Klebsiella variicola and *Klebsiella quasipneumoniae* with capacity to adapt to clinical and plant settings

Esperanza Martínez-Romero, PhD,[1] Nadia Rodríguez-Medina, MSc,[2]
Marilú Beltrán-Rojel, Biol,[2] Jeiry Toribio-Jiménez, PhD,[3] Ulises Garza-Ramos, PhD.[2]

Martínez-Romero E, Rodríguez-Medina N, Beltrán-Rojel M, Toribio-Jiménez J, Garza-Ramos U. *Klebsiella variicola* and *Klebsiella quasipneumoniae* with capacity to adapt to clinical and plant settings. Salud Publica Mex 2018;60:29-40.
http://doi.org/10.21149/8156

Abstract

Objective. To compare the genetic determinants involved in plant colonization or virulence in the reported genomes of *K. variicola, K. quasipneumoniae* and *K. pneumoniae*. **Materials and methods.** *In silico* comparisons and Jaccard analysis of genomic data were used. Fimbrial genes were detected by PCR. Biological assays were performed with plant and clinical isolates. **Results.** Plant colonization genes such as cellulases, catalases and hemagglutinins were mainly present in *K. variicola* genomes. Chromosomal β-lactamases were characteristic of this species and had been previously misclassified. *K. variicola* and *K. pneumoniae* isolates produced plant hormones. **Conclusions.** A mosaic distribution of different virulence- and plant-associated genes was found in *K. variicola* and in *K. quasipneumoniae* genomes. Some plant colonizing genes were found mainly in *K. variicola* genomes. The term plantanosis is proposed for plant-borne human infections.

Keywords: bacterial infections; Gram-negative bacterial infections; *Enterobacteriaceae* infections

Martínez-Romero E, Rodríguez-Medina N, Beltrán-Rojel M, Toribio-Jiménez J, Garza-Ramos U. *Klebsiella variicola* y *Klebsiella quasipneumoniae* con capacidad para adaptarse al ambiente clínico y a las plantas. Salud Publica Mex 2018;60:29-40.
http://doi.org/10.21149/8156

Resumen

Objetivo. Comparar genes de colonización de plantas o de virulencia en los genomas reportados de *K. variicola*, *K. quasipneumoniae* y *K. pneumoniae*. **Material y métodos.** Se utilizaron análisis *in silico* y de Jaccard. Por PCR se detectaron genes de fimbrias. Se realizaron ensayos biológicos con aislados de plantas y clínicos. **Resultados.** Los genes de colonización de plantas como celulasas, catalasas y hemaglutininas se encontraron principalmente en genomas de *K. variicola*. Las β-lactamasas cromosómicas son características de la especie y en algunos casos estaban mal clasificadas. *K. variicola* y *K. pneumoniae* producen hormonas vegetales. **Conclusiones.** Se encontró una distribución en mosaico de los genes de asociación con plantas y de virulencia en *K. variicola* y *K. quasipneumoniae*. Principalmente en *K. variicola* se encontraron algunos genes involucrados en la colonización de plantas. Se propone el término plantanosis para las infecciones humanas de origen vegetal.

Palabras clave: infecciones bacterianas; infecciones de bacterias Gram-negativas; infecciones *Enterobacteriaceae*

(1) Centro de Ciencias Genómicas, Universidad Nacional Autónoma de México. Cuernavaca, Morelos, Mexico.
(2) Laboratorio de Resistencia Bacteriana, Centro de Investigación sobre Enfermedades Infecciosas, Instituto Nacional de Salud Pública. Cuernavaca, Morelos, México.
(3) Laboratorio de Biotecnología y Genética Microbiana, Unidad Académica de Ciencias Químico Biológicas, Universidad Autónoma de Guerrero. Chilpancingo, Guerrero, México.

Received on: August 22, 2016 • **Accepted on:** February 24, 2017
Corresponding author. PhD. Ulises Garza-Ramos. Departamento de Diagnóstico Epidemiológico, Instituto Nacional de Salud Pública.
Av. Universidad 655 col. Santa María Ahuacatitlán. 62100, Cuernavaca, Morelos, Mexico.
E-mail: ulises.garza@insp.mx

The taxonomy of the genus *Klebsiella* has been periodically revised. Currently, this genus includes *K. pneumoniae* subsp. *pneumoniae*, also known as KpI, and four novel species: *Klebsiella quasipneumoniae*, also known as KpII - with two subspecies: *quasipneumoniae* (KpII-A) and *similipneumoniae* (KpII-B); *Klebsiella variicola* (KpIII) and *Klebsiella michiganensis*.[1] *K. variicola* and *K. quasipneumoniae* are sister species of *K. pneumoniae*. *K. variicola* has been isolated from plant tissues,[2,3] fungal gardens of leaf-cutter ant colonies,[4] cotton disease vectors (specifically of the insect, *Nezara viridula*),[5] human[6-10] and animal infections[11] including bovine mastitis.[12] Unlike *K. variicola*, *K. quasipneumoniae* has been described exclusively in hospital settings.[13] A recent genomic comparative study with a large number of *Klebsiella* genomes showed that the *nif* operon was detected in all the genomes of *K. variicola*, in half of those of *K. quasipneumoniae*, but only in one strain of *K. pneumoniae*.[13] Additionally, *K. pneumoniae* and *K. variicola* shared some virulence determinants that cause infections in humans.[14,15]

Some *Klebsiella* isolates have been misclassified[16,17] and consequently the taxonomic descriptions of their genome sequences in NCBI repositories are inaccurate. Here we used correctly identified *K. variicola* and *K. quasipneumoniae* genomes for the comparison of virulence and plant colonization determinants of this genus of enterobacteria.

Materials and methods

K. variicola and *K. quasipneumoniae* genomes included in the study

A total of 41 *Klebsiella* genomes were included in the study, 31 of which corresponded to *K. variicola*, eight to *K. quasipneumoniae* and two to *K. pneumoniae* (table I). Of these genomes, 19 of *K. variicola* and six *K. quasipneumoniae* were originally misidentified.

In silico analysis of virulence, plant-associated determinants and efflux pumps, regulators, heavy metal resistance and β-lactamases

A set of 138 proteins was selected from previous reports[18,19] and the Pasteur Institute webpage (http://bigsdb.web.pasteur.fr/klebsiella/klebsiella.html) for their putative roles in host interactions. Among these proteins, 84 were included for their involvement in virulence. Two efflux pump clusters, 11 efflux pump regulators, four heavy metal clusters, three β-lactamase family proteins (LEN-, SHV- and OKP-type), as well as 35 plant-association proteins that included nitrogen fixation enzymes (encoded by the *nif* operon with 20 genes), were also included. All genes encoding these proteins/enzymes were further searched in 41 *Klebsiella* spp. genomes by BLAST20 (Genome BLASTp option with default values). A cluster analysis using the average linkage based on the Jacquard similarity coefficient and the respective dendrogram (constructed with UPGMA) were constructed using the DendroUPGMA program.[20] In the case of the LEN-, SHV- and OKP-type families of β-lactamases, a phylogenic reconstruction was performed using the maximum-likelihood approach based on the JTT matrix-based model and 100 bootstrap replications (Mega v6.06).[21] In addition, phylogenetic analyses were carried out on the amino acid sequences of FimV fimbrial proteins (KVR801v1_60088) and NifH nitrogenase (KVR801v1_120019) proteins.

PCR screening of *fimV* gene

Specific oligonucleotides for the *fimV* gene (fimV-F 5′-TTTGCGGATACTGACCAGGG-3′ and fimV-R 5′-GGTTACCACGGTCAGCGTAA-3′) were designed. Twenty-one *K. variicola* isolates from plants and humans were analyzed by PCR as previously described.[8]

In vitro assays of plant growth promotion

The mechanisms involved in plant growth promotion, such as phytohormone production (auxin [indole-3-acetic acid] and gibberellins), phosphate solubilization, siderophore production and lytic enzyme activities were evaluated in six *K. variicola* isolated from plants, and in 15 *K. variicola* and eight *K. pneumoniae* clinical isolates. The production of auxin was analyzed according to Khalid and collagues[22] with *Azotobacter vinelandii* and *Salmonella enteritidis* as positive and negative controls, respectively. Phosphate solubilization was evaluated using the methodology described by Mehta and colleagues[23] using *Azotobacter vinelandii* and *E. coli* DH5-α as bacterial controls. Siderophore production was evaluated using chrome azurol S according to Schwyn and Neilands[24] with *Pseudomonas fluorescens* and *Staphylococcus aureus* as controls. Activities of lytic enzymes, such as lipases, proteases, esterases and amylases, were evaluated using the methodology described by Malleswari and Bagyanarayana[25] using *Bacillus subtilis* and *E. coli* DH5-α as controls.

Ethical committee approval

This study is part of the *K. variicola* project that was revised and approved by the ethical commission from the *Instituto Nacional de Salud Pública* on the 21th of June 2011. The present study was carried out at the *Instituto*

Table I

ORIGIN AND ACCESSION NUMBER DATA OF *K. VARIICOLA*, *K. QUASIPNEUMONIAE* AND *K. PNEUMONIAE* GENOMES INCLUDED IN THE STUDY

Bacterial specie on GenBank	*True bacterial specie**	*Isolate*	*Origin of isolates*	*Accession number data base*	*PUBMED (PMID)*
K. pneumoniae	*K. pneumoniae*	NTUH-K2044	Human (blood)	AP006725.1	19447910
K. pneumoniae	*K. pneumoniae*	MGH-78578	Human	CP000647.1	Unpublished
K. pneumoniae	*K. variicola*	223/14	Human (pus)	JRTV00000000	Unpublished
K. pneumoniae	*K. variicola*	B1	Plant (bitter gourd)	JSWX00000000	Unpublished
K. pneumoniae	*K. variicola*	CH4	Plant (chili)	JSXA00000000	Unpublished
K. pneumoniae	*K. variicola*	MGH20	Human (respiratory)	AYJK00000000	Unpublished
K. pneumoniae	*K. variicola*	MGH40	Unknown (urine)	AYIX00000000	Unpublished
K. pneumoniae	*K. variicola*	UCICRE10	Unknown	AYIF00000000	Unpublished
K. pneumoniae	*K. variicola*	UCI18	Human (urine)	JCML00000000	Unpublished
K. pneumoniae	*K. variicola*	BIDMC61	Human (urine)	JMWB00000000	Unpublished
K. pneumoniae	*K. variicola*	MGH68	Human (urine)	JMZD00000000	Unpublished
K. pneumoniae	*K. variicola*	MGH76	Human (bile)	JMZK00000000	Unpublished
K. pneumoniae	*K. variicola*	MGH80	Human (urine)	JMZM00000000	Unpublished
K. pneumoniae	*K. variicola*	BIDMC88	Human	LFBA00000000	Unpublished
K. pneumoniae	*K. variicola*	BIDMC90	Human	LFBC00000000	Unpublished
K. pneumoniae	*K. variicola*	MGH114	Human	LFAP00000000	Unpublished
K. pneumoniae	*K. variicola*	MGH92	Human	LFAD00000000	Unpublished
Klebsiella sp.	*K. variicola*	1.1.55	Human	ACXA00000000	Unpublished
Klebsiella sp.	*K. variicola*	KTE92	Human	ASQN00000000	Unpublished
K. pneumoniae	*K. variicola*	342	Plant (maize stems)	CP000964.1	18654632
K. pneumoniae	*K. variicola*	KP5-1	Insect (*Nezara viridula*)	CP008700.1	25146146
K. variicola	*K. variicola*	BZ19	Human (faeces)	JDWA00000000	25135672
K. variicola	*K. variicola*	DX120E	Plant (banana roots)	CP009274	Unpublished
K. variicola	*K. variicola*	DSM15968	Plant (banana roots)	CP010523	Unpublished
K. variicola	*K. variicola*	CAG:634	Human (gut)	CBBA000000000	Unpublished
K. variicola	*K. variicola*	801	Human (blood)	CDMV00000000	25886267
K. variicola	*K. variicola*	8917	Human (sputum)	CEGG00000000	25858850
K. variicola	*K. variicola*	06-268	Human (abscess)	CXOZ00000000	This work
K. variicola	*K. variicola*	3	Plant (maize shoots)	CXOY00000000	26358599
K. variicola	*K. variicola*	4880	Human (blood)	CXPB00000000	This work
K. variicola	*K. variicola*	6A2	Plant (banana root)	CXPC00000000	26358599
K. variicola	*K. variicola*	T29A	Plant (sugar cane stem)	CXPA00000000	26358599
K. variicola	*K. variicola*	At-22	Insect (fungus matrix)	CP001891.1	19965433
K. quasipneumoniae‡	*K. quasipneumoniae*	18A069	Human	CBZM000000000	24958762
K. quasipneumoniae§	*K. quasipneumoniae*	07A044	Human	CBZR000000000	24958762
K. pneumoniae	*K. quasipneumoniae*‡	UCICRE14	Unknown	AYIC00000000	Unpublished
K. pneumoniae	*K. quasipneumoniae*§	12-3578	Human (blood)	AQOC00000000	Unpublished
K. pneumoniae	*K. quasipneumoniae*§	ATCC 700603	Human	AOGO00000000	23723407
K. pneumoniae	*K. quasipneumoniae*§	MGH44	Unknown (respiratory)	AYIV00000000	Unpublished
K. variicola	*K. quasipneumoniae*§	HKUOPLA	Panda (feces)	CP012252	26472841
K. pneumoniae	*K. quasipneumoniae*§	HKUOPLC	Panda (feces)	CP012300	26564041

* The true bacterial specie was described previously by Chen and colleagues,[16] and Martinez-Romero and colleagues.[17]

‡ *Klebsiella quasipneumoniae* subsp. *quasipneumoniae*

§ *Klebsiella quasipneumoniae* subsp. *similipneumoniae*

Nacional de Salud Pública and at the *Centro de Ciencias Genómicas* of UNAM, Cuernavaca, Morelos.

Results and discussion

Virulence-associated determinants in *K. variicola and K. quasipneumoniae* genomes

In silico screening of 84 virulence-determinant proteins in reported genomes of *K. variicola*, *K. quasipneumoniae* and *K. pneumoniae* showed a mosaic distribution in different isolates (table I). Most (>98%, except for FimV) of the virulence determinants that were selected here were originally described in *K. pneumoniae* NTUH-K2044. From the 84 virulence determinants included in the *in silico* analysis, 20 were found in at least one *K. variicola* genome and 11 in at least one *K. quasipneumoniae* genome. The proteins coded by the urease (*UreA*) and fimbriae gene cluster *MrkABCDFHIJ* were present in all the isolates of *K. variicola* and *K. quasipneumoniae*. The gene coding glucuronic acid transferase (WabG) was also present in all *K. variicola* and *K. pneumoniae*, but was absent in a single *K. quasipneumoniae* genome.

Siderophores, such as salmochelin (IroN), aerobactin (IucA and receptor IutA), KfuABC cluster, enterobactin (EntB) and yersiniabactin cluster YbtAESTX, were unequally distributed among *Klebsiella*. The most prevalent siderophores were enterobactin and KfuABC, in 93.5% (29/31) of *K. variicola* and in 100% (8/8) of the *K. quasipneumoniae* isolates. The aerobactin (IucA) was not identified, however the receptor of aerobactin (IutA) was present in 87% (27/31) and salmochelin (IroN) in 39% (12/31) of *K. variicola* isolates; however, in the *K. quasipneumoniae* genomes, both aerobactin receptor and salmochelin siderophore were present in 100% of the isolates. The siderophore yersiniabactin YbtAESTX cluster was identified in only one *K. variicola* isolate (MGH20) and in none of the *K. quasipneumoniae* genomes.

The genes encoding enzymes from the glycerate pathway (Gcl, GlxK), GlxR, and Hyi) were identified in *K. variicola* isolates (41.9 to 45.1%). The allantoinase cluster *AllABCDRS* genes and the YlbE-F and YbbW were present in few *K. variicola* genomes at 3.2, 6.4 and 3.2%, respectively. The GlxK-R and Hyi encoding genes were present only in *K. quasipneumoniae* genomes, in 75% of them. The genes for the two-component system KvgAS proteins were identified in 12.9% (4/31) of the *K. variicola* genomes and were absent from the *K. quasipneumoniae* genomes. The mucoviscosity-associated protein Wzy-K1 was identified in one *K. variicola* isolate (3.2%-1/31).

The Jacquard index analysis of the virulence-associated determinants that are shared between the *K. variicola*, *K. quasipneumoniae* and *K. pneumoniae* genomes is shown in figure 1. Three main clusters were obtained, two of which were closely related and could be distinguished only by a single siderophore difference. Cluster 1 grouped sixteen *K. variicola* and two *K. quasipneumoniae* subsp. *similipneumoniae* genomes that shared the siderophores IroN (11/18 genomes); receptor IutA, KfuABC and EntB (15/18 genomes); WabG, Uge, the fimbriae cluster MrkABCDFHIJ, FimV (15/18 genomes) and urease UreA. Cluster 2 includes ten *K. variicola* genomes and mostly contains the virulence-determinants proteins of cluster 1, except for IroN. Genes associated with all of the *K. variicola* genomes from cluster 2 were those encoding Glc, GlxKR and Hyi proteins that are involved in the glycerate pathway. This pathway together with allantoinase cluster genes are involved in allantoin metabolism as a nitrogen source.[26] Cluster 3 consists of seven *K. quasipneumoniae* and one *K. variicola* genomes. All of the genomes of cluster 3 contain the Uge, WabG, IroN (except *K. variicola* KTE92 genome), IutA, KfuABC, EntB, UreA and MrkABCDFHIJ proteins. FimV protein was present in three (42.8%-3/7) of the *K. quasipneumoniae* genomes, corresponding to *Klebsiella quasipneumoniae* subsp. *similipneumoniae* (KpII-B).

K. pneumoniae NTUH-K2044 and *K. pneumoniae* MGH78578 each independently clustered together with *K. pneumoniae* BIDMC90 and *K. variicola* DSM15298, respectively (figure 1). The *K. pneumoniae* BIDMC90 genome shared with *K. pneumoniae* NTUH-K2044 fifteen virulence determinants. Finally, *K. pneumoniae* MGH78578 and *K. variicola* DSM15298 were different from the other clusters due to the absence of siderophore KfuABC, which was present in all of the *K. variicola, K. quasipneumoniae* and *K. pneumoniae* NTUH-K2044, and to the presence of the two-component system KvgAS, which was only present in the DX120E, MGH76 and UCICRE10 genomes (figure 1).

In silico analysis and prevalence of Fimbrial FimV protein

Previously a total of 114 unique proteins were identified in *K. variicola* genomes.[8] Excluding transposons and *tra*-genes that are involved in horizontal transfer and selecting those that are involved in metabolism and cellular structure, 13 structural proteins were found to be unique for *K. variicola*, from which a fimbrial protein was chosen for further analysis (fimbria have been shown to mediate host interaction). The fimbrial FimV protein was identified in 90.3% (28/31) of the *K. variicola* (except for KTE92, B1 and BIDMC90) and in 37.5% (3/8) of *K. quasipneumoniae* but not in *K. pneumoniae* NTUH-K2044 and MGH78578 genomes. An additional BLASTp search in GenBank

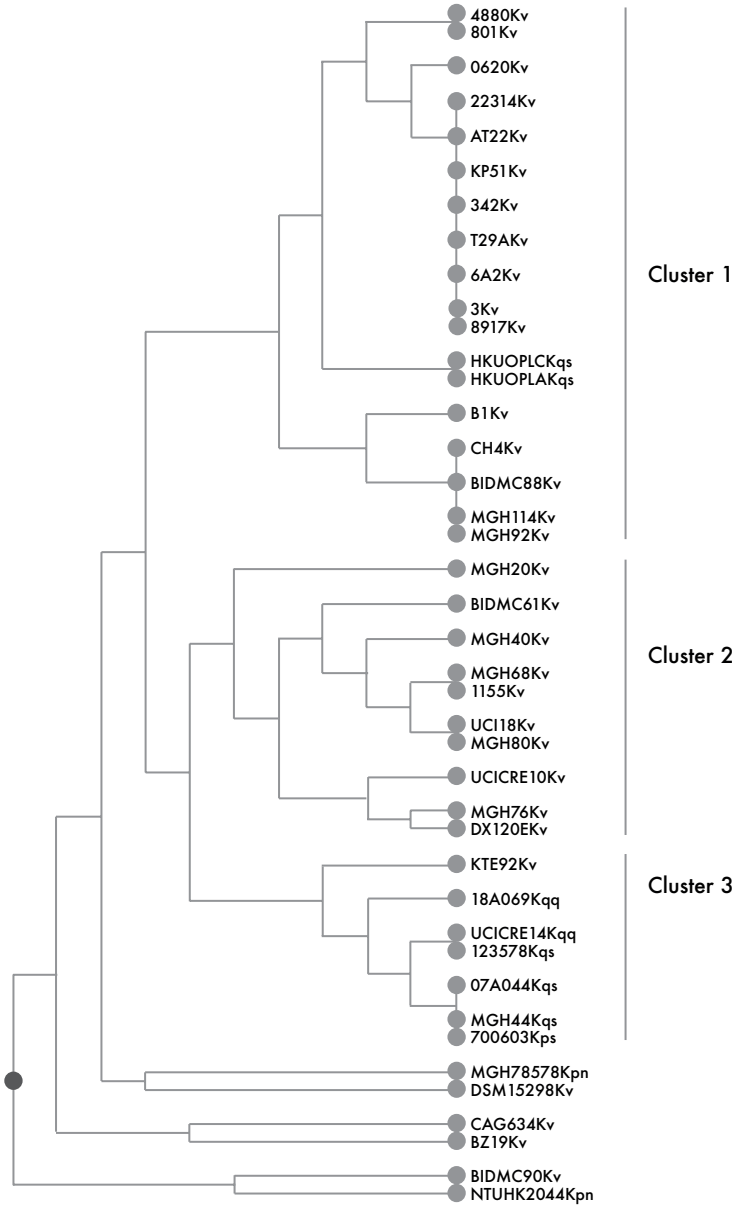

FIGURE I. JACQUARD INDEX OF THE VIRULENCE-DETERMINANT PROTEINS OF THE *K. VARIICOLA* (KV), *K. QUASIPNEUMONIAE* (KQ) AND *K. PNEUMONIAE* (KPN) GENOMES

using a non-redundant protein sequence (nr-BLASTp protein-protein BLAST) identified the FimV protein in *K. variicola* genomes described in GenBank (except for KTE92, B1 and BIDMC90) and in the misclassified *K. pneumoniae* genomes that actually corresponded to *K. variicola* species (table I). The phylogenetic analysis of FimV proteins generally showed high amino acid conservation in *K. variicola* (>98%) (data not shown); however, in a few *K. variicola* strains, for example the wrongly named *"K. pneumoniae* 342*"*, a lower amino acid identity was observed (92 to 94%). Conversely, *K. quasipneumoniae* subsp. *similipneumoniae*, 07A044, 700603 and MGH44 FimV amino acid sequence had 91% similarity. FimV protein sequence was not found in *bona fide K. pneumoniae*; nevertheless it was identified in related *Enterobacteriaceae* such as *Citrobacter koseri*,

Citrobacter freundii, *Enterobacter*, *Salmonella enterica* and *Escherichia albertii*, all with an amino acid similarity less than 53%. The FimV protein is also widely distributed in *Escherichia coli* and *Shigella sonnei*, with an amino acid identity of 33% to that from *K. variicola*.

The PCR screening of the *fimV* gene was carried out in a collection of twenty-one *K. variicola* isolates (see Materials and methods). The *fimV* gene was identified in 86.9% (20/23) of the *K. variicola* isolates. This prevalence is similar to that from *in silico* analysis (90.3%). This indicates that the *fimV* gene is not universally encoded in *K. variicola* genomes.

Plant-associated proteins in the *K. variicola*, *K. quasipneumoniae* and *K. pneumoniae* genomes

Plant-associated proteins were identified in *K. variicola* and *K. quasipneumoniae* isolates. Twenty-seven plant-associated determinants were found in at least one *K. variicola* or *K. quasipneumoniae* isolate. In general, 21 plant-associated determinants were contained in all of the *K. variicola*, *K. quasipneumoniae* and *K. pneumoniae* genomes that were included in the study. The NifJ-NifQ nitrogen fixation cluster was identified in all of the *K. variicola* genomes, in 62.5% (5/8) of *K. quasipneumoniae* genomes and was absent from the *K. pneumoniae* NTUH-K2044 and MGH78578 genomes. The NifH protein has high amino acid identity among all *K. variicola* genomes and in a phylogenetic analysis it is grouped in a cluster different from the corresponding sequence from the *K. quasipneumoniae* or other species. Holt and colleagues[13] identified only one *K. pneumoniae* genome that contained the *nifJ-nifQ* operon. In this work, we identified another *K. pneumoniae* genome (from KPNIN29) that contains the *nif* operon in addition to that described by Holt *and colleagues*.[13] Hazen *and colleagues*[27] identified the absence of *nifJ-nifQ* gene cluster in *K. pneumoniae* NTUH-K2044, MGH-78578, 1162281, JH1, MS 92-3, 1191100241, ATCC13884 and KCTC-2242 genomes. Here we identified a conserved genetic context and gene synteny of *nifJ-nifQ* operon both in *K. variicola* and *K. quasipneumoniae* genomes (data not shown). Our results confirm that nitrogen fixation seems to be a characteristic trait of *K. variicola*.

Other differences identified in this work between the *K. variicola*, *K. quasipneumoniae* and *K. pneumoniae* genomes are genes encoding cellulases (CelK and BglH), catalases (KPK_4954 and KPK_2333) and hemagglutinins (HecA). While BglX cellulose is present in all *K. variicola*, *K. quasipneumoniae* and *K. pneumoniae* isolates, CelK and BglH were absent from *K. pneumoniae* NTUH-K2044 and MGH78578 and from 87.5% of the *K. quasipneumoniae* isolates.

The BglH enzyme, which has specificity towards 1,4-b glucosidic bonds and that most likely acts by hydrolyzing short cello-oligosaccharides, was present in 61.2% (19/31) of *K. variicola* isolates. The CelK gene encoding an enzyme for the decomposition of highly ordered forms of insoluble cellulose was present in *K. variicola* strains obtained from three different plants 342, T29A and 6A2 representing 10% (3/31) of the isolates. KPK_2233 (catalase) gene was present in 22.5% of the *K. variicola* genomes (7/31). KPK_4954 gene encoding a cyclic beta 1-2 glycan synthase possibly playing a role in osmotic adaptation was present in 16.1% of the *K. variicola* isolates (5/31).

HecA/B hemolysin/hemagglutinin secretion protein is involved in plant attachment in *Erwinia chrysanthemis*. A *hecA* gene mutant had reduced attachment, cell aggregate formation, and virulence in its host. HecA protein was present in 64.5% (20/31) of the *K. variicola* isolates. In *K. quasipneumoniae*, HecA protein is present only in HKUOPLA[28] and HKUOPLC[29] isolates obtained from giant panda feces.

The Jacquard index analysis showed two main clusters (A and B) that were grouped by the absence or presence of HecA, CelK or BglH protein (figure 2). Cluster A included the *K. variicola* and *K. quasipneumoniae* genomes that did not contain the HecA and CelK proteins. The *K. quasipneumoniae* 18A069 genome is the unique genome of this species that does not contain the YqeF (acetyl-CoA acetyltransferase), induce plant colonization and Ada (a regulatory protein of adaptive response) and DinF (DNA-damage-inducible protein F) putative stress response proteins.[18]

Efflux pump, regulators, heavy metal resistance and chromosomal β-lactamase proteins in *K. variicola*, *K.* quasipneumoniae and *K. pneumoniae*

The OqxABR efflux pump gene clusters are present in all *K. variicola* and *K. quasipneumoniae* genomes examined. The AcrABR protein efflux was identified in 58.0% (18/31) and 100% (8/8), of *K. variicola* and *K. quasipneumoniae*. All of the protein regulators were identified (MarAR, SoxSR, RamAR, Rob, SdiA, Fis, EnvR, and RarA) in >98% and 100%, of the *K. variicola* and *K. quasipneumoniae* isolates. All efflux pumps and efflux pump regulator proteins were identified in the *K. pneumoniae* NTUH-K2044 and MGH78578 genomes.

The copper (PcoABCDERS), silver (SilCERS) and tellurium (TerABCDEWXYZ) protein clusters for heavy metal resistance were analyzed. The PcoABCDERS and SilCERS protein clusters were identified together in 22.5% (7/31) and 37.5% (3/8) of the *K. variicola* and

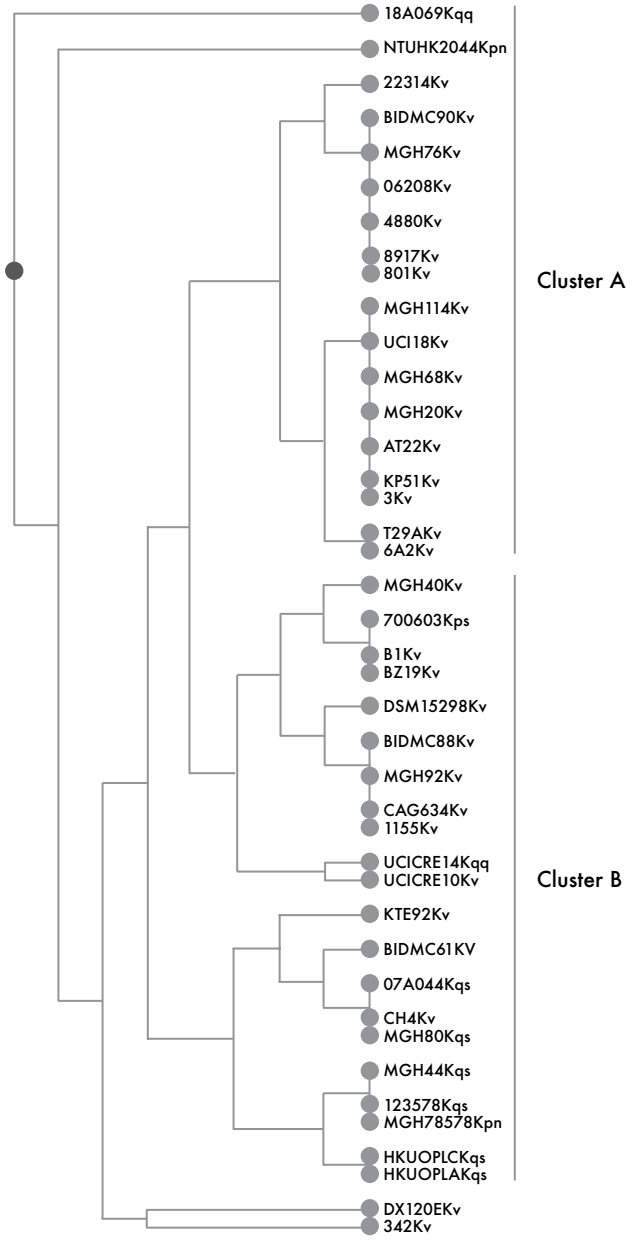

FIGURE 2. JACQUARD INDEX OF THE PLANT-DETERMINANT PROTEINS OF THE *K. VARIICOLA* (KV), *K. QUASIPNEUMONIAE* (KQ) AND *K. PNEUMONIAE* (KPN) GENOMES

K. quasipneumoniae genomes, respectively, while the *K. variicola* Bz19 genome contained only the PcoABCDERS protein cluster.

Tellurite resistance is conferred by the TerABCDEWXYZ protein cluster and seemingly is strongly associated with the hypervirulent clonal groups.[30] Tellurite resistance is needed to colonize macrophages.[31] The TerABCDEWXYZ proteins cluster was identified in 12.9% (4/31) and 12.5% (1/8), of K. variicola and K. quasipneumoniae, all of human origin. K. variicola 8917 isolate was described as hypermucoviscous.[15]

In the analysis of chromosomal β-lactamase, we found that all the *K. pneumoniae* genomes that were identified as *K. variicola* contained the LEN-type β-lactamases

allele. The *K. pneumoniae* genomes that were correctly identified as *K. quasipneumoniae* contained the OKP-type alleles, while the *K. pneumoniae* NTUH-K2044 and MGH78578 genomes contained the SHV-type alleles. The chromosome-encoded β-lactamases corresponded to constitutive genes of these bacterial species (figure 3A). The OKP-A1 and OKP-B1 β-lactamases is characteristic of *K. quasipneumoniae* subsp. *quasipneumoniae* (KpII-A) and *K. quasipneumoniae* subsp. *similipneumoniae* (KpII-B), respectively.

β-lactamase genes were present in most of the *K. variicola* and *K. quasipneumoniae* genomes and were previously annotated as "Beta-lactamase or class A beta-lactamase". In *K. pneumoniae* B1 and *K. variicola,* BZ19 the corresponding sequence was annotated as "Beta-lactamase TEM" and had an amino acid identity of 99 and 100%, respectively, with LEN-2 β-lactamase. In the *Klebsiella* sp. 1.1.55 genome, β-lactamase was annotated as "LEN family Class A β-lactamase" and had a 99% of amino acid identity with LEN-2. Figure 3B shows the phylogenetic analysis of *K. quasipneumoniae* chromosomal β-lactamase genes. The β-lactamases were annotated as class A beta-lactamase, or SHV-1 or TEM proteins. However, all these proteins corresponded to OKP-A1 and OKP-B1 β-lactamases, respectively, in subspecies of *K. quasipneumoniae* (KpII-A, KpII-B)

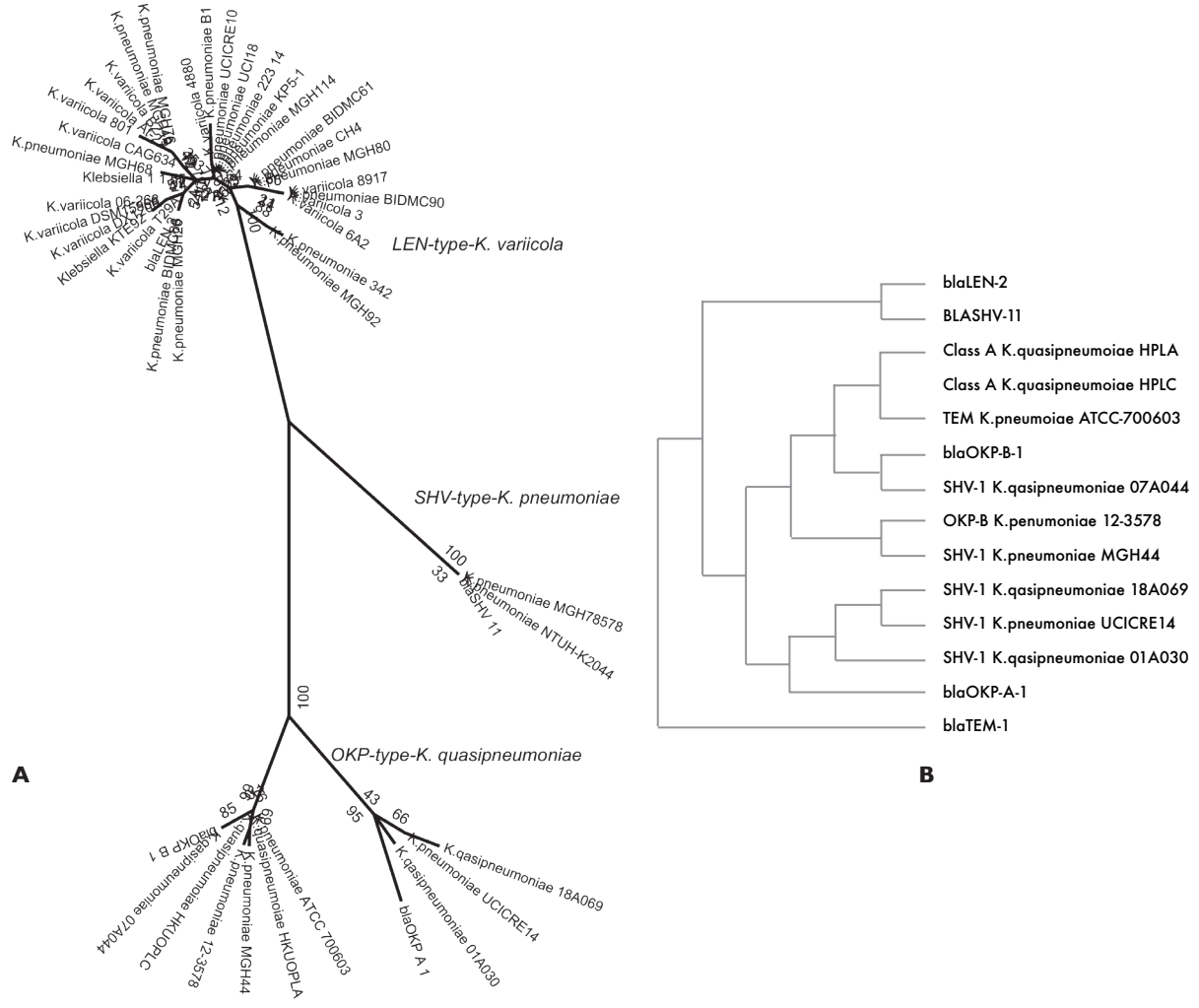

Figure 3. Phylogenetic analysis of chromosomal β-lactamases. A) LEN-, SHV- and OKP-type β-lactamases that were identified in the *K. variicola*, *K. quasipneumoniae* and *K. pneumoniae* genomes. B) Chromosomal β-lactamase protein that was annotated in *K. quasipneumoniae* genomes and that corresponded to the OKP family. HPLA and HPLC, respectively, corresponds to *K. quasipneumoniae* subsp. *similipneumoniae* HKUOPLA and HKUOPLC

(figure 3B). Chromosomal LEN- and OKP-β-lactamase proteins were not properly annotated and could have contributed to the bacterial species misclassification, if these chromosomal β-lactamase proteins are considered for *Klebsiella* classification.

In vitro assays of plant growth-promoting mechanisms in *K. variicola* and *K. pneumonia*

Indole acetic acid (an auxin) has the ability to regulate root growth, differentiation of vascular tissues, elongation, apical dominance, initiation of lateral roots and maturation.[32] This molecule also functions as an important signaling molecule in the regulation of plant growth, expansion, cell differentiation and cell division regulation.[33] *K. variicola* and *K. pneumoniae* isolates were capable of producing indole acetic acid (table II).

Gibberellins (GA) are involved in seed germination, seedling emergence, stem and leaf growth, flower induction, vegetative regulation of reproductive dormancy buds and fruit growth.[34,35] Almost all of the *K. variicola* and *K. pneumoniae* isolates (except for two *K. pneumoniae* isolates) were capable of producing gibberellins.

In addition to nitrogen, phosphorus is a nutrient that is required by plants.[36] *K. variicola* and *K. pneumoniae* isolates were not able to solubilize phosphate (table II), which is a common characteristic among plant associated bacteria. *Klebsiella* genus, unlike other plant growth promoting bacteria, did not produce chitinases or proteases that may help fighting pathogenic fungi of plants (table II). Siderophores are synthesized mainly by Gram-negative bacteria, fungi, yeast and some plants (phytosiderophores) and act as specific chelating agents. They are soluble in aqueous solutions at neutral pH and are considered as secondary metabolites.[37] Siderophores may be important in the process of colonizing plant and human tissues.[38] Different siderophores were identified *in silico* in the genome of *K. variicola*, *K. quasipneumoniae* and *K. pneumoniae*. In agreement with these findings, the *in vitro* siderophore assay was positive for all of the *K. variicola* and *K. pneumoniae* isolates that were tested.

The name *K. variicola* means from various places and novel reports confirmed the adequacy of its name. This species has now been identified in diverse plants, such as bitter gourd and chili, as well as in insects such as *Nezara viridula*. It is also significant that isolates have been obtained from different human samples, such as urine, pus and bile.[17]

K. variicola isolates were additionally obtained from the fecal microbiota of two cohorts of Malawian infants/children, but their genomes could not be included in this work because they are fragmented into a high number of contigs. Of note, the *K. variicola* strain has been identified in giant panda feces,[28] and corresponds to other example of misclassification and corresponds to *K. quasipneumoniae*.[17] The first OXA-181 (carbapenemase)-producing *K. variicola* isolate was identified (by *rpoB* gene analysis) in fresh vegetables that were commercialized between different continents (Asia-Europe).[39] This work showed the fist carbapenem-resistance *K. variicola* strain and may suggest possible emergence paths of resistance in *Enterobacteriaceae*.

Gene flow may occur among different related *Klebsiella* species, mediated by plasmid transfer. The plasmid of *K. variicola* pBz19 showed high identity to pl9 (both associated with IncN incompatibility) from *K. pneumoniae* isolate.[14] *K. variicola* DX120E contains pKV1 and pKV2 plasmids. The pKV1 plasmid is very similar to the pKp5-1 plasmid that was identified in *K. pneumoniae* KP5-1.[19] Plasmid pKV2 is most similar to plasmid pKOXM1C from *K. oxytoca* strain M1, suggesting plasmid exchange between these bacterial species. Additional analyses of plasmids are required specially in *K. variicola* from different niches to further understand their role in host interaction.

Conclusions

This work provides a new molecular framework for distinguishing different organisms of the *Klebsiella* genus. The *in silico* analyses showed that the genomes of *K. variicola* and *K. quasipneumoniae* shared a set of virulence-associated genes encoding mainly siderophores (aerobactin and enterobactin KfuABC cluster), urea metabolism, lipopolysaccharide biosynthesis enzymes (WabG and Uge), and fimbriae genes encoding the *MrkABCDFHIJ* gene cluster. The *in vitro* analysis indicated the presence of different siderophores in distinct isolates and species. Siderophores may have redundant functions given that they all are involved in bacterial iron acquisition. The number of siderophores in a single cell and their type however, may be important to regulate the flow of iron incorporation into the bacterial cell, a process that is linked to invasion of host tissues and thus virulence.

FimV, on the other hand, could be considered a common character in *K. variicola* strains distinguishing them from those of *K. pneumoniae*; however, not all *K. variicola* genomes contained this gene, as shown by both *in silico* and PCR assays. To explain the mosaic distribution of putative host-interaction genes in *Klebsiella* genomes, we envisage the possibility that *Klebsiella* cells may acquire virulence genes by horizontal transfer from closely related bacteria. However, this possibility is tempered by the alternative process of gene loss in bacterial evolution, which could be rather significant in the conformation of

Table II
***In vitro* assays of plants growth-promoting mechanisms in *K. variicola* and *K. pneumoniae* isolates**

Isolates	Direct mechanism			Indirect mechanism	
	Indole acetic acid (mg/ml)*	Solubilization of phosphorus index (mm)	Gibberellins	Siderophores	Lytic enzymes
K. variicola					
F2R9^T	17.19 ± 0.740	1.83	+	+	-
CFNE 2600	20.63 ± 0.191	1.63	+	+	-
T29A	19.96 ± 2.014	1.99	+	+	-
VI	21.81 ± 1.155	2.22	+	+	-
3	19.47 ± 2.532	2.1	+	+	-
6A2	20.03 ± 0.172	1.3	+	+	-
1171	18.09 ± 0.871	1.46	+	+	-
1258	20.01 ± 0.827	2.04	+	+	-
4880	18.58 ± 1.022	2.18	+	+	-
8917	19.63 ± 0.894	1.88	+	+	-
9326	18.28 ± 0.806	1.68	+	+	-
9351	19.95 ± 1.375	2.12	+	+	-
9352	18.65 ± 0.841	2.4	+	+	-
9353	19.17 ± 0.477	1.81	+	+	-
9387	17.27 ± 0.376	1.59	+	+	-
9388	20.22 ± 0.543	2.63	+	+	-
9635	18.00 ± 0.455	1.48	+	+	-
9925	19.63 ± 0.463	0	+	+	-
K. pneumoniae					
6419	17.63 ± 2.13	1.6	-	+	-
6421	17.09 ± 2.57	1.2	+	+	-
9018	16.71 ± 1.56	1.9	-	+	-
1610	16.32 ± 0.73	2.2	+	+	-
06-208	16.88 ± 2.53	2.1	+	+	-
01-250	15.90 ± 0.97	1.5	+	+	-
2989	18.12 ± 1.54	2.6	+	+	-
3407	15.80 ± 1.19	1.4	+	+	-
Control strains					
Azotobacter vinellandii	5.65 ± 0.342	5.8	+	+	NA
Salmonella enteritidis	0.002 ± 0.00	NA	-	NA	NA
Escherichia coli DH5-α	NA	0.0	NA	NA	-
Staphylococcus aureus	NA	NA	NA	-	NA
Pseudomonas fluorecens	NA	NA	NA	+	NA
Bacillus subtilis	NA	NA	NA	NA	+

* The *in vitro* assay was carried by triplicated and the standard errors are shown

NA; not applied, +; positive, -; negative

Klebsiella genomes as has been shown in other bacteria.[40] *K. variicola* and *K. quasipneumoniae* strains with different genetic repertoires have the capacity to adapt to plant and clinical settings. Previously, we suggested that a different epidemiological dynamics occurred with *K. variicola* in comparison to *K. pneumoniae* with *K. variicola* coming from the environment,[41] we further refine here that the environment refers most probably to the plant environment. Plants can be a reservoir for *K. variicola* isolates that may opportunistically infect humans or

animals. *K. variicola* infection in humans thus seems to be a case of "phytonosis", a term we suggest for symbiotic bacteria plant-borne, parallel to the term zoonosis bacteria pathogen animal-borne.

Acknowledgments

This work was funded by the *Consejo Nacional de Ciencia y Tecnología* (Conacyt), SEP-Conacyt – grant number 256988. We thank Angelina Quiterio for her excellent laboratory assistance and Michael Dunn and J. Martínez-Romero for their help in reviewing the manuscript.

Declaration of conflict of interests. The authors declare that they have no conflict of interests.

References

1. Brisse S, Passet V, Grimont PA. Description of *Klebsiella quasipneumoniae* sp. nov., isolated from human infections, with two subspecies, *Klebsiella quasipneumoniae* subsp. *quasipneumoniae* subsp. nov. and *Klebsiella quasipneumoniae* subsp. *similipneumoniae* subsp. nov., and demonstration that *Klebsiella singaporensis* is a junior heterotypic synonym of *Klebsiella variicola.* Int J Syst Evol Microbiol 2014;64(9):3146-3152. https://doi.org/10.1099/ijs.0.062737-0

2. Martinez-Romero E, Silva-Sanchez J, Barrios H, Rodriguez-Medina N, Martinez-Barnetche J, Tellez-Sosa J, *et al*. Draft genome sequences of *Klebsiella variicola* plant isolates. Genome Announc 2015;3(5):e01015-15. https://doi.org/10.1128/genomeA.01015-15

3. Rosenblueth M, Martinez L, Silva J, Martinez-Romero E. *Klebsiella variicola*, a novel species with clinical and plant-associated isolates. System Appl Microbiol 2004;27(1):27-35. https://doi.org/10.1078/0723-2020-00261

4. Pinto-Tomas AA, Anderson MA, Suen G, Stevenson DM, Chu FS, Cleland WW, *et al.* Symbiotic nitrogen fixation in the fungus gardens of leaf-cutter ants. Science 2009;326(5956):1120-1123.

5. Medrano EG, Forray MM, Bell AA. Complete genome sequence of a *Klebsiella pneumoniae* strain isolated from a known cotton insect boll vector. Genome Announc 2014;2(4):e00850-14. https://doi.org/10.1128/genomeA.00850-14

6. Garza-Ramos U, Martinez-Romero E, Silva-Sanchez J. SHV-type extended-spectrum beta-lactamase (ESBL) are encoded in related plasmids from enterobacteria clinical isolates from Mexico. Salud Publica Mex 2007;49(6):415-421. https://doi.org/10.1590/S0036-36342007000600008

7. Garza-Ramos U, Moreno-Dominguez S, Hernandez-Castro R, Silva-Sanchez J, Barrios H, Reyna-Flores F, *et al*. Identification and characterization of imipenem-resistant *Klebsiella pneumoniae* and susceptible *Klebsiella variicola* isolates obtained from the same patient. Microb Drug Resist 2016;22(3):179-184. https://doi.org/10.1089/mdr.2015.0181

8. Garza-Ramos U, Silva-Sanchez J, Martinez-Romero E, Tinoco P, Pina-Gonzales M, Barrios H, *et al*. Development of a multiplex-PCR probe system for the proper identification of *Klebsiella variicola*. BMC Microbiol 2015;15:64. https://doi.org/10.1186/s12866-015-0396-6

9. Maatallah M, Vading M, Kabir MH, Bakhrouf A, Kalin M, Naucler P, *et al*. *Klebsiella variicola* is a frequent cause of bloodstream infection in the stockholm area, and associated with higher mortality compared to *K. pneumoniae*. PloS One 2014;9(11):e113539. https://doi.org/10.1371/journal.pone.0113539

10. Seki M, Gotoh K, Nakamura S, Akeda Y, Yoshii T, Miyaguchi S, *et al*. Fatal sepsis caused by an unusual *Klebsiella* species that was misidentified by an automated identification system. J Med Microbiol 2013;62(5):801-803. https://doi.org/10.1099/jmm.0.051334-0

11. Brisse S, Duijkeren E. Identification and antimicrobial susceptibility of 100 *Klebsiella* animal clinical isolates. Vet Microbiol 2005;25:105(3-4):307-312.

12. Podder MP, Rogers L, Daley PK, Keefe GP, Whitney HG, Tahlan K. *Klebsiella* species associated with bovine mastitis in Newfoundland. PloS One 2014;9(9):e106518. https://doi.org/10.1371/journal.pone.0106518

13. Holt KE, Wertheim H, Zadoks RN, Baker S, Whitehouse CA, Dance D, *et al*. Genomic analysis of diversity, population structure, virulence, and antimicrobial resistance in *Klebsiella pneumoniae*, an urgent threat to public health. Proc Natl Acad Sci USA 2015;112(27):E3574-E3581. https://doi.org/10.1073/pnas.1501049112

14. Andrade BG, de Veiga Ramos N, Marin MF, Fonseca EL, Vicente AC. The genome of a clinical *Klebsiella variicola* strain reveals virulence-associated traits and a pI9-like plasmid. FEMS Microbiol Lett 2014;360(1):13-16. https://doi.org/10.1111/1574-6968.12583

15. Garza-Ramos U, Silva-Sanchez J, Barrios H, Rodriguez-Medina N, Martinez-Barnetche J, Andrade V. Draft genome sequence of the first hypermucoviscous *Klebsiella variicola* clinical isolate. Genome Announc 2015;3(2):e01352-14. https://doi.org/10.1128/genomeA.01352-14

16. Chen M, Li Y, Li S, Tang L, Zheng J, An Q. Genomic identification of nitrogen-fixing *Klebsiella variicola, K. pneumoniae* and *K. quasipneumoniae*. J Basic Microbiol 2016;56(1):78-84 https://doi.org/10.1002/jobm.201500415

17. Martínez-Romero E, Rodríguez-Medina N, Beltrán-Rojel M, Pérez-Rueda E, Garza-Ramos U. Genome misclassification of Klebsiella variicola and Klebsiella quasipneumoniae isolated from plants, animals and humans. Salud Publica Mex. 2018;60. https://doi.org/10.21149/8149

18. Fouts DE, Tyler HL, DeBoy RT, Daugherty S, Ren Q, Badger JH, *et al*. Complete genome sequence of the N2-fixing broad host range endophyte *Klebsiella pneumoniae* 342 and virulence predictions verified in mice. PLoS Genetics 2008;4(7):e1000141. https://doi.org/10.1371/journal.pgen.1000141

19. Lin L, Wei C, Chen M, Wang H, Li Y, Li Y, *et al*. Complete genome sequence of endophytic nitrogen-fixing *Klebsiella variicola* strain DX120E. Stand Genomic Sci 2015;10:22. https://doi.org/10.1186/s40793-015-0004-2

20. Garcia-Vallve S, Palau J, Romeu A. Horizontal gene transfer in glycosyl hydrolases inferred from codon usage in *Escherichia coli* and *Bacillus subtilis*. Mol Biol Evol 1999;16(9):1125-1134. https://doi.org/10.1093/oxfordjournals.molbev.a026203

21. Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol Bol Evol 2011;28(10):2731-2739. https://doi.org/10.1093/molbev/msr121

22. Khalid A, Tahir S, Arshad M, Ahmad Z. Relative efficiency of rhizobacteria for auxin biosynthesis in rhizosphere and non-rhizosphere soils. Aust J Soil Res 2004;42(8):921-926. https://doi.org/10.1071/SR04019

23. Mehta S, Nautiyal CS. An efficient method for qualitative screening of phosphate-solubilizing bacteria. Curr Microbiol 2001;43(1):51-56. https://doi.org/10.1007/s002840010259

24. Schwyn B, Neilands JB. Universal chemical assay for the detection and determination of siderophores. Anal Biochem 1987;160(1):47-56. https://doi.org/10.1016/0003-2697(87)90612-9

25. Malleswari D, Bagyanarayana G. Plant growth-promoting activities and molecular characterization of rhizobacterial strains isolated from medicinal and aromatic plants. IOSR-JPBS 2013;6(6):30-37. https://doi.org/10.9790/3008-0663037

26. Cusa E, Obradors N, Baldoma L, Badia J, Aguilar J. Genetic analysis of a chromosomal region containing genes required for assimilation of allantoin nitrogen and linked glyoxylate metabolism in *Escherichia coli.* J Bacteriol 1999;181(24):7479-7484.

27. Hazen TH, Zhao L, Sahl JW, Robinson G, Harris AD, Rasko DA, *et al.* Characterization of *Klebsiella sp.* strain 10982, a colonizer of humans that contains novel antibiotic resistance alleles and exhibits genetic similarities to plant and clinical *Klebsiella isolates*. Antimicrob Agents Chemother 2014;58(4):1879-1888. https://doi.org/10.1128/AAC.01605-13
28. Lu MG, Jiang J, Liu L, Ma AP, Leung FC. Complete genome sequence of *Klebsiella variicola* Strain HKUOPLA, a cellulose-degrading bacterium isolated from giant panda feces. Genome Announc 2015;3(5):e01200-15. https://doi.org/10.1128/genomeA.01200-15
29. Lu MG, Jiang J, Liu L, Ma AP, Leung FC. Complete genome sequence of *Klebsiella pneumoniae* strain HKUOPLC, a cellulose-degrading bacterium isolated from giant panda feces. Genome Announc 2015;12;3(6):e01318-15. https://doi.org/10.1128/genomeA.01318-15
30. Passet V, Brisse S. Association of tellurite resistance with hypervirulent clonal groups of *Klebsiella pneumoniae*. J Clin Microbiol 2015;53(4):1380-1382: 1380-1382. https://doi.org/10.1128/JCM.03053-14
31. Ponnusamy D, Clinkenbeard KD. Role of tellurite resistance operon in filamentous growth of *Yersinia pestis* in macrophages. PloS One 2015;10(11):e0141984. https://doi.org/10.1371/journal.pone.0141984
32. Woodward AW, Bartel B. Auxin: regulation, action, and interaction. Ann Bot 2005;95(5):707-735. https://doi.org/10.1093/aob/mci083
33. Ryu RJ, Patten CL. Aromatic amino acid-dependent expression of indole-3-pyruvate decarboxylase is regulated by TyrR in *Enterobacter cloacae* UW5. J Bacteriol 2008;190(21):7200-7208. https://doi.org/10.1128/JB.00804-08
34. Cassan F, Bottini R, Schneider G, Piccoli P. *Azospirillum brasilense* and *Azospirillum lipoferum* hydrolyze conjugates of GA20 and metabolize the resultant aglycones to GA1 in seedlings of rice dwarf mutants. Plant Physiol 2001;125(4):2053-2058. https://doi.org/10.1104/pp.125.4.2053
35. Yamaguchi S. Gibberellin metabolism and its regulation. Annu Rev Plant Biol 2008;59:225-251. https://doi.org/10.1146/annurev.arplant.59.032607.092804
36. Naik PR, Raman G, Narayanan KB, Sakthivel N. Assessment of genetic and functional diversity of phosphate solubilizing fluorescent pseudomonads isolated from rhizospheric soil. BMC Microbiol 2008;8:230. https://doi.org/10.1186/1471-2180-8-230
37. Ahmed E, Holmstrom SJ. Siderophores in environmental research: roles and applications. Microbial Biotechnol 2014;7(3):196-208. https://doi.org/10.1111/1751-7915.12117
38. D'Onofrio A, Crawford JM, Stewart EJ, Witt K, Gavrish E, Epstein S, *et al.* Siderophores from neighboring organisms promote the growth of uncultured bacteria. Chemistry & Biology 2010;17(3):254-264. https://doi.org/10.1016/j.chembiol.2010.02.010
39. Zurfluh K, Poirel L, Nordmann P, Klumpp J, Stephan R. First detection of *Klebsiella variicola* producing OXA-181 carbapenemase in fresh vegetable imported from Asia to Switzerland. Antimicrob Resist Infect Control 2015;4:38. https://doi.org/10.1186/s13756-015-0080-5
40. McNally A, Thomson NR, Reuter S, Wren BW. 'Add, stir and reduce': *Yersinia spp.* as model bacteria for pathogen evolution. Nature Rev MicroBiol 2016;14(3):177-190. https://doi.org/10.1038/nrmicro.2015.29
41. Martinez J, Martinez L, Rosenblueth M, Silva J, Martinez-Romero E. How are gene sequence analyses modifying bacterial taxonomy? The case of *Klebsiella*. Int Microbiol 2004;7(4):261-268.