

RESEARCH ARTICLE

Open Access

# Pyrosequencing-based analysis reveals a novel capsular gene cluster in a KPC-producing *Klebsiella pneumoniae* clinical isolate identified in Brazil

Pablo Ivan Pereira Ramos<sup>1</sup>, Renata Cristina Picão<sup>2,4</sup>, Eliana Carolina Vespero<sup>3</sup>, Marsileni Pelisson<sup>3</sup>, Luiz Fernando Goda Zuleta<sup>1</sup>, Luiz Gonzaga P Almeida<sup>1</sup>, Alexandra L Gerber<sup>1</sup>, Ana Tereza R Vasconcelos<sup>1</sup>, Ana Cristina Gales<sup>4</sup> and Marisa Fabiana Nicolás<sup>1\*</sup>

## Abstract

**Background:** An important virulence factor of *Klebsiella pneumoniae* is the production of capsular polysaccharide (CPS), a thick mucus layer that allows for evasion of the host's defense and creates a barrier against antibacterial peptides. CPS production is driven mostly by the expression of genes located in a locus called *cps*, and the resulting structure is used to distinguish between different serotypes (K types). In this study, we report the unique genetic organization of the *cps* cluster from *K. pneumoniae* Kp13, a clinical isolate recovered during a large outbreak of nosocomial infections that occurred in a Brazilian teaching hospital.

**Results:** A pyrosequencing-based approach showed that the *cps* region of Kp13 (*cps*<sub>Kp13</sub>) is 26.4 kbp in length and contains genes common, although not universal, to other strains, such as the *rm*/BADDC operon that codes for L-rhamnose synthesis. *cps*<sub>Kp13</sub> also presents some unique features, like the inversion of the *wzy* gene and a unique repertoire of glycosyltransferases. In silico comparison of *cps*<sub>Kp13</sub> RFLP pattern with 102 previously published *cps* PCR-RFLP patterns showed that *cps*<sub>Kp13</sub> is distinct from the C patterns of all other K serotypes. Furthermore, in vitro serotyping showed only a weak reaction with capsular types K9 and K34. We confirm that K9 *cps* shares common genes with *cps*<sub>Kp13</sub> such as the *rm*/BADDC operon, but lacks features like *uge* and Kp13-specific glycosyltransferases, while K34 capsules contain three of the five sugars that potentially form the Kp13 CPS.

**Conclusions:** We report the first description of a *cps* cluster from a Brazilian clinical isolate of a KPC-producing *K. pneumoniae*. The gathered data including K-serotyping support that Kp13's K-antigen belongs to a novel capsular serotype. The CPS of Kp13 probably includes L-rhamnose and D-galacturonate in its structure, among other residues. Because genes involved in L-rhamnose biosynthesis are absent in humans, this pathway may represent potential targets for the development of antimicrobial agents. Studying the capsular serotypes of clinical isolates is of great importance for further development of vaccines and/or novel therapeutic agents. The distribution of K-types among multidrug-resistant isolates is unknown, but our findings may encourage scientists to perform K-antigen typing of KPC-producing strains worldwide.

**Keywords:** Capsular gene cluster, Capsular polysaccharide, K-antigen, KPC-producing *K. pneumoniae*, Molecular serotyping, Monosaccharide biosynthesis pathways

\* Correspondence: marisa@lncc.br

<sup>1</sup>Laboratório Nacional de Computação Científica (LNCC), Petrópolis, Rio de Janeiro, Brazil

Full list of author information is available at the end of the article

## Background

*Klebsiella pneumoniae* is a Gram-negative, rod-shaped bacterium frequently associated with nosocomial and community-acquired infections [1]. Over the past decade, healthcare practitioners have observed the rapid evolution of antimicrobial resistance among *K. pneumoniae* clinical isolates worldwide. The emergence and subsequent global spread of strains producing *Klebsiella pneumoniae* carbapenemase (KPC) represents a significant threat to public health [2]. The gene encoding this  $\beta$ -lactam resistance factor is frequently carried along with genes conferring resistance to multiple classes of antimicrobial agents. As a result, the therapeutic options to treat infections caused by KPC-producing *K. pneumoniae* are generally scarce and in some instances limited to polymyxins [2].

The development of an effective response against *K. pneumoniae* infections depends on the integrity of the immune system. Indeed, many authors have provided evidence that activation of the inflammatory response is required to clear such infections [3–5]. Unfortunately, most patients infected by multidrug-resistant *K. pneumoniae* have serious underlying conditions and/or a compromised immune status [1,6]. Capsule production is believed to be one of the most important virulence factors for this species. The polysaccharide matrix found on its cell surface may prevent desiccation, confer adherence to host cells and protect it against both non-specific and specific host immunity [7]. However, there are differences in the degree of virulence conferred by different *Klebsiella* capsule types, possibly depending on the mannose and/or rhamnose content of the CPS [1]. The *K. pneumoniae* capsule is generally composed of acidic polysaccharides, including uronic acid repeats and, in several instances, mannose, rhamnose,

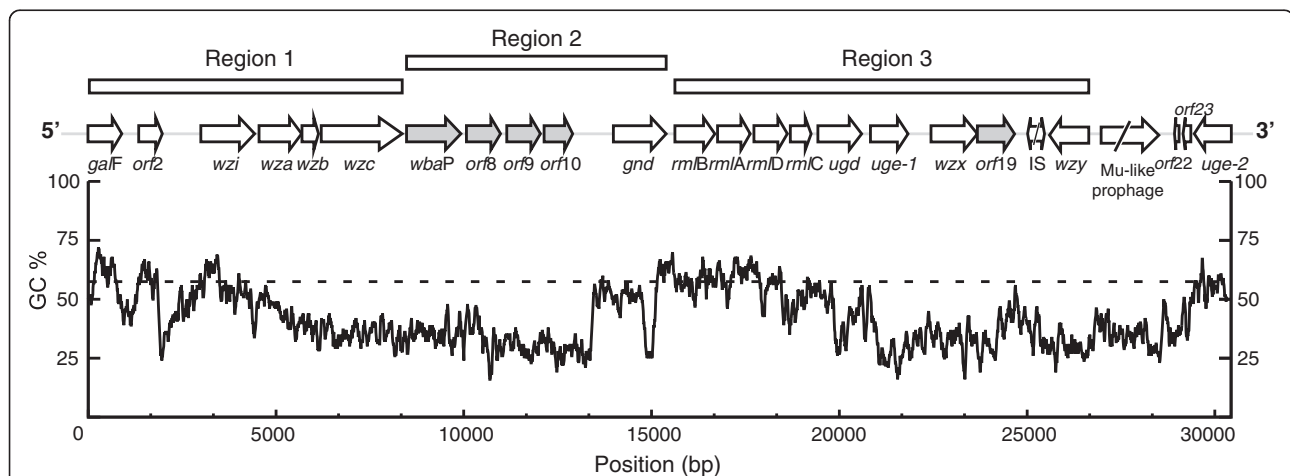
galactose, pyruvate and fucose residues [8]. The genes involved in the biosynthesis, transport and assembly of *K. pneumoniae* and *E. coli* group 1 capsules are found at a locus called *cps*, which is organized similarly in the two species [9]. The biosynthetic process of both types of capsules is also related between the two bacteria. Briefly, CPS synthesis initially takes place on the cytoplasmic side of the inner membrane with the assembly of individual sugar repeat residues which are linked by the sequential activities of specific glycosyltransferases (GTs) [10]. These are then flipped across the inner membrane by the action of the Wzx protein and undergo polymerization by the Wzy protein [11]. Polymerization control and translocation of the nascent polymer to the cell surface occurs with the coordinated action of Wza, Wzb and Wzc proteins [12].

To date, a variety of *cps* gene clusters have been characterized in *Klebsiella* spp., mostly from isolates recovered in the USA, Asia and Europe [13–15]. To our knowledge, there have been no studies on the *cps* organization of *K. pneumoniae* isolates from Brazil, KPC-producing or otherwise. Here, we report the unique *cps* organization of a KPC-producing *K. pneumoniae* isolate showing multidrug resistance. This bacterium was responsible for a large nosocomial outbreak in a teaching hospital located in Southern Brazil (Ana C. Gales, personal communication).

## Results and Discussion

### General features of the *cps*<sub>Kp13</sub> gene cluster

The *cps*<sub>Kp13</sub> gene cluster is 26.4 kbp in length and contains 20 open reading frames (ORFs) from *galF* to *wzy* (Figure 1, Table 1). The average GC content of these genes is 42%, which is lower than the average GC content of the entire Kp13 genome (57.5%, data not shown).



**Figure 1** Overall organization of the *cps* cluster of *K. pneumoniae* Kp13. The *cps*<sub>Kp13</sub> spans *galF* to *wzy*. ORFs are represented by arrows (gray for those encoding glycosyltransferases and double-headed for possible mobile elements). Rectangles above the ORFs represent distinct variably conserved regions of the *cps* cluster as discussed in the text. A plot of the GC content of the region using a 100-bp sliding window is shown below. The dashed horizontal line represents the mean GC content of the entire Kp13 chromosome.

**Table 1 General features of the 20 coding sequences identified in the Kp13 cps gene cluster**

ORF	Size (bp)	%GC	Gene name	Product	EC number	Best BLASTP hit (accession number) (identity)
KP03136	900	59.02	<i>galF</i>	UTP-glucose-1-phosphate uridylyltransferase	2.7.7.9	<i>K. pneumoniae</i> strain NK8 (BAI43699) (100%)
KP03135	627	58.41	<i>orf2</i>	Uncharacterized phosphatidic acid phosphatase protein	3.1.3.4	<i>K. pneumoniae</i> strain MGH 78578 (ABR77932) (100%) and strain VGH404 serotype K5 (BAI43755) (100%).
KP03809	1,431	55.99	<i>wzi</i>	Capsule assembly 55.8 kDa protein		<i>K. pneumoniae</i> strain VGH484 serotype K9 (BAI43775) (98%)
KP03808	1,131	45.15	<i>wza</i>	Capsule polysaccharide export protein		<i>K. pneumoniae</i> strain VGH484 serotype K9 (BAI43776) (97%)
KP03807	438	39.69	<i>wzb</i>	Protein-tyrosine-phosphatase	3.1.3.48	<i>K. pneumoniae</i> strain MGH 78578 (ABR77929) (78%)
KP03806	2,154	35.61	<i>wzc</i>	Uncharacterized tyrosine-protein kinase	2.7.10.-	<i>K. pneumoniae</i> strain MGH 78578 (ABR77928) (79%)
KP31533	1,446	35.2	<i>wbaP</i>	Undecaprenolphosphate Gal-1-P transferase	2.-.-	<i>K. pneumoniae</i> strain MGH 78578 (ABR77927) (79%)
KP03804	906	37.51	<i>orf8</i>	Uncharacterized glycosyltransferase family 2	2.4.1-	<i>K. pneumoniae</i> strain A1517 (BAF75773) (67%)
KP03803	894	30.99	<i>orf9</i>	Uncharacterized glycosyltransferase family 2	2.4.1-	<i>Dickeya dadantii</i> (ADM97617) (63%)
KP03802	759	29.79	<i>orf10</i>	Uncharacterized glycosyltransferase	2.4.1.-	<i>D. dadantii</i> (ADM97619) (57%)
KP31534	1,404	51.46	<i>gnd</i>	6-phosphogluconate dehydrogenase, decarboxylating	1.1.1.44	<i>K. pneumoniae</i> strain VGH484 serotype K9 (BAI43786) (99%)
KP31530	1,062	59.25	<i>rmIB</i>	dTDP-D-glucose 4,6-dehydratase	4.2.1.46	<i>K. pneumoniae</i> strain VGH484 serotype K9 (BAI43787) (98%)
KP03797	867	58.74	<i>rmlA</i>	Glucose-1-phosphate thymidyltransferase	2.7.7.24	<i>Escherichia coli</i> HS (EFK17576) (98%)
KP03796	888	61.5	<i>rmlD</i>	dTDP-4-dehydrorhamnose reductase	1.1.1.133	<i>K. pneumoniae</i> strain MGH 78578 (ABR77913) (98%)
KP03795	552	54.41	<i>rmlC</i>	dTDP-4-dehydrorhamnose 3,5-epimerase	5.1.3.13	<i>K. pneumoniae</i> strain VGH484 serotype K9 (BAI43790) (99%)
KP03794	1,164	50.82	<i>ugd</i>	UDP-glucose 6-dehydrogenase	1.1.1.22	<i>K. pneumoniae</i> strain NK8 (BAI43716) (100%) and strain VGH404 serotype K5 (BAI43755) (100%)
KP03793	999	41.92	<i>uge-1</i>	Uridine diphosphate galacturonate 4-epimerase	5.1.3.6	<i>K. pneumoniae</i> subsp. <i>rhinoscleromatis</i> ATCC 13884 (EEW43608) (97%)
KP31531	1,233	31.57	<i>wzx</i>	K-antigen flippase Wzx		<i>E. coli</i> TA27 (ZP_07523140) (64%)
KP03791	990	31.32	<i>orf19</i>	Uncharacterized glycosyltransferase family 2	2.4.1.-	<i>Cronobacter sakazakii</i> (ABX51890) (33%)
KP03789	1,044	29.61	<i>wzy</i>	K antigen polymerase Wzy		<i>Thermoanaerobacter wiegeli</i> (ACF14522) (35%)

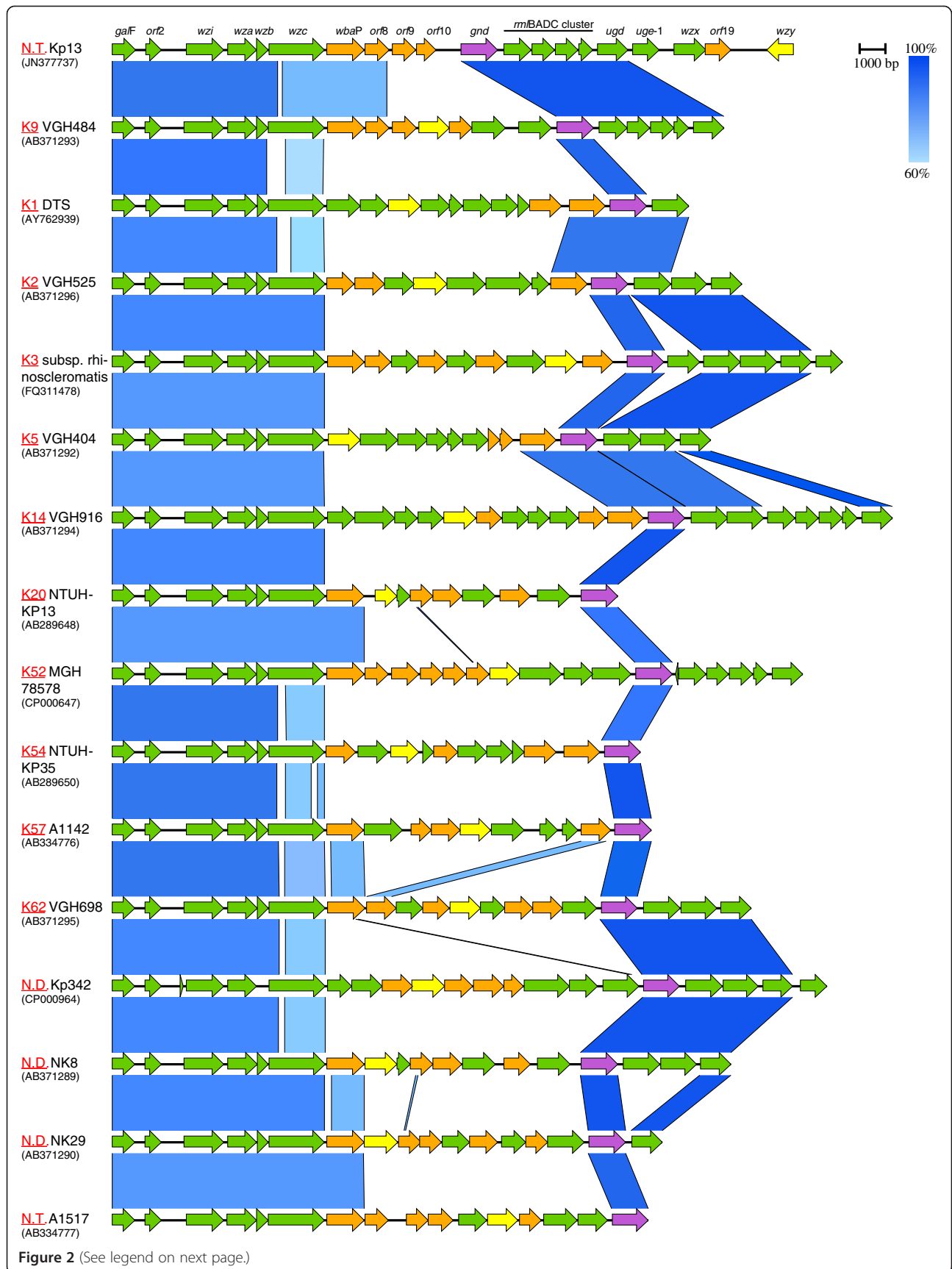
Comparable GC content has been reported for twelve other *K. pneumoniae* cps clusters [15].

The *cps*<sub>Kp13</sub> has a genomic organization similar to other *K. pneumoniae* cps clusters, and it can be divided into three regions as shown in Figure 1. The 5' end or region 1 (from *galF* to *wbaP*) contains conserved genes responsible for polymer assembly and translocation [12]. The central region or region 2 contains genes encoding serotype-specific GTs and *gnd*. The 3' end or region 3 is more variable among different capsular types, with some containing the *manCB* operon that encodes GDP-D-mannose, like serotypes K1 and K5 [15]. Similarly to serotypes K9 and K52, the 3' end of the *cps*<sub>Kp13</sub> gene cluster contains the *rmIBADC* operon for the synthesis of dTDP-L-rhamnose instead of the *manCB* operon [15]. The genes *wzx* and *wzy* are also found in the 3' region of the Kp13 cps cluster. This region is succeeded by defective IS elements and a prophage fragment (Figure 1). The discussed conservation of region 1 and variability of region 2 can be readily observable on

a comparison of the cps loci of different K-types deposited in NCBI (Figure 2).

#### The *cps*<sub>Kp13</sub> monosaccharide biosynthesis pathways: UDP-D-glucuronate, UDP-D-galacturonate and L-rhamnose

As in other bacteria that produce group-1 capsules, *galF* delimits the 5' region of *cps*<sub>Kp13</sub>. This gene shows 100% identity to the *galF* sequence present in *K. pneumoniae* NK8 [GenBank:BAI43699], which codes for a UTP-glucose-1-phosphate uridylyltransferase (EC 2.7.7.9, Figure 3). This enzyme belongs to the nucleotidyltransferase family and catalyzes the reaction UTP + α-D-glucose 1-phosphate ↔ diphosphate + UDP-D-glucose. This enzyme is important because UDP-D-glucose serves as a precursor for the biosynthesis of bacterial lipopolysaccharides and capsular polysaccharides. It is also possible that the *galF* product interacts with the product of *galU*, thus elevating UDP-D-glucose concentration in the cell and providing more material for the synthesis of capsular polysaccharides [11]. In fact, a *galU* homolog found in Kp13 outside the cps region



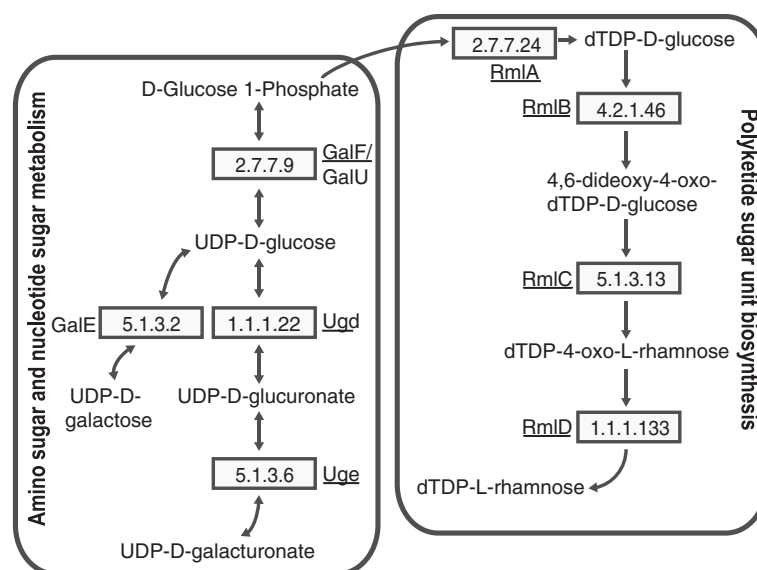
(See figure on previous page.)

**Figure 2 Comparison of sequenced *K. pneumoniae* cps loci.** For each *cps* cluster, a two-way comparison with the clusters immediately above and/or below is presented. The K-type of each compared cluster is shown in red, followed by the strain/isolate identification and its NCBI accession number in parentheses. The blue segments connecting each cluster represent variably conserved (60–100% identity) regions among them (from a BLASTN comparison with  $e\text{-value} \leq 10^{-4}$ ). Predicted glycosyltransferases are colored in orange, *wzy* and *gnd* homologs in yellow and purple, respectively. N.T., new K-type; N.D., K-type not determined.

(KP04702) shows 94% identity (BLASTP) to GalU from *Shigella flexneri* [Swiss-Prot:POAEP6]. Immediately downstream of the *rmlBADC* operon, the gene *ugd* is found (Figure 1). It encodes a UDP-glucose 6-dehydrogenase (EC 1.1.1.22). As depicted in Figure 3, this enzyme converts UDP-D-glucose to UDP-D-glucuronate, a common constituent of bacterial capsules [7]. As with other sequences located in the 3' region of the *cps*<sub>Kp13</sub> gene cluster, this coding sequence exhibits remarkable amino acid conservation. It is 100% identical to Ugd from *K. pneumoniae* strains NK8 [GenBank:BAI43716] and VGH404 serotype K5 [GenBank:BAI43755] (Table 1), both studied by Shu et al. [15]. Uge catalyzes the conversion of UDP-D-glucuronate to UDP-D-galacturonate (Figure 3), which is also present in both bacterial capsules and LPS. In fact, Kp13 has two copies of this gene, *uge-1* (KP03793) and *uge-2* (KP03786). A NAD-dependent epimerase domain (Pfam accession no. PF01370) is predicted to occupy amino acids 4 to 230 in both Uge sequences. Two copies of *uge* are also found in the genome of *K. pneumoniae* subsp. *rhinoscleromatis* (which produces a K3 capsule), one in the *cps* cluster and an inverted adjacent copy in the cluster for LPS synthesis [16]. As the K3 CPS contains D-galacturonate in its composition, *uge* was

considered the last gene of its *cps* cluster [16] instead of *ugd* as usually regarded [15,17]. In Kp13 *uge-1* should also be considered within the *cps* since the genes necessary to the flippase and polymerase activities, *wzx* and *wzy* respectively, are located downstream (Figure 1); therefore, D-galacturonate could also form the Kp13 CPS composition. The effects of a *uge* null mutation on colonization and virulence were studied in *K. pneumoniae* 52145, which is a highly virulent strain able to colonize different surfaces [18]. A *uge* deletion reduced colonization and rendered the strain completely avirulent in an experimental model of pneumonia [18]. This suggests that the *uge-1* and/or *uge-2* mutation in Kp13 could have important, measurable effects on colonization and virulence.

In the *cps*<sub>Kp13</sub> cluster, genes encoding enzymes that participate on the synthesis of dTDP-L-rhamnose from glucose 1-phosphate are found immediately downstream of the *gnd* gene (Figure 1). The *rmlBADC* genes were found in three capsular serotypes studied by Shu et al. [15]: K9, K14 and K52. In serotypes K9 and K52, these genes are also found downstream of *gnd*. The lengths of the products encoded by *rmlA*, *rmlB*, *rmlC* and *rmlD* are shown in Table 1, along with the best BLAST hits for these genes. The gene *rmlA* codes for a glucose-



**Figure 3 Amino- and polyketide sugar production in *K. pneumoniae* Kp13.** Pathways leading to UDP-D-galacturonate, UDP-D-galactose and dTDP-L-rhamnose are shown, as these residues could be present in the capsular structure of Kp13. Enzymes coded by genes present in the *cps*<sub>Kp13</sub> cluster are underlined.

1-phosphate thymidyltransferase (EC 2.7.7.24), which catalyzes the first reaction of L-rhamnose synthesis: dTTP +  $\alpha$ -D-glucose 1-phosphate  $\rightarrow$  diphosphate + dTDP-D-glucose (Figure 3). The second reaction is performed by dTDP-D-glucose 4,6-dehydratase (EC 4.2.1.46, Figure 3), the product of *rmlB*, which catalyzes the dehydration of dTDP-D-glucose to dTDP-4-keto 6-deoxy-D-glucose. Epimerization at the C3' and C5' positions of this molecule is performed by dTDP-4-dehydrorhamnose 3,5-epimerase (*rmlC*, EC 5.1.3.13, Figure 3), producing dTDP-4-oxo-L-rhamnose. Finally, dTDP-4-dehydrorhamnose reductase (EC 1.1.1.133, Figure 3), encoded by *rmlD*, catalyzes the reduction of dTDP-4-oxo-L-rhamnose to dTDP-L-rhamnose, which can be subsequently linked to the capsular polymer by a specific rhamnosyltransferase. All three conserved regions (the Y-X<sub>3</sub>-K loop, the Wierenga motif G-X<sub>2</sub>-G-X<sub>2</sub>-G and the STDYVF sequence) discussed by Giraud and Naismith [19] are present in Kp13's RmlD.

Whereas the chemical composition of the Kp13 capsule remains to be determined, the pyrosequencing-based genomic analysis of *cps*<sub>Kp13</sub> allowed the identification of sugar metabolic pathways. Genes encoding enzymes for the biosynthesis of sugar nucleotide precursors in the Kp13 capsule, such as UDP-D-glucose, UDP-D-glucuronate, UDP-D-galacturonate and dTDP-L-rhamnose, are found in the *cps* cluster. Thus, the capsule of Kp13 may contain any of these sugar nucleotide precursors. In particular, the presence of two genes in the *cps* cluster encoding possible rhamnosyltransferases suggests that L-rhamnose makes up part of the Kp13 capsule (see discussion below). On the other hand, *galE* (KP02995) was identified outside the *cps* region, and it encodes a UDP-glucose 4-epimerase with roles in the amino sugar and nucleotide sugar pathways producing UDP-D-galactose from UDP-D-glucose (Figure 3). The presence of this gene suggests that the capsule composition of Kp13 could also include UDP-D-galactose derivatives. Neither the *manA*, *manB* and *manC* genes of the *cps* cluster nor other genes of the mannose and fucose biosynthesis pathways were identified in the Kp13 genome. This suggests that the CPS of Kp13 does not contain GDP-D-mannose or GDP-L-fucose derivatives.

#### **Proteins involved in translocation, surface assembly and polymerization: Wzi, Wza, Wzb, Wzc, Wzx and Wzy**

The deduced amino acid sequences of the *wzi* and *wza* genes found in *cps*<sub>Kp13</sub> show 98% and 97% identity, respectively, with homologs from *K. pneumoniae* VGH484 (Table 1), and both proteins were predicted to localize in the outer membrane (PSORTb scores: Wzi, 9.52; Wza, 9.92). Moreover, a signal peptide was predicted for the *wzi* gene product. Analysis of the secondary structure of the Kp13 Wzi protein using PSIPRED showed that it is rich in  $\beta$ -sheet regions (data not shown), an observation that has been experimentally confirmed for a Wzi

homolog in *E. coli* [GenBank:AAD21561.1] [20] which shares 98% identity with that of Kp13. Also, Rahn et al. [20] established the importance of the Wzi outer membrane protein for capsule synthesis by showing that *wzi* mutants have lower amounts of cell-associated capsular polysaccharide.

The *wza* product of Kp13 has 92% identity with Wza from *E. coli* [GenBank:AAD21562.1], which has been shown to be an integral lipoprotein with exposed regions on the cell surface. The *E. coli* protein forms a ring-like structure responsible for polymer translocation through the outer membrane [12]. Wzc and Wzb are a tyrosine autokinase and its cognate acid phosphatase, respectively, and they are ubiquitously found in group 1 capsule clusters [12,21]. The Kp13 Wzc protein was predicted to have two transmembrane regions, like its counterpart in the *K. pneumoniae* strain Chedid, with which it shares 72% amino acid identity [Swiss-Prot:Q48452]. The inner membrane is the probable location of Kp13' Wzc (PSORTb score 9.99), in agreement with its role in capsule synthesis. Wzc is involved in the translocation of capsular polysaccharide from the periplasm to the cellular surface through formation of a complex with Wza [22]. Wzc undergoes autophosphorylation of its tyrosine-rich C-terminal residues (of the last 17 residues in Kp13 Wzc, eight are Tyr) potentially modulating the opening and closing of the translocation channel [12]. The Wzb protein (EC 3.1.3.48) of Kp13 is probably located in the cytoplasm (PSORTb score: 9.26). Wzb catalyzes the removal of a phosphate group from phosphorylated Wzc and is necessary for continued polymerization of the repeat units [12]. Sequence conservation of the Wzy and Wzx proteins is usually low [23]. The localization of *wzx* and *wzy* in Kp13 is different from that observed in various K-serotypes by Shu et al. [15], in which the genes usually mapped upstream of *gnd*. In Kp13, both genes are located downstream of *gnd*, in region 3 of the *cps* cluster, and *wzy* is transcribed in the opposite direction relative to other *cps* genes. Wzx is an inner membrane protein that transfers the polysaccharide units, assembled in the cytoplasm, into the periplasm, thus acting as a flippase [12]. The Wzx protein from *cps*<sub>Kp13</sub> has 10 predicted transmembrane segments and is 411 aa long, which is in agreement with a previous study of this protein in *E. coli* that predicted 10–12 transmembrane segments [23]. BLASTP against the NCBI database shows that the best hit (64% identity) is a putative Wzx protein from *E. coli* TA271 (NCBI accession no. ZP\_07523140, Table 1). A polysaccharide biosynthesis domain (Pfam accession no. PF01943), common to Wzx proteins, was found spanning amino acids 8 to 275 of Kp13 Wzx.

Wzy from Kp13 is 348 aa long and also had 10 predicted transmembrane segments, similar to the Wzy proteins of other *Enterobacteriaceae* that have 10–11

transmembrane segments [24]. This protein is believed to be a polysaccharide polymerase, although experimental evidence for this activity has not yet been reported due to the technical difficulty of working with Wzy in vitro [12]. NCBI BLASTP searches show that the best hit (35% identity) for Wzy is a conserved protein from *Thermoaerobacter wiegeleri* [GenBank:ACF14522.1] (Table 1).

It is remarkable that the *wzy* gene from isolate Kp13 is transcribed in the opposite direction compared to other genes of the *cps* cluster, a characteristic that to our knowledge has not been reported for previously studied *cps* clusters, as can be observed in Figure 2, where the position of *wzy* within different *K. pneumoniae cps* loci is highlighted.

Downstream *wzy*, we have identified an 862-bp region showing 70% identity to an IS element of the IS3 family [GenBank:CP002438.1]. No terminal inverted repeats or target site duplications were found in this element. Although three ORFs identified within this putative IS showed significant identity to distinct transposases, these structures do not seem to encode functional enzymes. The occurrence of mutations leading to premature stop codons and/or frameshifts might have rendered this transposase non-functional. Alternatively, this chimeric structure could have resulted from homologous recombination events with other transposase-encoding genes. Upstream *wzy*, there is a 1539-bp ORF whose deduced amino acid sequence shows 31% identity to a defective tail fiber protein of a Mu-like prophage identified in *Dickeya dadantii* [GenBank:ADM97620]. Notably, other prophage genes were absent. The location of *wzy* between two defective mobile genetic elements suggests that this gene may have been incorporated into Kp13's *cps* via an ancient horizontal gene transfer event. Subsequently, these mobile genetic elements may have been truncated in order to assure that *wzy* would be permanently present in the *cps*<sub>Kp13</sub> region, which is desirable because capsule assembly is a Wzy-dependent process.

The conserved gene *gnd*, found in the central region of *cps*<sub>Kp13</sub>, encodes a 468 aa protein (6-phosphogluconate dehydrogenase, EC 1.1.1.44, Figure 3) that catalyzes the conversion of 6-phospho-D-gluconate to D-ribulose 5-phosphate during the third step of the pentose phosphate pathway. This gene was found in all of the *cps* gene clusters studied by Shu et al. [15] and shows a high degree of conservation among them, which would be expected from an evolutionary standpoint due to the central role of this metabolic pathway. At the protein sequence level, the best hit (99% identity) for Kp13's *gnd* product is an ortholog from strain VGH484, serotype K9 [GenBank:BAI43786.1] (Table 1).

#### **Kp13's *cps* gene cluster has five GTs: WbaP, Orf8, Orf9, Orf10 and Orf19**

The products of *wbaP*, *orf8*, *orf9*, *orf10* and *orf19* are GTs, enzymes specialized on the polymerization of sugar

molecules into existing molecules, which can be carbohydrates, lipids or proteins. Because of the variety of modifications catalyzed by GTs it is difficult, based on sequence analysis alone, to define the exact outcome of each reaction [25], even though they may play an important part on the diversity of capsular structures encountered in *K. pneumoniae*. The number of GTs in *K. pneumoniae*'s *cps* cluster is variable, ranging from three (serotypes K1 and K2) to six as reported by Shu et al. [15]. Kp13 has a total of five GTs, four of these located contiguously (*wbaP*, *orf8*, *orf9* and *orf10*) and one of them found on the 3' end of the cluster (*orf19*). All the GTs found on Kp13's *cps* gene cluster have been predicted to belong to the family 2 GTs, comprising enzymes that use an inverting catalytic mechanism which modifies the anomeric configuration of the transferred sugar [26]. *wbaP* (formerly *rfbP*) is the first GT on Kp13's *cps* gene cluster and encodes a 482 aa long UDP-Gal::undecaprenolphosphate Gal-1-P transferase, which catalyzes the initial transfer of galactose-1-phosphate to an undecaprenol phosphate acceptor, thus initiating the capsule polymer synthesis. This protein was predicted to be located in the cytoplasmic membrane (PSORTb score: 10.0) and may contain five transmembrane-spanning regions. A conserved WbaP phosphotransferase domain (IPR017472, e-value 7.5e-194) is also found ranging from amino acids 21 to 482. NCBI BLASTP searches showed identity of up to 80% with WbaP from other *K. pneumoniae* and *E. coli*. The protein presents two conserved DxD motifs, which are widespread in GTs and are thought to be involved in metal/nucleotide binding and catalysis [27,28]: DED, ranging from amino acids 356–358 and DVD, 442–444 aa. The latter has been found in all but one of 12 different capsular serotypes studied by Shu et al. [15]. *orf8* (KP03804) encodes a 302 aa uncharacterized GT whose amino acid sequence shows 67% identity (Table 1) with putative rhamnosyltransferases from *K. pneumoniae* strain A1517 showing a unique capsular serotype [GenBank:BAF75773.1] [14]. The GT encoded by *orf9* (KP03803) is predicted to be 298 aa long, with a best hit on NCBI BLASTP with a putative dTDP-rhamnosyltransferase from *D. dadantii* [GenBank:ADM97617.1] (63% identity, Table 1). *D. dadantii* is a distantly related plant pathogen of the *Enterobacteriaceae* family. Interestingly, there is little similarity between *orf9* and other *K. pneumoniae* sequences. The highest identity match (31%) is with a putative rhamnosyltransferase from strain VGH484 [GenBank:BAI43783.1]. The presence of the *rmIBADC* genes (previously discussed) together with the possible rhamnosyltransferases provides appealing evidence that L-rhamnose makes part of Kp13's capsular structure. *orf10*, the third gene encoding a putative GT located in region 2 of the Kp13 *cps* cluster, is predicted to code for a 253 aa long protein with a conserved domain of unknown function spanning amino acids 36 to 193

(Pfam accession no. PF04765). As with *orf9*, the best hit (57% identity, Table 1) is also with a sequence encoding a putative GT from *D. dadantii* [GenBank: ADM97619.1]. There was no similarity between the *orf10* (KP03802) product and other published *Klebsiella* sequences.

Finally, the last GT from *cps*<sub>Kp13</sub>, termed *orf19*, is located on the 3' end of the *cps* cluster and encodes a predicted 330 aa product. This protein has similarity with several uncharacterized GTs family 2 from different *Enterobacteriaceae*, including *E. coli* TA271 [GenBank: EGI36158.1] (58% identity), *D. dadantii* [GenBank: ADM97622.1] (38%) and *Cronobacter sakazakii* [GenBank: ABX51890.1] (34%). Only a general domain of the GTs family 2 was found in this protein, spanning amino acids 7 to 145 (Pfam accession no. PF00535).

#### In silico serotyping

Using molecular serotyping for the *cps* cluster, Brisse et al. [29] showed that very distinct PCR-RFLP patterns (C patterns) were obtained for most of the K serotypes, indicating that differences in antigenic specificity among serotypes are due to differences in *cps* gene content. Thus, we have also applied in silico molecular serotyping to determine the capsular serotype of isolate Kp13. For this approach, the sequence between the primers published by Brisse et al. [29] was used to search in silico for restriction sites of the HincII endonuclease. This sequence spanned 12,031 bp from *wzi* to *gnd*, and the in silico restriction analysis identified 12 restriction sites, corresponding to 11 restriction fragments (Table 2). The fragments, ranging in size from 368 to 1,777 bp, were selected for analysis as suggested by Brisse et al. [29] (Table 2). The *cps*<sub>Kp13</sub> RFLP pattern was compared to 102 previously published C patterns [29]. None of the

reference patterns matched the one displayed by Kp13 (see Additional file 1). The similarity score for Kp13 was greater than 10.4 (MST cutoff value score  $\geq 0.75$ ), thus providing additional evidence that Kp13 K-type is a new serotype.

#### In vitro K-serotyping

Kp13 showed a weak positive reaction with both K9 and K34 antisera that could not be resolved by modifying antiserum dilution or quellung reaction. This result is not surprising since cross-reactions with the type-specific antisera is commonly observed among *K. pneumoniae* clinical isolates due to the activity of common genetic elements among distinct *cps* clusters [30]. In fact, the *rmlBADC* genes are also present in the *cps* cluster displayed by serotype K9 [15], and its CPS is composed of D-glucuronate, D-galactose and L-rhamnose residues [31]. Given the gene content of *cps*<sub>Kp13</sub> and the presence of *galE* on the Kp13 genome, these residues could all be synthesized by this isolate, hence cross-reactions were not unexpected. From the comparison of *cps*<sub>Kp13</sub> and *cps*<sub>VGH484</sub> (K9, Figure 2) it is clear that they have common genes, but the Kp13 *cps* also has distinguishing features like its repertoire of GTs, the presence of *uge-1* and a different cluster organization (e.g. the positions of *wzy* and *wzx*). In the same line of evidence, the CPS of serotype K34 is composed of L-rhamnose, D-glucose and D-galacturonate residues [32], all of which also potentially present in the Kp13 CPS as discussed earlier, and D-galacturonate being produced by the epimerase activity from the *uge-1* product. No *cps* sequences from K34 isolates were found on public databases. Nevertheless, our results indicate that Kp13 possess a unique serotype since it showed a distinct RFLP pattern compared to those 102 patterns, including representatives of serotypes K9 and

**Table 2 In silico HincII restriction pattern obtained for the 12,031 bp sequence spanning *wzi* to *gnd* in the Kp13 *cps* gene cluster**

Start	End	Cut site	Restriction fragment size between adjacent sites* (bp)
548	553	550	
1,561	1,566	1,563	<u>1,013</u>
1,638	1,643	1,640	77
2,458	2,463	2,460	<u>820</u>
2,550	2,555	2,552	92
7,129	7,134	7,131	4,579
7,260	7,265	7,262	131
7,266	7,271	7,268	6
7,634	7,639	7,636	<u>368</u>
9,411	9,416	9,413	<u>1,777</u>
10,798	10,803	10,800	<u>1,387</u>
10,863	10,868	10,865	65

\* Fragments used for this analysis are underlined.



K34, previously described [29]. It has also been observed that *cps*-PCR genotyping seems to be a more sensitive and specific way for detecting novel serotypes [14], and our pyrosequencing-based approach together with the careful scrutinization of each CDS in the cluster and the in vitro results supports the finding that Kp13 synthesizes a novel CPS.

### Regulation of *cps* gene expression in Kp13

The transcriptional regulation of *cps* genes is thought to be under the control of three promoters, P1, P2 and P3, which are located upstream of *galF*, *wzi* and *rmlB*, respectively [13,15]. As previously shown for other strains by Shu et al. [15], in the *cps*<sub>Kp13</sub> cluster the transcripts driven by P1 and P2 should consist of *galF/orf2* and *wzi* to *gnd*, respectively (Figure 4). Regulatory elements have been identified within the promoters P1 and P2 of the *cps*<sub>Kp13</sub> cluster. Promoter P1 contains the regulatory RcsAB box operator (5-TAAGATTATTCTCA-3') that is essential for the induction of Rcs-regulated promoters by way of its interaction with the regulators RcsB and RcsA [33]. Predicted *rscB* and *rscA* genes are present in the Kp13 genome, encoded, respectively, by predicted coding sequences KP00953 and KP04844.

The JUMPStart element was found within promoter P2 (Figure 4). This element was identified upstream of a number of bacterial *cps* clusters [15,34]. The 8-bp *ops* element (5'-GGCGGTAG-3') is located within JUMPStart and has been reported to function as a binding site for the RfaH activator protein [35]. Indeed, *rfaH* is found elsewhere in the Kp13 genome (KP31625), and its deduced amino acid sequence displays 80% identity with an ortholog from *E. coli* K12 [Swiss-Prot:P0AFW0]. A possible stem-loop structure (Figure 4) related to the Rho-independent transcription attenuator is located in the intergenic region between *wzc* and *wbaP* of the *cps*<sub>Kp13</sub> cluster, as predicted by the ARNold web server [36] with a calculated free energy of -8.49 kcal/mol. Similar features have also been identified in other *cps* clusters from *K. pneumoniae* [9,15]. Additionally, a second putative stem-loop structure (Figure 4) was predicted downstream of *orf10* ( $\Delta G = -8.20$  kcal/mol). Further studies are necessary to confirm the implications

of this finding; a stem-loop in this position has not been previously described. The transcription of *cps*<sub>Kp13</sub> region 3 may occur from different promoters. For instance, the P3 promoter upstream *rmlB* may transcribe a polycistronic mRNA from this gene up to *orf19* or, alternatively, each individual promoter predicted in this region may drive the transcription of a limited number of genes (Figure 4).

Notably, *wzy* is located between defective mobile elements and is transcribed in the opposite direction of other genes in the *cps* cluster (Figure 1). Thus, it should have its own promoter (possibly P7). A putative -10 box was found, separated by 15 bp from its -35 counterpart, but no obvious RBS could be identified. This observation raises the question of how Kp13 coordinates expression of *wzy*, since this protein is also essential for the formation of CPS.

Deviations from the -10 and -35 consensus sequences significantly modify the strength of each promoter [37], so the number of promoters could in fact be different from that proposed here. Still, their roles in *cps* cluster regulation deserve further experimental study, and our sequence-based bioinformatic approach provides candidates for follow-up.

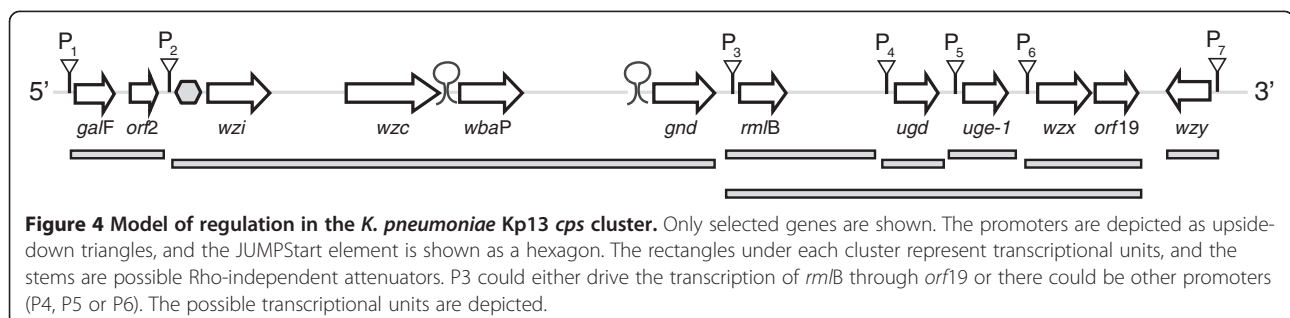
### Conclusions

In this study, we report a unique *cps* cluster organization in Kp13, a multidrug-resistant, KPC-producing *K. pneumoniae* strain that caused a large outbreak in a Brazilian teaching hospital. The Kp13 *cps* cluster contains all of the genes necessary for capsule biosynthesis. Based on the sugar metabolic pathways identified in *cps*<sub>Kp13</sub> and in other genomic regions, we have predicted that the capsule composition of Kp13 may include D-glucose, D-glucuronate, D-galacturonate, D-galactose and L-rhamnose residues.

### Methods

#### Ethics statement

This study was approved by the Ethics Committee of the Universidade Estadual de Londrina (UEL) under reference number CAAE: 3356.0.000.268-09. Clinical assessment and blood sampling were performed after diagnostic



routine procedures in the intensive care unit of the Hospital Universitário-UEL, with written informed consent of the patient.

#### Bacterial strain

Between February and May 2009, a teaching hospital located in Southern Brazil experienced its first outbreak of nosocomial infections due to KPC-producing *K. pneumoniae*. The KPC-producing *K. pneumoniae* isolate Kp13 was recovered from the blood culture of a patient admitted to the intensive care unit with diabetes mellitus and cranial encephalic trauma. Automated bacterial identification was conducted with a MicroScan WalkAway apparatus (Dade Behring, Sacramento, CA, USA). Kp13 was phenotypically detected as a carbapenemase producer by the modified Hodge [38], and the specific *bla*<sub>KPC-2</sub> gene was identified by PCR and amplicon sequencing using previously described primers and cycling conditions [39].

Kp13 was identified as *K. pneumoniae* subsp. *pneumoniae* by showing that its *rpoB* gene has 99% identity to *rpoB* of *K. pneumoniae* subsp. *pneumoniae* strain MGH 78578 [GenBank:ABR79724.1].

#### DNA sequencing, assembly and sequence analysis

Genome sequencing of Kp13 was performed at the Unidade Genômica Computacional - UGC/LNCC Facility (<http://www.labinfo.lncc.br/index.php/ugc>) located in Petrópolis, Rio de Janeiro, Brazil, using the Genome FLX sequencer (454 Life Science/Roche). Both shotgun and 3 kb paired-end libraries were constructed, and sequencing was carried out using FLX-Titanium chemistry. A paired-end (PE) library analysis was applied to determine the orientation and relative position of contigs produced by de novo shotgun sequencing. The data consisted of a total of 1,336,815 whole-genome shotgun reads and 558,997 paired-end reads.

Assembly of the sequence data into contigs and scaffolds was performed using the GS De Novo Assembler software provided by 454 Life Sciences/Roche (v 2.5). The high-quality reads were assembled into 151 contigs and 15 scaffolds, comprising 5.9 Mb of sequence. For the *cps*<sub>Kp13</sub> region from *galF* to *wzy*, 99.9% of the bases had Phred-like quality  $\geq 60$ . The SABIA annotation pipeline [40] was used to predict protein-coding genes and non-coding RNA genes. With the aim of detecting complete *cps*<sub>Kp13</sub> genes, functional annotation of the ORFs was performed by searching the NCBI non-redundant protein database using BLASTX, followed by manual curation.

Protein subcellular localizations and signal peptides were predicted using PSORTb 3.0 [41] with default parameters for Gram-negative bacteria. A score of 7.5 was considered to be the cutoff for identification of protein localization. Transmembrane regions were analyzed using

TMHMM [42]. Protein secondary structures were predicted using the PSIPRED web server [43], available at <http://bioinf.cs.ucl.ac.uk/psipred>. Prediction of promoters was performed using the in-house SABIA platform as well as the BPROM program (<http://linux1.softberry.com>), which searches for promoters under the control of the sigma factor 70. Ribosome binding sites search was performed using the RBS finder software that is included in the SABIA platform. EasyFig [44] was used to generate the structural comparison of *cps*<sub>Kp13</sub> and other sequenced *cps* loci.

#### In silico serotyping

An in silico serotyping approach was applied using the Molecular Serotyping Tool (MST) [45]. MST is a program for computer-assisted molecular identification of restriction fragment length polymorphisms (RFLP) patterns, in which the concepts of similarity and alignment between RFLP patterns were adapted from Needleman and Wunsch's dynamic programming algorithm. By analogy, RFLP patterns represented by ordered fragment sizes can be aligned, and their similarity can be calculated as the sum of penalties for edit operations (insertions, deletions or substitutions) that transform one pattern into another [45]. MST, available at [www.cebio.org/mst](http://www.cebio.org/mst), was originally designed for the identification of RFLP patterns from *Escherichia coli* and the *Shigella* O-antigen gene clusters [46,47]. At present, identification of *K. pneumoniae* serotypes can also be achieved because the RFLP patterns of the amplified capsular antigen gene clusters of all known *Klebsiella* serotypes were published by Brisse et al. [29]. The RFLP of Kp13 was determined and compared to those already described. All scores were used to build a distance matrix in a PHYLIP compatible format [48]. The distance matrix was used to reconstruct a phylogeny by the UPGMA method with the NEIGHBOR program, available in the PHYLIP package. The tree generated by UPGMA was visualized with the graphical viewer FIGTREE (<http://tree.bio.ed.ac.uk/software/figtree/>). To improve the analysis of the UPGMA tree, the two-time-scales were applied. The MST distance cutoff that is able to distinguish between two serotypes is 1.5, and the scale-adjusted measure should be interpreted as 0.75.

#### In vitro K-serotyping

Isolate Kp13 was sent to the International *Escherichia* and *Klebsiella* Reference Center (WHO), Statens Serum Institut, Copenhagen, Denmark, for serotyping. Briefly, K-typing was done by counter-current immunoelectrophoresis (CCIE) against antiserum pools as previously described [49]. Then, Kp13 was tested against with the specific K-antisera of the reacting pool. The presence of a visible capsule by wet-mount microscopy with Indian

Ink, quellung reaction, was also carried out with specific antisera since a cross-reaction had occurred.

#### Nucleotide sequence accession numbers

The *cps*<sub>Kp13</sub> sequence and annotations are available from Genbank (<http://www.ncbi.nlm.nih.gov/Genbank>) under accession number [GenBank:JN377737]. The GenBank accession numbers for other sequences discussed in the manuscript are [GenBank:JN377738] (*galE*), [GenBank:JN377739] (*galU*), [GenBank:JN377740] (*rfaH*), [GenBank:JN377741] (*rCSB*) and [GenBank:JN377742] (*rCSA*).

#### Additional file

**Additional file 1: Cluster analysis of 103 RFLP patterns after MST analysis.** MST distances between serotypes are represented as alignment scores, with 0.75 used as the scale-adjusted threshold for distinguishing two serotypes. *K. pneumoniae* Kp13 is labeled as KP13, while the other serotypes follow the C-pattern nomenclature from Brisse et al. [29].

#### Competing interests

The authors declare that there are no competing interests.

#### Authors' contributions

ECV and MP provided the Kp13 isolate and performed bacterial identification. ALG, MFN and ATRV conceived the pyrosequencing strategy. Annotation and bioinformatics analyses were performed by LGPA, LFGZ, PIPR, RCP, ACG and MFN. The manuscript was prepared by PIPR, RCP, ACG and MFN. All authors read and approved the final manuscript.

#### Acknowledgements

The authors thank Dr. Roney S. Coimbra, Dr. Fabiano S. Pais and Dr. Angela Volpini for performing the in silico serotyping. We thank Eva Møller Nielsen from the Serum Institut for their technical assistance with K-serotyping. We thank Alex Sandro Mundstein and Oberdan de Lima Cunha for carrying out the automatic genome annotation at the SABIA platform. PIPR had a Masters scholarship from Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), Brazil. ACG would like to thank the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Brazil (Process number: 307816/2009-5). MFN thanks the CNPq, Brazil (Process number: 309370/2009-4) and the Fundação Carlos Chagas Filho de Amparo à Pesquisa do Estado do Rio de Janeiro (FAPERJ), Brazil (Process number: E-26/102.214/2009). Finally, we thank the anonymous reviewers whose comments and suggestions greatly improved our manuscript.

#### Author details

<sup>1</sup>Laboratório Nacional de Computação Científica (LNCC), Petrópolis, Rio de Janeiro, Brazil. <sup>2</sup>Instituto de Microbiologia Paulo de Góes, Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brazil. <sup>3</sup>Departamento de Patologia Clínica, Análises Clínicas e Toxicologia, Universidade Estadual de Londrina, Londrina, Brazil. <sup>4</sup>Laboratório ALERTA, Divisão de Doenças Infecciosas, Universidade Federal de São Paulo, São Paulo, Brazil.

Received: 8 February 2012 Accepted: 23 May 2012

Published: 11 August 2012

#### References

1. Podschun R, Ullmann U: *Klebsiella* spp. as nosocomial pathogens: epidemiology, taxonomy, typing methods, and pathogenicity factors. *Clin Microbiol Rev* 1998, **11**:589–603.
2. Nordmann P, Cuzon G, Naas T: The real threat of *Klebsiella pneumoniae* carbapenemase-producing bacteria. *Lancet Infect Dis* 2009, **9**:228–236.
3. Greenberger MJ, Kunkel SL, Strieter RM, Lukacs NW, Bramson J, Gaudie J, Graham FL, Hitt M, Danforth JM, Standiford TJ: IL-12 gene therapy protects mice in lethal *Klebsiella pneumoniae*. *J Immunol* 1996, **157**:3006–3012.
4. Standiford TJ, Wilkowski JM, Sisson TH, Hattori N, Mehrad B, Bucknell KA, Moore TA: Intrapulmonary tumor necrosis factor gene therapy increases bacterial clearance and survival in murine gram-negative pneumonia. *Hum Gene Ther* 1999, **10**:899–909.
5. Ye P, Garvey PB, Zhang P, Nelson S, Bagby G, Summer WR, Schwarzenberger P, Shellito JE, Kolls JK: Interleukin-17 and lung host defense against *Klebsiella pneumoniae* infection. *Am J Respir Cell Mol Biol* 2001, **25**:335–340.
6. Tumbarello M, Spanu T, Sanguinetti M, Citton R, Montuori E, Leone F, Fadda G, Cauda R: Bloodstream infections caused by extended-spectrum-beta-lactamase-producing *Klebsiella pneumoniae*: risk factors, molecular epidemiology, and clinical outcome. *Antimicrob Agents Chemother* 2006, **50**:498–504.
7. Roberts IS: The biochemistry and genetics of capsular polysaccharide production in bacteria. *Annu Rev Microbiol* 1996, **50**:285–315.
8. Sahly H, Keisari Y, Crouch E, Sharon N, Ofek I: Recognition of bacterial surface polysaccharides by lectins of the innate immune system and its contribution to defense against infection: the case of pulmonary pathogens. *Infect Immun* 2008, **76**:1322–1332.
9. Rahn A, Drummelsmith J, Whitfield C: Conserved organization in the *cps* gene clusters for expression of *Escherichia coli* group 1 K antigens: relationship to the colanic acid biosynthesis locus and the *cps* genes from *Klebsiella pneumoniae*. *J Bacteriol* 1999, **181**:2307–2313.
10. Whitfield C, Roberts IS: Structure, assembly and regulation of expression of capsules in *Escherichia coli*. *Mol Microbiol* 1999, **31**:1307–1319.
11. Whitfield C, Paiment A: Biosynthesis and assembly of Group 1 capsular polysaccharides in *Escherichia coli* and related extracellular polysaccharides in other bacteria. *Carbohydr Res* 2003, **338**:2491–2502.
12. Whitfield C: Biosynthesis and assembly of capsular polysaccharides in *Escherichia coli*. *Annu Rev Biochem* 2006, **75**:39–68.
13. Arakawa Y, Wacharotayankun R, Nagatsuka T, Ito H, Kato N, Ohta M: Genomic organization of the *Klebsiella pneumoniae cps* region responsible for serotype K2 capsular polysaccharide synthesis in the virulent strain Chedid. *J Bacteriol* 1995, **177**:1788–1796.
14. Pan YJ, Fang HC, Yang HC, Lin TL, Hsieh PF, Tsai FC, Keynan Y, Wang JT: Capsular polysaccharide synthesis regions in *Klebsiella pneumoniae* serotype K57 and a new capsular serotype. *J Clin Microbiol* 2008, **46**:2231–2240.
15. Shu HY, Fung CP, Liu YM, Wu KM, Chen YT, Li LH, Liu TT, Kirby R, Tsai SF: Genetic diversity of capsular polysaccharide biosynthesis in *Klebsiella pneumoniae* clinical isolates. *Microbiology* 2009, **155**:4170–4183.
16. Favre C, Passet V, Deletoile A, Barbe V, Frangeul L, Almeida AS, Sansonetti P, Tournebise R, Brisse S: PCR-based identification of *Klebsiella pneumoniae* subsp. *rhinoscleromatis*, the agent of rhinoscleroma. *PLoS Negl Trop Dis* 2011, **5**:e1052.
17. Ho JY, Lin TL, Li CY, Lee A, Cheng AN, Chen MC, Wu SH, Wang JT, Li TL, Tsai MD: Functions of some capsular polysaccharide biosynthetic genes in *Klebsiella pneumoniae* NTUH K-2044. *PLoS One* 2011, **6**:e21664.
18. Regue M, Hita B, Pique N, Izquierdo L, Merino S, Fresno S, Benedi VJ, Tomas JM: A gene, *uge*, is essential for *Klebsiella pneumoniae* virulence. *Infect Immun* 2004, **72**:54–61.
19. Giraud MF, Naismith JH: The rhamnose pathway. *Curr Opin Struct Biol* 2000, **10**:687–696.
20. Rahn A, Beis K, Naismith JH, Whitfield C: A novel outer membrane protein, *Wzi*, is involved in surface assembly of the *Escherichia coli* K30 group 1 capsule. *J Bacteriol* 2003, **185**:5882–5890.
21. Lin MH, Hsu TL, Lin SY, Pan YJ, Jan JT, Wang JT, Khoo KH, Wu SH: Phosphoproteomics of *Klebsiella pneumoniae* NTUH-K2044 reveals a tight link between tyrosine phosphorylation and virulence. *Mol Cell Proteomics* 2009, **8**:2613–2623.
22. Cuthbertson L, Mainprize IL, Naismith JH, Whitfield C: Pivotal roles of the outer membrane polysaccharide export and polysaccharide copolymerase protein families in export of extracellular polysaccharides in gram-negative bacteria. *Microbiol Mol Biol Rev* 2009, **73**:155–177.
23. Marolda CL, Li B, Lung M, Yang M, Hanuszkiwicz A, Rosales AR, Valvano MA: Membrane topology and identification of critical amino acid residues in the *Wzx* O-antigen translocase from *Escherichia coli* O157:H4. *J Bacteriol* 2010, **192**:6160–6171.
24. Nakhamchik A, Wilde C, Rowe-Magnus DA: Identification of a *Wzy* polymerase required for group IV capsular polysaccharide and

- lipopolysaccharide biosynthesis in *Vibrio vulnificus*. *Infect Immun* 2007, **75**:5550–5558.
25. Coutinho PM, Deleury E, Davies GJ, Henrissat B: **An evolving hierarchical family classification for glycosyltransferases.** *J Mol Biol* 2003, **328**:307–317.
26. Liu J, Mushegian A: **Three monophyletic superfamilies account for the majority of the known glycosyltransferases.** *Protein Sci* 2003, **12**: 1418–1431.
27. Hurtado-Guerrero R, Zusman T, Pathak S, Ibrahim AF, Shepherd S, Prescott A, Segal G, van Aalten DM: **Molecular mechanism of elongation factor 1A inhibition by a *Legionella pneumophila* glycosyltransferase.** *Biochem J* 2010, **426**:281–292.
28. Unligil UM, Rini JM: **Glycosyltransferase structure and mechanism.** *Curr Opin Struct Biol* 2000, **10**:510–517.
29. Brisse S, Issenhuth-Jeanjean S, Grimont PA: **Molecular serotyping of *Klebsiella* species isolates by restriction of the amplified capsular antigen gene cluster.** *J Clin Microbiol* 2004, **42**:3388–3398.
30. Murcia A, Rubin SJ: **Reproducibility of an indirect immunofluorescent-antibody technique for capsular serotyping of *Klebsiella pneumoniae*.** *J Clin Microbiol* 1979, **9**:208–213.
31. Lindberg B, Lonngren J, Thompson JL: **Structural studies of the *Klebsiella* type 9 capsular polysaccharide.** *Carbohydr Res* 1972, **25**:49–57.
32. Joseleau JP, Michon F, Vignon M: **Structural investigation of the capsular polysaccharide from *Klebsiella* serotype K-34 and its characterization by N.M.R. spectroscopy.** *Carbohydr Res* 1982, **101**:175–185.
33. Wehland M, Bernhard F: **The RcsAB box. Characterization of a new operator essential for the regulation of exopolysaccharide biosynthesis in enteric bacteria.** *J Biol Chem* 2000, **275**:7013–7020.
34. Nakhmchik A, Wilde C, Chong H, Rowe-Magnus DA: **Evidence for the horizontal transfer of an unusual capsular polysaccharide biosynthesis locus in marine bacteria.** *Infect Immun* 2010, **78**:5214–5222.
35. Bailey MJ, Hughes C, Koronakis V: **In vitro recruitment of the RfaH regulatory protein into a specialised transcription complex, directed by the nucleic acid ops element.** *Mol Gen Genet* 2000, **262**:1052–1059.
36. Naville M, Ghuillot-Gaudeffroy A, Marchais A, Gautheret D: **ARNold: a web tool for the prediction of Rho-independent transcription terminators.** *RNA Biol* 2011, **8**:11–13.
37. Hawley DK, McClure WR: **Compilation and analysis of *Escherichia coli* promoter DNA sequences.** *Nucleic Acids Res* 1983, **11**:2237–2255.
38. Clinical and Laboratory Standards Institute: *Performance standards for antimicrobial susceptibility testing. 21st informational supplement.* Wayne, Pa: Clinical and Laboratory Standards; 2011.
39. Woodford N, Tierno PM Jr, Young K, Tysall L, Palepou MF, Ward E, Painter RE, Suber DF, Shungu D, Silver LL, Inglima K, Kornblum J, Livermore DM: **Outbreak of *Klebsiella pneumoniae* producing a new carbapenem-hydrolyzing class A beta-lactamase, KPC-3, in a New York Medical Center.** *Antimicrob Agents Chemother* 2004, **48**:4793–4799.
40. Almeida LG, Paixao R, Souza RC, Costa GC, Barrientos FJ, Santos MT, Almeida DF, Vasconcelos AT: **A System for Automated Bacterial (genome) Integrated Annotation–SABIA.** *Bioinformatics* 2004, **20**:2832–2833.
41. Yu NY, Wagner JR, Laird MR, Melli G, Rey S, Lo R, Dao P, Sahinalp SC, Ester M, Foster LJ, Brinkman FS: **PSORTb 3.0: improved protein subcellular localization prediction with refined localization subcategories and predictive capabilities for all prokaryotes.** *Bioinformatics* 2010, **26**: 1608–1615.
42. Krogh A, Larsson B, von Heijne G, Sonnhammer EL: **Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes.** *J Mol Biol* 2001, **305**:567–580.
43. Jones DT: **Protein secondary structure prediction based on position-specific scoring matrices.** *J Mol Biol* 1999, **292**:195–202.
44. Sullivan MJ, Petty NK, Beatson SA: **Easyfig: a genome comparison visualizer.** *Bioinformatics* 2011, **27**:1009–1010.
45. Coimbra RS, Artiguenave F, Jacques LS, Oliveira GC: **MST (molecular serotyping tool): a program for computer-assisted molecular identification of *Escherichia coli* and *Shigella* O antigens.** *J Clin Microbiol* 2010, **48**:1921–1923.
46. Coimbra RS, Grimont F, Grimont PA: **Identification of *Shigella* serotypes by restriction of amplified O-antigen gene cluster.** *Res Microbiol* 1999, **150**:543–553.
47. Coimbra RS, Grimont F, Lenormand P, Burguiere P, Beutin L, Grimont PA: **Identification of *Escherichia coli* O-serogroups by restriction of the amplified O-antigen gene cluster (rfb-RFLP).** *Res Microbiol* 2000, **151**: 639–654.
48. Felsenstein J: **PHYLIP - Phylogeny Inference Package (version 3.2).** *Cladistics* 1989, **5**:164–166.
49. Hansen DS, Skov R, Benedi JV, Sperling V, Kolmos HJ: ***Klebsiella* typing: pulsed-field gel electrophoresis (PFGE) in comparison with O:K-serotyping.** *Clin Microbiol Infect* 2002, **8**:397–404.

doi:10.1186/1471-2180-12-173

**Cite this article as:** Ramos *et al.*: Pyrosequencing-based analysis reveals a novel capsular gene cluster in a KPC-producing *Klebsiella pneumoniae* clinical isolate identified in Brazil. *BMC Microbiology* 2012 **12**:173.

**Submit your next manuscript to BioMed Central and take full advantage of:**

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at  
[www.biomedcentral.com/submit](http://www.biomedcentral.com/submit)

