pathology and Medical Research (ICPMR), Westmead, NSW, Australia and* 3) *Life Science College, Peking University, Beijing, China*

Abstract

The prevalence of the newly discovered pneumococcal serotype 6C has increased in some countries since the introduction of seven-valent conjugate pneumococcal vaccine (PCV7). The distribution of invasive serogroup 6 serotypes, in Australia, including 6C and 6D, has not been reported previously. During the period 1999 to 2008, 6097 isolates were referred to the New South Wales Pneumococcal Reference Laboratory for serotyping. Of these, 847 were identified by Quellung reaction as belonging to serogroup 6 and 702 were available for further study. Serotypes were determined by serotype-specific PCR as follows: 6A, 197 (28.1%); 6B, 452 (64.4%); 6C, 52 (7.4%) and one 6D. The average numbers of invasive serogroup 6 isolates, per annum, fell from 62.2 before (2000–2005) to 49.7 after (2006–2008) the introduction of PCV7. The proportions of invasive 6B fell (from 72.4% to 47.3%, p 0.03), those of 6C rose (from 3.3% to 17%, p 0.02) significantly and those of 6A remained fairly constant (24.3% vs 27%, p 0.69) between the two periods. All 6C and 6D and selected 6A and 6B isolates were further characterized by multilocus sequence typing and sequence analysis of *cps* genes *cpsA-cpsB* (*wzg-wzh*) and *wchA-wciN*$_{beta}$*-wciO*, *wciP*. Results showed considerable diversity within serotype 6C, apparently as a result of both mutation and recombination. Sequence typing indicates that, in Australia, 6C has been largely derived from 6A. The genetic diversity and rapid increase in incidence of serotype 6C causing invasive pneumococcal disease has potential implications for vaccine efficacy.

Keywords: Antibiotic susceptibility, diversity, genotypes, serotypes 6C and 6D, *Streptococcus pneumoniae*
Original Submission: 3 May 2010; **Revised Submission:** 3 October 2010; **Accepted:** 6 October 2010
Editor: J.-L. Mainardi
Article published online: 15 October 2010
Clin Microbiol Infect 2011; **17:** 1246–1253
10.1111/j.1469-0691.2010.03404.x

Corresponding author: Gwendolyn L. Gilbert, Centre for Infectious Diseases and Microbiology (CIDM), Institute of Clinical Pathology and Medical Research (ICPMR), Westmead Hospital, Darcy Road, Westmead, NSW 2145, Australia
E-mail: lyn.gilbert@sydney.edu.au

†Present address: Department of Dermatology, Beijing Friendship Hospital, Capital Medical University, Beijing, China
‡These authors contributed equally to this article.

Introduction

The capsular polysaccharide of *Streptococcus pneumoniae* is its most important virulence factor [1] and serotyping has been used as the primary criterion for classification of pneumococci for decades [2]. A new serotype, 6C, was described in 2007 based on differences in reactions to monoclonal antibodies between two groups of isolates that had been identified by Quellung reaction as serotype 6A [3,4]. Several recent studies have shown that the proportion of serotype 6C isolates has increased while those of serotype 6B—and to a lesser extent 6A—have decreased since the introduction of the seven-valent pneumococcal conjugate vaccine (PCV7) [5–7]. The existence of a new serotype, 6D (identified as 6B by Quellung reaction, but carrying the variant *wciN* found in serotype 6C) was predicted [8], shown experimentally to be feasible [9] and, subsequently, putatively identified by us among isolates carried in the nasopharynx by Fijian children enrolled in an immunization study [10]. Bratcher *et al.* [11] later demonstrated the unique capsular and antigenic structure of two naturally occurring serotype 6D isolates from Korea.

The aim of this study was to determine the distribution and characteristics of *S. pneumoniae* serotypes 6C and 6D among isolates referred to the New South Wales (NSW) Pneumococcal Reference Laboratory between 1999 and 2008 and changes following the introduction of routine infant immunization with PCV7 in Australia, in 2005.

Materials and Methods

Streptococcus pneumoniae isolates

All isolates from patients with invasive pneumococcal disease in NSW and the Australian Capital Territory (combined population ~ 7.4 million) have been referred, routinely, for serotyping to the NSW Pneumococcal Reference Laboratory, Centre for Infectious Disease and Microbiology (CIDM), Westmead Hospital, since 1999 (although referral was incomplete early in this period). Non-sterile site isolates are referred sporadically. A total of 6097 invasive isolates were referred in the period 1999 to 2008. Of these, 679 (11.1%) were identified by Quellung reaction as belonging to serogroup 6 (6A, 202; 6B, 477); 135 were non-viable or unavailable, leaving 544 (80%) invasive isolates belonging to serogroup 6, of which 168 (83%) had been identified as 6A and 376 (79%) as 6B by Quellung reaction. All 168 available serogroup 6 isolates from non-sterile (mainly upper respiratory) sites (63 serotype 6A and 105 serotype 6B) were also studied, giving a total of 712 presumed serogroup 6 isolates (231 6A and 481 6B).

PCR amplification

All isolates were tested by serogroup 6 serotype-specific PCRs, which were performed as previously described [10,12]. Forty randomly selected serotype 6A, 27 randomly selected serotype 6B and all serotype 6C and 6D isolates were subjected to multilocus sequence typing (MLST) by a standard method [13] with modifications [14]. Three *cps* gene regions (*wzg-wzh*(*cpsA-cpsB*), *wchA-wciN*$_{beta}$*-wciO* and *wciP*) were amplified as previously described [12,15].

Sequencing and sequence analysis

The PCR products of seven alleles in MLST and three *cps* gene regions were purified and sequenced as described previously [12,15]. The MLST sequences were compared at https://www.mlst.net/. Sequences for three regions of *cps* genes were compared with GenBank sequences of serotype 6C and 6D *cps* loci (accession numbers EF538714 and EU714777, respectively) and our previously submitted sequences [10]. All sequences were analysed with the BLASTn tool in BioManager (http://biomanager.info/).

Antimicrobial susceptibility tests

Susceptibilities to chloramphenicol, erythromycin and tetracycline, of all serotype 6C and 6D isolates, were tested by the Kirby–Bauer technique and MICs of penicillin and ceftriaxone were determined by E-test (AB Biodisk, Solna, Sweden). All susceptibility tests were performed and interpreted according to CLSI recommendations and definitions [16]. To allow comparison with previous studies, penicillin MICs were interpreted according to the new oral penicillin breakpoints, established in 2008 (susceptible, ≤0.06 μg/mL; intermediate, 0.12–1 μg/mL; and resistant, ≥2 μg/mL).

GenBank submission

Fourteen new 6C and 6D nucleotide sequences were submitted to GenBank (Table 3).

Results

Serotype-specific PCR identification, distribution and sources of serogroup 6 isolates in NSW

Ten of 712 presumed serogroup 6 isolates were not amplified by the 6A/6B PCR; repeat Quellung testing confirmed that they did not belong to serogroup 6, so they were excluded. Serotype-specific PCR for the remaining 702 isolates showed that 197 belonged to serotype 6A (28.1%; 140 invasive, 57 from non-sterile sites); 452 to 6B (64.4%; 355 invasive); 52 to 6C (7.4%; 42 invasive) and one to 6D (isolated in 2008 from a blood culture of a 68-year-old man). Discrepant PCR and Quellung results for 26 isolates were resolved by retesting, which confirmed the PCR results; 22 isolates previously identified as 6B were reassigned to 6A and four previously identified as 6A were reassigned to 6B.

We divided 536 invasive serogroup 6 isolates into two groups according to the ages of patients from whom they were isolated: ≤5 years (the main high-risk group targeted by infant vaccine) and >5 years. In addition, we divided the total time period (excluding 1999, when referral of isolates was incomplete) into two time periods—prevaccine: 2000–2005 and postvaccine: 2006–2008. The annual average numbers of serogroup 6 isolates referred fell from 62.1 in 2000–2005 to 49.7 in 2006–2008, following the introduction of PCV7 in 2005. In both age groups, serotype 6B was dominant among serogroup 6 isolates in the first period, but the proportion fell significantly, especially in the younger age group, in 2006–2008 (Table 1). In both age groups the differences in proportions of 6B isolates between prevaccine and postvaccine periods (2000–2005 vs 2006–2008) were highly significant (age group ≤5 years: χ^2 = 15.36, df = 2, p 0.0005; age group ≥5 years: χ^2 = 50.43, df = 2, p <0.0005).

The proportions of serotype 6A among serogroup 6 isolates did not change significantly overall or in either age group between periods. However, the average annual numbers of 6A isolates referred fell between the prevaccine (2000–2005) and postvaccine (2006–2008) periods from 5.2

TABLE 1. Distribution of invasive serogroup 6 (SG-6) in New South Wales, Australia, by year and in different age groups

	Numbers of referred invasive isolates											
	All ages					Age ≤5 years[a]				Age >5 years[a]			
			Serotypes (% of SG-6)				Serotypes (% of total SG-6)				Serotypes (% of total SG-6)		
Year	Total isolates[b]	SG-6[c] (% of total)	6A	6B	6C	Total SG-6	6A	6B	6C	Total SG-6	6A	6B	6C
1999	95	14 (15)	2 (14)	12 (86)	0	7	1	6	0	7	1	6	0
2000	631	61 (10)	9 (15)	46 (75)	6 (10)	2000–2005							
2001	668	67 (10)	13 (19)	54 (81)	0	152	31 (20)	119 (78)	2 (1)	221	60 (27)	151 (68)	10 (5)
2002	781	61 (8)	15 (25)	46 (75)	0								
2003	729	50 (7)	9 (18)	40 (80)	1(2)								
2004	873	77 (9)	22 (29)	53 (69)	2 (3)								
2005	613	57 (9)	23 (40)	31 (54)	3 (5)								
2006	702	65 (9)	13 (20)	45 (69)	7 (10)	2006–2008							
2007	489	46 (9)	17 (37)	18 (39)	11(24)	19	6 (32)	7 (37)	6 (32)	130[d]	39 (30)	65(50)	25 (20)
2008	516	38[d] (7)	15 (39)	9 (24)	13 (34)								
Total	6097	536[d] (9)	138 (26)	354 (66)	43 (8)	178	38 (21)	132 (74)	8 (4)	358[d]	100 (28)	222 (62)	35 (10)

[a]Because of small numbers in each age group, years were grouped into 2-year periods (excluding 1999, when referral was incomplete).
[b]All *Streptococcus pneumoniae* invasive isolates referred for serotyping during the period 1999–2008.
[c]Invasive serogroup 6 isolates which were available for testing and confirmed as serotypes 6A, 6B, 6C or 6D (one isolate only) by PCR.
[d]Among 38 invasive serogroup 6 isolates collected in 2008 (130 in ≥5-year age group, 2006–2008), one from an adult was identified as serotype 6D.

to 2 in the vaccine target group whereas it rose from 10 to 13 in the >5-year age group. The absolute numbers and proportions of serotype 6C increased, overall and in both age groups, between time periods (Table 1). This difference was statistically significant in the ≥5 year age group (χ^2 = 19.98, df = 2, p <0.0005) but numbers of isolates in the <5 year age group were too small to demonstrate statistical significance.

Antibiotic susceptibility of invasive and non-invasive serotype 6C and 6D isolates

Two of the 52 (4%) serotype 6C isolates showed intermediate susceptibility to penicillin and one (2%) was resistant to erythromycin. The remaining 6C isolates and the one 6D isolate were susceptible to all five antibiotics tested.

MLST distribution among serotype 6A, 6B, 6C and 6D isolates in NSW

The distribution of MLSTs, for 120 serogroup 6 isolates, is shown in Table 2. There were 12 sequence types (STs), including five new ones (ST4237–4240 and ST4242) among 52 serotype 6C isolates; 18 STs, including seven new ones (ST4237, 4269–4274) among 40 serotype 6A isolates; 15 STs, including ten new ones (ST3999, ST4275–4283) among 27 serotype 6B isolates; the serotype 6D isolate belonged to new ST4241. All new ST sequences were submitted at http://www.mlst.net/. Based on a minimum similarity of five identical loci, eBURST analysis (Fig. 1) showed two clonal complexes (CCs, in which a predicted founder was identified), seven groups (in which there was no predicted founder) and 18 singletons with different STs (Table 2; Fig. 1).

Diversity of capsular genes for serotype 6C isolates in NSW

Sequence types based on variation in three different *cps* gene cluster regions are shown in Table 3. Among serotype 6C isolates, there were ten *wchA-wciN*$_{beta}$*-wciO* STs, one of which (type 10) we had identified previously among 6C isolates from Fiji [10]. Nine STs were distinguished by only one or two single nucleotide polymorphisms (SNPs) but type 10 differed by four base mutations. There were four *cpsA-cpsB* STs, including two new ones (types 2 and 3), which varied by multiple SNPs, and three *wciP* STs, including one new type (type 3), which differed by one or two SNPs. *wciP* type 1 (FJ899601) had been found previously among serotype 6C isolates from Fiji [10].

Comparison of MLST and three *cps* sequence types of serotype 6C isolates

In our previous study, we demonstrated sequence polymorphisms in the *cpsA-cpsB* region, among serogroup 6 isolates [15] and in the *wchA-wciN-wciO* region between serotypes 6C and 6D among isolates from Fiji [10]. For this reason, we used these regions to identify *cps* profiles (concatenated in the order *cpsA-cpsB*, *wchA-wciN-wciO*, *wciP*). Our *cps* profiles differ from those described by Mavroidi *et al.*, which refer to genes *wciP-wzx-wzy* [17]. Comparison of MLST and three *cps* STs among serotype 6C isolates is shown in Table 4. In all, the 52 serotype 6C isolates were divided into 17 *cps* profiles, based on all three *cps* gene regions, and 12 MLSTs. Generally, there was a fair correlation between *cps* profiles and MLST. However, 16 isolates belonging to ST1715 were divided among four *cps* STs and ten isolates belonging to a single *cps* profile (1-2-2) were divided among five STs (ST600, ST1379, ST1390, ST4238 and ST4242).

TABLE 2. Distribution of multilocus sequence types among 120 serotype 6A, 6B, 6C and 6D isolates in New South Wales, Australia

Clonal complexes/ groups (*n*)[a]	MLST allelic profile for 6A, 6B, 6C and 6D							Sequence types[d]	No. of isolates for serotype (*n*)				Serotype shared by sequence type[i]
	aroE	*gdh*	*gki*	*Recp*	*spi*	*xpt*	*ddl*		6A (40)	6B (27)	6C (52)	6D (1)	
gr1 (9)	7	13	9	6	10	6	14	4241[e]				1	6D
	7	13	8	10	10	6	14	4282[f]		1			6B
	11	13	8	6	10	6	14	4283[f]		1			6B
	7	13	8	6	10	246	14	4278[f]		2			6B
	7	97	27	6	10	6	14	1645		2			6B
	7	97	27	6	10	246	14	4277[f]		2			6B
gr2 (17)	2	40	4	19	10	1	27	2068	1				6A/10A[j]
	2	7	4	10	10	1	27	65	7				6A
	5	7	4	10	10	1	27	460	9				6A
CC1692 (7)	1	5	74	12	17	158	14	4239[g]			5		6C
	1	5	7	12	17	158	14	1692	1				6A
	1	5	7	12	17	1	14	395	1				6A
CC1092 (21)	2	13	2	1	9	19	14	4237[h,g]	1		14		6A/6C
	2	13	2	1	6	19	14	1092	1		2		6A/6B/6C/19A[l]
	2	13	9	1	6	19	14	490	3				6A
gr7 (6)	5	6	1	2	6	3	4	90		5			6B
	5	6	9	2	6	3	4	3999[f]		1			6B
gr6 (4)	7	25	4	4	15	20	28	473	3				6A/6B/6C
	8	25	4	4	15	20	28	4270[h]	1				6A
gr8 (5)	1	5	9	12	94	28	20	1379			4		6A/6C
	1	5	9	12	6	28	20	4238[g]			1		6C
gr5 (6)	33	1	40	13	6	1	5	4273[h]	5				6A
	10	1	40	13	6	1	5	4272[h]	1				6A
gr9 (3)	32	28	1	1	15	19	14	4280[f]		2			6B
	32	28	1	2	15	19	14	4281[f]		1			6B
Singleton STs (42)	1	25	72	1	15	20	28	690	1				6A
	8	8	19	16	6	1	14	4271[h]	1				6A
	2	8	4	131[b]	6	1	11	4274[h]	1				6A
	7	5	4	5	6	1	18	4269[h]	1				6A
	4	4	2	4	4	1	15	3604	1				6A/23F[k]
	7	6	1	1	6	15	14	145		3			6B
	7	22	1	2	5	1	14	185		3			6B
	7	6	4	2	10	15	14	1518		1			6B
	6	28	1	2	6	1	5	4275[f]		1			6B
	7	62	1	2	6	1	31	4276[f]		1			6B
	7	5	8	5	10	246	6	4279[f]		1			6B
	1	5	9	43	94	1	346[c]	4242[g]			2		6C
	93	25	4	15	17	20	8	4240[g]			1		6C
	10	13	1	43	98	1	20	1390			1		6A/6C
	2	10	54	4	6	3	14	1669			1		6A/6C
	13	1	1	5	6	1	18	1715	1		16		6A/6C
	7	25	8	6	25	6	8	1150			2		6B/6C
	5	10	9	43	13	1	14	600			3		6A/6C

MLST, multilocus sequence types; ST, sequence types.
[a]Clustering by eBURST analysis of 120 isolates in this study showed two clonal complexes or seven groups and 18 singleton STs (represented by 1–16 isolates) based on a minimum similarity of five identical loci.
[b,c]New allele of *recp* 131 and new allele of *ddl* 346.
[d]Sequence types of 120 isolates in this study.
[e–h]Represent MLST not previously described among serotypes 6A, 6B, 6C and 6D.
[i]Sequence type/serotype combinations previously described in published literature and http://www.mlst.net.
[j–l]ST2608 previously described in serotypes 6A and 10A, ST3604 in serotype 6A and 23F. ST1092 in serotype 6A, 6B, 6C and 19A.

Discussion

In countries where PCV7 has been used extensively in infants, there have been highly significant falls in the rates of invasive pneumococcal disease caused by vaccine (and some related) serotypes, particularly in the vaccine target age group but also, because of herd immunity, in older age groups [18–22]. However, serotype replacement under the pressure of vaccine use has been reported [23–25]. *Streptococcus pneumoniae* serotype 6C was first described in 2007, but was identified among stored isolates from as early as 1979 [3]. Subsequently, several studies have shown that it was generally uncommon (representing 2–3% of *S. pneumoniae* isolates) among both colonizing and invasive isolates before the introduction of PCV7 [7,26] but has since increased in several countries [5–7,27,28]. For example, the Centers for Disease Control and Prevention Active Bacterial Core surveillance showed an increase in the proportion of serotype 6C isolates among those initially identified as 6A by Quellung reaction, from 16.4% in 1999 to 69.1% in 2006 [5] and an increase in the rate of invasive disease caused by serotype 6C from 0.22 to 0.58 per 100 000 population, whereas those caused by serotypes 6A and 6B decreased

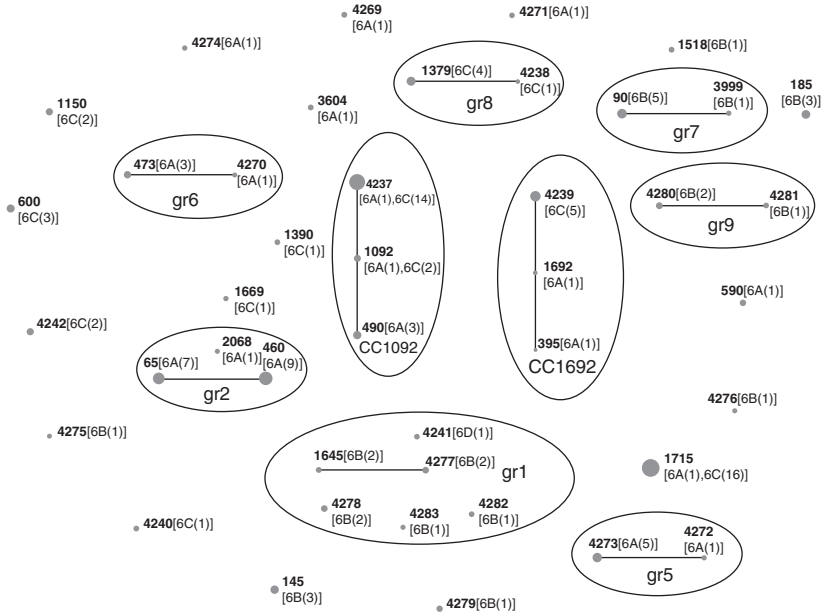

FIG. 1. Genetic relatedness of 120 serogroup 6 *Streptococcus pneumoniae* isolates in New South Wales, Australia. Clonal complexes (CCs), groups and singletons were determined by eBURST analysis using all allelic profiles reported in the multilocus sequence typing (MLST) database. The 42 sequence types (STs; black dots) detected in this study are represented. The sizes of the dots correspond to the number of isolates sharing the same ST. Within CCs (elliptic shapes) missing links (STs present in the MLST database but not found in the present study) are not represented. CC names refer to the predicted founder if any (i.e. CC1092 refers to ST1092 as predicted founder); 'gr' refers to the eBURST group if no founder is predicted.

significantly [5]. In Massachusetts the proportion of serotype 6A among nasopharyngeal pneumococcal isolates fell from 9.6% before 2001 to 2.9% in 2007 and that of serotype 6C increased from 0.6% to 8.7% [6].

In common with other reports, we have demonstrated a dramatic fall in number and proportion of serotype 6B among invasive pneumococcal isolates and increases in the numbers and proportions of serotype 6C since the introduction of PCV7 in NSW. Serotype 6C increased from 11.4% (12/105) of invasive isolates, originally serotyped as 6A by Quellung reaction, in the period 1999–2005 (before routine use of PCV7), to 40.8% (31/76), in the period 2006–2008. The explanation for the emergence of serotype 6C apparently lies in the observation that PCV7 elicits little opsonophagocytic activity against serotype 6C, compared with 6A and 6B strains [6,29]. Contrary to the findings of others, we did not demonstrate any significant change in numbers or proportions of invasive serotype 6A isolates, overall, after the introduction of PCV7. However, there were differences between age groups in that the average annual numbers of referred 6A isolates from older children and adults rose, whereas the proportion of 6A isolates from children under five fell, reflecting the effect of immunization. The possibility that the relatively modest changes in 6A prevalence could reflect a lower level of protection from the three-dose infant immunization schedule, as used in Australia, compared with the four-dose schedules used elsewhere, deserves consideration. On the other hand, previous studies have questioned the need for a fourth (booster) dose [30].

Several studies have reported relatively high levels of multidrug resistance among serotype 6C isolates [7,26] and, in one, the rates increased significantly between 1999 and 2007 [7]. Again, contrary to these findings, most of our 6C isolates were fully susceptible to penicillin and erythromycin and none was multidrug resistant. Nahm *et al.* also found that 6C isolates were more susceptible to penicillin than 6A isolates [6]. Presumably, these differences reflect local antibiotic use and indicate that the emergence of serotype 6C is primarily the result of immune pressure from vaccine use, rather than antibiotic pressure.

In common with other studies, we found considerable genotypic diversity among serogroup 6, including serotype 6C, isolates (Table 2). Several 6C-related STs were shared with 6A, including two (ST1715 and ST4237) that accounted for 58% of the serotype 6C isolates in the study. This finding supports the conclusion that serotype 6C is frequently derived from 6A as a result of recombination [3,5]. None of our 6C isolates shared STs with the 6B isolates tested in this

TABLE 3. Components of the three *cps* gene profiles for serotype 6C and 6D isolates[a]

Sequence types[b]	No. of isolates	GenBank accession no.[c]	Nucleotide differences
cpsA-cpsB-6C-1	35	1- EF538714	Reference sequence[d]
cpsA-cpsB-6C-2	15	2- GQ871848[b]	Multiple
cpsA-cpsB-6C-3	1	3- GQ871849[b]	Multiple
cpsA-cpsB-6C-4	1	4- AF246897	Multiple
cpsA-cpsB-6D	1	5- GQ871850[b]	Single
wchA-wciN$_{beta}$-wciO-6C-1	23	1- GQ871838[b]	Reference sequence[c]
wchA-wciN$_{beta}$-wciO-6C-2	11	2- GQ871839[b]	Single
wchA-wciN$_{beta}$-wciO-6C-3	7	3- GQ871840[b]	Single
wchA-wciN$_{beta}$-wciO-6C-4	4	4- GQ871841[b]	Single
wchA-wciN$_{beta}$-wciO-6C-5	2	5- GQ871842[b]	Two nucleotides
wchA-wciN$_{beta}$-wciO-6C-6	1	6- GQ871843[b]	Two nucleotides
wchA-wciN$_{beta}$-wciO-6C-7	1	7- GQ871844[b]	Single
wchA-wciN$_{beta}$-wciO-6C-8	1	8- GQ871845[b]	Single
wchA-wciN$_{beta}$-wciO-6C-9	1	9- GQ871846[b]	Single
wchA-wciN$_{beta}$-wciO-6C-10	1	10- FJ899597	Four nucleotides
wchA-wciN$_{beta}$-wciO-6D	1	2- GQ871847[b]	Single
wciP-6C-1	33	1- FJ899601	Reference sequence[c]
wciP-6C-2	12	2- EF538714	Two nucleotides
wciP-6C-3	7	3- GQ871851[b]	Single
wciP-6D -	1	4- EU714777	Multiple

[a]*cps* gene profiles consist of alleles of each of three genes/regions in the order shown: *cpsA-cpsB*-6C 1–4; *wchA-wciN$_{beta}$-wciO*-6C 1–10; *wciP*-6C 1–3 (and corresponding sequence for 6D, for which only a single isolates/genotype has so far been characterized (see Table 4).
[b]Sequence type nomenclature is based on the predominance of individual sequence types (1, 2, 3, etc. in decreasing order of prevalence).
[c]New sequences have been submitted to GenBank.
[d]The commonest allele for each *cps* gene sequence is characterized as the reference sequence.

study, but ST1092 (shared by 6A and 6C in our study) has been previously associated with both 6B and 6C [27]. Our single 6D isolate belonged to a previously unknown ST4241, which is a single locus variant of ST176, commonly associated with serotype 6B [27]. This finding is compatible with the assumption that *wciN$_{beta}$* has been directly introduced into 6B. The *wchA-wciN$_{beta}$-wciO* region sequence for this isolate differs by one SNP, compared with the reference sequence.

The combination of MLST and *cps* profiles demonstrated a high genetic diversity of serotype 6C isolates. We identified only two CCs among 6C isolates (CC1092 and gr8) in this study; all other STs were singletons. CC1092 included 14 isolates belonging to ST4237 and two belonging to ST1092, which differed by 14 bases between *spi* 6 and *spi* 9 alleles, suggesting recombination. The 14 ST4237 isolates had three *cps* profiles that differed from each other by only one SNP in the *wchA-wciN$_{beta}$-wciO* region. Group 8 includes four ST1379 isolates and one ST4238 isolate. The differences between these STs include one SNP between *spi* alleles. Most singleton STs, except those with only one isolate, showed variation (Table 4). For example, 16 isolates belonging to ST1715 produced four *cps* profiles that differed from each other by one or two SNPs in the *wchA-wciN$_{beta}$-wciO* region. These SNPs could be the result of either recombination or point mutation, but it has been shown that, in *Escherichia coli*, an SNP in a housekeeping gene is 10 to 50 times more likely to result from recombination than mutation [31].

In conclusion, all four pneumococcal serogroup 6 serotypes (6A, 6B, 6C and 6D) were represented among isolates referred to the NSW Pneumococcal Reference Laboratory between 1999 and 2008. The changes in numbers and proportions of different serotypes, including a significant increase in 6C and fall in 6B between two time periods

TABLE 4. Comparison of multilocus sequence types and *cps* gene profiles of serotype 6C isolates

MLST (No. of isolates)	*cps* gene sequence type profiles[a] (No. of isolates)	*cps* gene target[a] allele number- GenBank accession no. *cpsA-cpsB* (4 alleles)	*wchA-wciN$_{beta}$-wciO* (10 alleles)	*wciP* (3 alleles)
600 (3); 1379 (4); 1390 (1); 4238 (1); 4242 (1)	1-2-2 (10)	1- EF538714	2- GQ871839	2- EF538714
4242 (1)	1-1-2 (1)	1- EF538714	1- GQ871838	2- EF538714
1092 (2)	2-1-1 (1)	2- Q871848	1- GQ871838	1- FJ899601
	4-1-1 (1)	4- AF246897	1- GQ871838	1- FJ899601
1150 (2)	1-5-2 (1)	1- EF538714	5- GQ871842	2- EF538714
	1-5-3 (1)	1- EF538714	5- GQ871842	3- GQ871851
1669 (1)	1-2-3 (1)	1- EF538714	2- GQ871839	3- GQ871851
1715 (16)	1-1-1 (5)	1- EF538714	1- GQ871838	1- FJ899601
	1-3-1 (6)	1- EF538714	3- GQ871840	1- FJ899601
	1-4-1 (4)	1- EF538714	4- GQ871841	1- FJ899601
	1-6-1 (1)	1- EF538714	6- GQ871843	1- FJ899601
4237 (14)	2-1-1 (12)	2- GQ871848	1- GQ871838	1- FJ899601
	2-3-1 (1)	2- GQ871848	3- GQ871840	1- FJ899601
	2-7-1 (1)	2- GQ871848	7- GQ871844	1- FJ899601
4239 (5)	1-1-3 (3)	1- EF538714	1- GQ871838	3- GQ871851
	3-8-3 (1)	3- GQ871849	8- GQ871845	3- GQ871851
	1-9-3 (1)	1- EF538714	9- GQ871846	3- GQ871851
4240 (1)	1-10-1 (1)	1- EF538714	10- GQ871847	1- FJ899601

MLST, multilocus sequence types.
[a]*cps* gene profile (or sequence type) is represented by combinations of different alleles at each of three *cps* gene regions sequenced; profiles are represented by a three-digit code in the order shown in Table 3. The corresponding GenBank accession number for each allele is shown.

(2000–2005 and 2006–2008), were apparently related to the widespread use of PCV7. Both mutation and recombination may have contributed to genotypic diversity among serotype 6C isolates and its rapid increase in prevalence suggests that it has the potential to replace 6B as the predominant invasive serogroup 6 serotype. Profiling of the *cps* gene is a useful genotyping method, which will potentially complement MLST in studies of epidemiology and population diversity of serotype 6C.

Authors' contributions

GLG initiated and supervised the molecular serotyping of *S. pneumoniae* project. FRK, FLZ and MX designed research. FLZ carried out the molecular work and sequence alignment, MX participated in the sequence alignment and SO performed the bacteriological work and serotyping. FRK helped to draft the manuscript and FLZ and MX wrote the manuscript. GLG, JZZ, FZ and FRK revised the manuscript.

Acknowledgements

The authors wish to thank laboratory staff at the Kids Research Institute (Children's Hospital at Westmead) and at the Pneumococcal Reference Laboratory, Centre for Infectious Diseases and Microbiology (ICPMR, Westmead Hospital).

Transparency Declaration

Fenglin Zhuo was supported by the Chinese Government Graduate Student Overseas Study Program. The project was supported by National Institutes of Health DMID Funding Mechanism: R01 AI 52337 and National Health and Medical Research Council (NHMRC) project grant #358351 (Chief Investigator A, G.L.Gilbert). GLG has received small research grants from Wyeth Australia (Pty Ltd) and GlaxoSmithKline. The other authors have no conflict of interest to declare.

References

1. Kadioglu A, Weiser JN, Paton JC *et al.* The role of *Streptococcus pneumoniae* virulence factors in host respiratory colonization and disease. *Nat Rev Microbiol* 2008; 6: 288–301.
2. Sorensen UB. Typing of pneumococci by using 12 pooled antisera. *J Clin Microbiol* 1993; 31: 2097–2100.
3. Park IH, Park S, Hollingshead SK *et al.* Genetic basis for the new pneumococcal serotype, 6C. *Infect Immun* 2007; 75: 4482–4489.
4. Park IH, Pritchard DG, Cartee R *et al.* Discovery of a new capsular serotype (6C) within serogroup 6 of *Streptococcus pneumoniae*. *J Clin Microbiol* 2007; 45: 1225–1233.
5. Carvalho Mda G, Pimenta FC, Gertz RE Jr *et al.* PCR-based quantitation and clonal diversity of the current prevalent invasive serogroup 6 pneumococcal serotype, 6C, in the United States in 1999 and 2006 to 2007. *J Clin Microbiol* 2009; 47: 554–559.
6. Nahm MH, Lin J, Finkelstein JA *et al.* Increase in the prevalence of the newly discovered pneumococcal serotype 6C in the nasopharynx after introduction of pneumococcal conjugate vaccine. *J Infect Dis* 2009; 199: 320–325.
7. Nunes S, Valente C, Sá-Leão R *et al.* Temporal trends and molecular epidemiology of recently described serotype 6C of *Streptococcus pneumoniae*. *J Clin Microbiol* 2009; 47: 472–474.
8. Hermans PW, Blommaart M, Park IH *et al.* Low prevalence of recently discovered pneumococcal serotype 6C isolates among healthy Dutch children in the pre-vaccination era. *Vaccine* 2008; 26: 449–450.
9. Bratcher PE, Park IH, Hollingshead SK *et al.* Production of a unique pneumococcal capsule serotype belonging to serogroup 6. *Microbiology* 2009; 155: 576–583.
10. Jin P, Kong F, Xiao M *et al.* First report of putative *Streptococcus pneumoniae* serotype 6D among nasopharyngeal isolates from Fijian children. *J Infect Dis* 2009; 200: 1375–1380.
11. Bratcher PE, Kim KH, Kang JH *et al.* Identification of natural pneumococcal isolates expressing serotype 6D by genetic, biochemical and serological characterization. *Microbiology* 2010; 156: 555–560.
12. Jin P, Xiao M, Kong F *et al.* Simple, accurate, serotype-specific PCR assay to differentiate *Streptococcus pneumoniae* serotypes 6A, 6B, and 6C. *J Clin Microbiol* 2009; 47: 2470–2474.
13. Enright MC, Spratt BG. A multilocus sequence typing scheme for *Streptococcus pneumoniae*: identification of clones associated with serious invasive disease. *Microbiology* 1998; 144: 3049–3060.
14. Moore MR, Gertaz RE Jr, Woodbury RL *et al.* Population snapshot of emergent *Streptococcus pneumoniae* serotype 19A in the United States. *J Infect Dis* 2008; 197: 1016–1027.
15. 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: 1047–1058.
16. CLSI. *Performance standards for antimicrobial susceptibility testing*. Eighteenth informational supplement, Vol. 28, No. 1. Approved standard M2-A9. Wayne, PA: Clinical and Laboratory Standards Institute, 2008.
17. Mavroidi A, Godoy D, Aanensen DM, Robinson DA, Hollingshead SK, Spratt BG. Evolutionary genetics of the capsular locus of serogroup 6 pneumococci. *J Bacteriol* 2004; 186: 8181–8192.
18. Lexau CA, Lynfield R, Danila R *et al.* Changing epidemiology of invasive pneumococcal disease among older adults in the era of pediatric pneumococcal conjugate vaccine. *JAMA* 2005; 294: 2043–2051.
19. O'Brien KL, Moulton LH, Reid R *et al.* Efficacy and safety of seven-valent conjugate pneumococcal vaccine in American Indian children: group randomised trial. *Lancet* 2003; 362: 355–361.
20. Roche PW, Krause V, Cook H *et al.* Invasive pneumococcal disease in Australia. *Commun Dis Intell* 2008; 32: 18–30.
21. Whitney CG, Farley MM, Hadler J *et al.* Decline in invasive pneumococcal disease after the introduction of protein–polysaccharide conjugate vaccine. *N Engl J Med* 2003; 348: 1737–1746.
22. Whitney CG, Pilishvili T, Farley MM *et al.* Effectiveness of seven-valent pneumococcal conjugate vaccine against invasive pneumococcal disease: a matched case–control study. *Lancet* 2006; 368: 1495–1502.
23. Byington CL, Samore MH, Stoddard GJ *et al.* Temporal trends of invasive disease due to *Streptococcus pneumoniae* among children in

the intermountain west: emergence of nonvaccine serogroups. *Clin Infect Dis* 2005; 41: 21–29.
24. Hicks LA, Harrison LH, Flannery B *et al.* Incidence of pneumococcal disease due to non-pneumococcal conjugate vaccine (PCV7) serotypes in the United States during the era of widespread PCV7 vaccination, 1998–2004. *J Infect Dis* 2007; 196: 1346–1354.
25. Singleton RJ, Hennessy TW, Bulkow LR *et al.* Invasive pneumococcal disease caused by nonvaccine serotypes among Alaska native children with high levels of 7-valent pneumococcal conjugate vaccine coverage. *JAMA* 2007; 297: 1784–1792.
26. Campos LC, Carvalho Mda G, Beall BW *et al.* Prevalence of *Streptococcus pneumoniae* serotype 6C among invasive and carriage isolates in metropolitan Salvador, Brazil, from 1996 to 2007. *Diagn Microbiol Infect Dis* 2009; 65: 112–115.
27. Jacobs MR, Bajaksouzian S, Bonomo RA *et al.* Occurrence, distribution, and origins of *Streptococcus pneumoniae* serotype 6C, a recently recognized serotype. *J Clin Microbiol* 2009; 47: 64–72.
28. Jacobs MR, Good CE, Bajaksouzian S *et al.* Emergence of *Streptococcus pneumoniae* serotypes 19A, 6C, and 22F and serogroup 15 in Cleveland, Ohio, in relation to introduction of the protein-conjugated pneumococcal vaccine. *Clin Infect Dis* 2008; 47: 1388–1395.
29. Park IH, Moore MR, Treanor JJ *et al.* Differential effects of pneumococcal vaccines against serotypes 6A and 6C. *J Infect Dis* 2008; 198: 1818–1822.
30. Mahon BE, Hsu K, Karumuri S *et al.* Effectiveness of abbreviated and delayed 7-valent pneumococcal conjugate vaccine dosing regimens. *Vaccine* 2006; 24: 2514–2520.
31. Guttman DS, Dykhuizen DE. Clonal divergence in *Escherichia coli* as a result of recombination, not mutation. *Science* 1994; 266: 1380–1383.