# Enhanced *Klebsiella pneumoniae* carbapenemase (KPC) expression from a novel Tn*4401* deletion

Anita Cheruvanky[1], Nicole Stoesser[2], Anna E. Sheppard[2], Derrick W.Crook[2], Paul S. Hoffman[1], Erin Weddle [1,3], Joanne Carroll[4], Costi D. Sifri[1], Weidong Chai[1], Katie Barry[1], Girija Ramakrishnan[1], Amy J. Mathers[1,4#]

[1]Division of Infectious Diseases and International Health, Department of Medicine, University of Virginia Health System, Charlottesville, Virginia

[2]Modernizing Medical Microbiology Consortium, Nuffield Department of Clinical Medicine, John Radcliffe Hospital, Oxford University, Oxford, United Kingdom

[3]Department of Biologic Sciences, Shenandoah University, Winchester, Virginia

[4]Clinical Microbiology, Department of Pathology, University of Virginia Health System, Charlottesville, Virginia

Abstract (≤250 words), running title (≤54 characters and spaces), Figures 1-4, Tables 1

Running Head: Enhanced KPC expression from novel Tn*4401* deletion

**#Corresponding Author:** Amy Mathers, M.D. Associate Professor of Medicine, Division of Infectious Diseases and International Health, P.O. Box 801361, Charlottesville, VA 22908-1361, Phone: (434) 982-4814, Fax: (434)-982-4047, Email: ajm5b@virginia.edu

**Abstract**

The *Klebsiella pneumoniae* carbapenemase gene ($bla_{KPC}$) is typically located within the mobile transposon Tn*4401*. Enhanced KPC expression has been associated with deletions in the putative promoter region upstream of $bla_{KPC}$. Illumina sequences from $bla_{KPC}$-positive clinical isolates from a single institution were mapped to a Tn*4401*b reference sequence, which carries no deletions. The novel isoform Tn*4401*h (188bp deletion [between *istB* and $bla_{KPC}$]) was present in 14% (39/281) clinical isolates. MICs for *Escherichia coli* strains containing plasmids with Tn*4401*a and Tn*4401*h were more resistant to meropenem (≥16, ≥16), ertapenem (≥8, 4) and cefepime (≥64, 4) than *E. coli* strains with Tn*4401*b (0.5, ≤0.5, ≤1). Quantitative RT-PCR demonstrated that Tn*4401*a had a 16-fold and Tn*4401*h a 4-fold increase in $bla_{KPC}$ mRNA levels compared to the reference Tn*4401*b. A *lacZ* reporter plasmid was used to test the activity of the promoter regions from the different variants and showed that the Tn*4401*a and Tn*4401*h promoter sequences generated higher β-galactosidase activity than the corresponding Tn*4401*b sequence. Further dissection of the promoter region demonstrated that putative promoter P1 was not functional. Activity of the isolated promoter P2 was greatly enhanced by inclusion of the P1-P2 intervening sequence. These studies indicate that gene expression could be an important consideration in understanding resistance phenotypes predicted by genetic signatures in the context of sequencing-based rapid diagnostics.

41 **INTRODUCTION**

42 Carbapenem-resistant Enterobacteriaceae (CRE) are an urgent public health threat because of
43 their increasing incidence and high mortality seen with infection (1). Currently, the most
44 commonly reported mechanism of carbapenem resistance in clinical Enterobacteriaceae is
45 *Klebsiella pneumoniae* carbapenemase (KPC), which has disseminated globally over the last
46 decade, becoming endemic in several geographic locations such as Italy, Greece and the United
47 States (2). KPC is a class A serine beta-lactamase which can hydrolyze penicillins,
48 cephalosporins, aztreonam and carbapenems, limiting treatment options in infected patients. The
49 KPC gene ($bla_{KPC}$) is typically located within a 10kb mobile transposon (Tn*4401*), which is most
50 often situated on conjugative plasmids (3, 4). The association of $bla_{KPC}$ with mobile genetic
51 elements like plasmids and transposons contributes to inter- and intra-species gene transfer and
52 dissemination of carbapenem resistance (2, 5-8).
53
54 Tn*4401* is composed of $bla_{KPC}$, a transposase gene (*tnpA*), a resolvase gene (*tnpR*) and two
55 insertion sequences IS*Kpn6* and IS*Kpn7*, all flanked by two 32bp inverted repeat sequences (3).
56 To date, seven unique Tn*4401* isoforms (a-g) have been identified (9-12), with Tn*4401*a and
57 Tn*4401*b most widespread (13). The original characterization of these isoforms demonstrated
58 deletions in the non-coding region upstream of $bla_{KPC}$ (a, -99 bp (14, 15); b, no deletion (3); c, -
59 215 bp (16); d, -68 bp (9)/-5.3 kb (9, 17, 18); e, -255 bp (17); f, truncated *tnpA* and lacking *tnpR*,
60 *ISKpn7* left, and Tn*4401* IRL-1 (10); g, lacking *tnpA*, *tnpR*, *ISKpn7* left with -215 bp deletion in
61 the non-coding region (Figure 1); most recently a truncated version of Tn*4401*e, lacking *tnpA*
62 and *tnpR* was also described (19).
63

Previously, the non-coding region between IS*Kpn7* and *bla*$_{KPC}$ was proposed to contain three putative promoter regions (P1, P2, P3), of which two were demonstrated to affect expression (P1, P2) while P3 was shown to not be a real promoter (5, 8, 20). Deletions in the non-coding region upstream of *bla*$_{KPC}$ have demonstrated variable expression of KPC *in vitro*, with the highest levels of expression observed with Tn*4401*a (presence of P1 and P2). However, the relative contribution of P1, P2 and regions of associated flanking sequence to KPC expression has not been fully characterized.

Here, we describe a novel Tn*4401* transposon variant (Tn*4401*h), which has a unique 188bp deletion in the non-coding region upstream of *bla*$_{KPC}$, and similar to Tn*4401*a, retains promoters P1 and P2. We explore the effect of this 188bp deletion on the level of KPC expression, the degree of phenotypic resistance compared to the most frequently seen and well-characterized isoforms Tn*4401*a and Tn*4401*b (4) and further assess promoter activity using fusions to a *lacZ* reporter. Understanding the impact of genetic variability in non-coding and/or promoter regions on resistance gene expression and its correlation with minimum inhibitory concentrations (MICs) will be important for the use of sequencing-based approaches in clinical diagnostics and antimicrobial susceptibility prediction (8, 21).

## RESULTS

### The molecular epidemiology of Tn*4401* in the study isolates

The Tn*4401* sequence was identified in 281 isolates across eight genera, mainly represented by *Klebsiella* spp. (n=129), *Enterobacter* spp. (n=101) and *Citrobacter* spp. (n=32) with *bla*$_{KPC-2}$ found in the majority (230/281; 82%) of the isolates as previously characterized (6). The same

work also reported that the most common Tn*4401* variant was Tn*4401*b (n=230); Tn*4401*a was seen in 8 isolates (all *K. pneumoniae*); and a novel variant (Tn*4401*h; all with $bla_{KPC-2}$ [Figure 1]) with a 188bp deletion identified in 39 isolates from 28 patients. Tn*4401*h was first identified in a *K. pneumoniae* strain (Feb 2009); all other Tn*4401*h variants were identified in highly-genetically related *Enterobacter cloacae* isolates from all 28 patients (including the patient with the initial *K. pneumoniae* with Tn*4401*h) over three years (May 2009- Dec 2012), consistent with local inter-species and inter-patient transmission (6).

Structural variation in published isoforms of Tn*4401* and the novel Tn*4401*h isoform, including variation in IS*Kpn6*, IS*Kpn7*, and the putative promoter (P1, P2, P3) regions, is shown in Figure 1(9). Notably isoforms a and h do not contain the putative promoter P3 and have nucleotide variation in the intervening sequence between P2 and P1 (Figure 2).

**Susceptibility testing**

Susceptibility results for selected parent strains and *E. coli* transformants from CAV1016, CAV1746 and CAV1438 containing Tn*4401*b, Tn*4401*a and Tn*4401*h respectively are presented in Table 1. Parent strains CAV1746 (Tn*4401*a) and CAV1438 (Tn*4401*h) were resistant to cefepime and meropenem by broth dilution and VITEK2 (Table 1). CAV1016 (Tn*4401*b) was resistant to meropenem by broth dilution and VITEK2 but was susceptible to cefepime by VITEK2.

Susceptibility testing by VITEK2 and broth dilution of *E. coli* transformants containing Tn*4401*a or Tn*4401*h to cefepime and meropenem demonstrated MICs that were all in the resistant range.

For the Tn*4401*b *E. coli* transformant, cefepime was in the susceptible dose dependent (SDD) range and meropenem was in the susceptible range by broth dilution. By VITEK2 both cefepime and meropenem demonstrated MICs which were in the susceptible range (2015 CLSI breakpoints, Table 1) (22). The plasmid backbones for each transformant were unique but both the Tn*4401*h and Tn*4401*a were large IncR plasmids whereas Tn*4401*b was on a non-typable 43kb plasmid (pKPC_UVA01) (Table 1).

For parent strains CAV1016, CAV1746, CAV1438 and *E. coli* transformants, carbapenemase production was indicated by both the modified Hodge Test and the indirect carbapenemase test (23).

**Expression Assays**

Two-step qRT-PCR was performed on all three *E. coli* transformants using primers as noted in Supplementary Table 1. Fold difference in $bla_{KPC}$ expression (normalized to *rpoB*) was quantified relative to Tn*4401*b. Tn*4401*a had 23-fold greater $bla_{KPC}$ expression compared to Tn*4401*b (student t-test: p=0.011) (Figure 3). Tn*4401*h had 4-fold greater $bla_{KPC}$ expression compared to Tn*4401*b (student t-test: p=0.03) (Figure 3).

Transcription fusions to the *lacZ* reporter in plasmid pRS551 (24) of the putative promoter sequences from plasmids carrying Tn*4401*a, Tn*4401*b and Tn*4401*h (Supplementary Table 1, Figure 4) were used to quantify β-galactosidase activity associated with each Tn*4401* variant. Promoter sequence Tn*4401*a had significantly higher β-galactosidase activity compared with Tn*4401*b (student t-test: p=0.0001) and Tn*4401*h (student t-test: p=0.0002), and the Tn*4401*h

putative promoter sequence had significantly higher β-galactosidase activity compared with Tn*4401*b (student t-test: $p<0.0001$) (Figure 4). These results were consistent with the effect each deletion had on the level of $bla_{KPC}$ mRNA in the transformants (Figure 3).

Tn*4401*h is missing an intervening sequence of 188 nucleotides (IVS) between P1 and P2, whereas Tn*4401*a carries a smaller deletion in this region (Fig. 2 and Fig. 4A). To better discern the contribution of the different sequence domains to gene expression, *lacZ* fusions were constructed of the isolated promoters P1 and P2, as well as of these promoter regions carrying the IVS. Reporter constructs containing P1 or P1+IVS showed no β-galactosidase activity. Reporter constructs containing P2 showed significantly higher β-galactosidase activity compared with vector (student t-test: $p=0.00025$) or P1 (student t-test: $p=0.0034$) and P1+IVS (student-t test: $p=0.0020$). Reporter constructs containing P2+IVS had significantly higher β-galactosidase activity relative to P2 alone (student t-test: $p=0.015$), P1 (student t-test: $p<0.001$) and P1+IVS (student t-test: $p<0.001$) (Fig4). Additionally promoter sequence Tn*4401*a had significantly higher β-galactosidase activity compared with reporter constructs containing P2+IVS (student t-test; $p=0.006$).

## DISCUSSION

As KPC-producing Enterobacteriaceae are seen in increasing numbers around the globe, understanding expression of $bla_{KPC}$ and its effect on the degree of associated carbapenem resistance could have significant clinical consequences. The Tn*4401* transposon has been the most consistently observed unit of $bla_{KPC}$ transfer, and the non-coding region upstream of $bla_{KPC}$ has been described as a regulatory region (3, 25, 26). We describe a novel isoform of Tn*4401*,

Tn*4401*h, which was first identified in a *K. pneumoniae* strain in a patient who also had an *E. cloacae* strain that was subsequently transmitted between several patients over a period of three years in our institution.

The 188bp deletion of Tn*4401*h, encompassing the region between P1 and P2, demonstrated increased $bla_{KPC}$ expression when compared with Tn*4401*b (no deletions), although to a lesser degree than Tn*4401*a (99bp deletion; also P1+, P2+ with smaller deletion between P1 and P2). To exclude the impact of variability of gene copy number and plasmid background on these results, we directly investigated the impact of promoter regions and intervening sequences on gene expression in a model system, by inserting these promoter regions and promoter/intervening sequence combinations into a *lacZ* reporter.

The Tn*4401*h promoter region showed significantly higher β-galactosidase activity compared with Tn*4401*b, again to a lesser degree than the Tn*4401*a promoter region, and consistent with the $bla_{KPC}$ expression results (Tn*4401*a>Tn*4401*h>Tn*4401*b). Given that Tn*4401*a and Tn*4401*h both have deletions in the intervening sequence between P1 and P2, with P1 and P2 intact, the differential levels of gene expression and promoter activity remain incompletely explained. One possible explanation for this variability is that the differing length of intervening sequence between P2 and P1 for Tn*4401*a (99bp deletion) and Tn*4401*h (188bp deletion) may potentially result in a more stable RNA structure for Tn*4401*a.

Previously, Naas *et al*. reported only P1 and P2 as true promoters involved in $bla_{KPC}$ expression. A review of the -10/-35 regions of the promoter P1 and P2 sequences shows the P2 promoter

(initially identified by Roth (27)) has a strong -35 consensus sequence (TTGACA) and only two minor differences from the canonical -10 sequence (TATctT), while the P1 promoter has a poor consensus sequence, namely 4 differences from the canonical -35 sequence (TaatCc) and 3 differences from the canonical -10 sequence (TtacAT). However, the relative contribution of putative promoter P1 by itself and in relation to the IVS remained previously unknown. Hence, promoter sequences for P1 alone and in combination with IVS were cloned into a *lacZ* reporter plasmid. Surprisingly, the reporter constructs containing P1 and P1+IVS showed no β-galactosidase activity, indicating that this promoter was not active under our assay conditions. We additionally attempted to characterize the relative contribution of promoter P2 alone and in combination with IVS. Higher β-galactosidase activity was observed with the promoter P2 *lacZ* construct compared with P1, and this was further enhanced when IVS was included. These results are consistent with findings of Naas *et al.* (20). The exact role of IVS in the enhanced activity of P2+IVS compared to P2 will require future evaluation and would be ideally varied in IVS length.

*E. coli* transformants containing Tn*4401*a, and h variants on native parent plasmid backgrounds showed resistant meropenem and cefepime MICs compared to *E. coli* transformants containing Tn*4401*b, consistent with the variations in $bla_{KPC}$ expression (i.e. Tn*4401*b<Tn*4401*h< Tn*4401*a). Previous studies have suggested that additional mechanisms aside from gene expression and copy number, namely production of porin channel defects, are primary contributors to the variability in susceptibility when testing KPC-producing isolates (3, 11, 20, 28-30). Here we excluded the impact of porin channel variation on susceptibility by evaluating plasmids bearing the Tn*4401*-$bla_{KPC}$ units in a fixed *E. coli* background. Although we did not exclude gene copy

number as a contributing factor to expression as the different Tn*4401* variants were located on unique plasmid backbones, the difference in the promoter activity seen in the β-galactosidase assay aligns with the differences seen in the *E. coli* transformants. In addition, Hanson *et al.* previously showed that in $bla_{KPC-2}$ transformants, $bla_{KPC}$ gene copy number did not correlate with expression and increases in β-lactam MIC; similarly Kitchel *et al*. demonstrated that factors other than $bla_{KPC}$ copy number likely contribute to elevated KPC production and high level carbapenem resistance in certain isolates (26, 27). Taken together, these data suggest that the promoter region upstream of $bla_{KPC}$ appears to have the most substantial effect on expression of carbapenem resistance in the transformants. Future studies utilizing methods such as chromatin immunoprecipitation with sequencing (ChIP-sequencing) could be used to examine the interaction of transcription factors and thus offer further insight into regulation of $bla_{KPC}$ expression.

By characterizing the Tn*4401*h isoform we have highlighted that the the promoter region may be more complex than previously described. Correctly ascertaining the role of promoter regions and 5' untranslated regions in resistance gene expression is of particular relevance to clinical diagnostic microbiology given the development of an increasing number of genotype-based approaches (e.g. microarray, multiplex, or whole genome sequencing) to resistance prediction (31-35). There is a trend towards incorporating rapid genotypic methods into routine laboratory diagnostics yet there is increasing recognition of the importance of pharmacokinetics/pharmacodynamics in relation to the MIC. Therefore, understanding the predicted impact of promoter regions on the MIC may be of increasing importance to clinical care.

In summary, we have found a novel isoform, Tn*4401*h, with a 188bp deletion in the upstream non-coding region between IS*Kpn7* and $bla_{KPC}$. We have shown that this deletion has a differential effect on carbapenem MICs compared with other common Tn*4401* variants. This work increases our understanding of the complexity of $bla_{KPC}$ expression and has implications for predicting the degree of resistance based solely on the presence or absence of a particular resistance gene.

## MATERIALS AND METHODS

### Sampling

The sampling of $bla_{KPC}$-positive Enterobacteriaceae clinical and surveillance isolates from the University of Virginia Health System (UVaHS) between August 2007 and December 2012 has been previously published (6). Sheppard *et al.* performed Illumina sequencing on the collected isolates and determined the Tn*4401* isoform of each isolate using BLASTn comparisons between the isolate's *de novo* assembly and the Tn*4401*b reference sequence (EU176013.1) (3, 6). Where available, ertapenem minimum inhibitory concentration (MIC) results from the VITEK2 AST GN-70 (Biomerieux Durham, NC) were obtained retrospectively from laboratory records. Purified plasmid DNA was electroporated into electrocompetent *E. coli* Genehog cells (Invitrogen, Carlsbad, CA) as previously described (25). To examine the plasmid background in the newly generated transformants incompatibility group PCR typing was performed (36). For CAV1746 and CAV1438 parent strains (those with unknown $bla_{KPC}$ plasmids), the Illumina contig containing the corresponding replicon sequence (1,960 bp and 15,241 bp respectively) and the contig containing Tn*4401* (18,813 bp and 10,011 bp respectively) were used as BLASTn

queries to determine similarity with previously described plasmids in GenBank. For CAV1016, $bla_{KPC}$ plasmid background has been previously described (25). For plasmid size estimation *Bam*HI and *Eco*RI (New England Ipswich, MA) digested and undigested plasmid extractions were run on a 0.8% agarose gel over 8 hours at 70V with V517 and Hyperladder I (Bioline Taunton, MA) to estimate plasmid size as previously described (37).

**Quantitative real-time polymerase chain reaction (qRT-PCR) experiments**

Total RNA was extracted from *E. coli* transformants containing parent plasmid Tn*4401*a, Tn*4401*b and Tn*4401*h using the RNeasy Mini Kit (Qiagen, GmBH, D 40724 Hilden, Germany). RNA products underwent column DNAase I digestion (Qiagen) and were quantified by measuring the optical density at 260nm using a Nanodrop 1000 spectrophotometer (Thermo Scientific, Wilmington, DE, USA). cDNA was synthesized from RNA using a commercial method in accordance with the manufacturer's instructions (Quanta Biosciences, qScript cDNA Supermix, Gaithersburg, MD).

$bla_{KPC}$-gene expression was examined in *E. coli* transformants utilizing qRT-PCR and the synthesized cDNA (primers listed in Supplementary Table 1). Twenty-five nanograms of RNA equivalent DNA were used in triplicate. qRT-PCR was performed using SsoFast EvaGreen Supermix (Bio Rad Laboratories, Hercules, CA), template DNA, 500 nM of each probe RT-KPC-F/RT-KPC-R or EcoliRPOB-F/EcoliRPOB-R (Supplementary Table 1). Cycling conditions included a 3 min enzyme activation step at 95ºC followed by 40 cycles of: melting (95ºC for 10 sec) and annealing/extension (56.3ºC for 30 sec/72ºC for 30 sec); followed by a final extension for 72ºC for 1 min. Standard curves were generated for both the target ($bla_{KPC}$)

and the endogenous control (*rpoB*) using 4-fold dilutions of template DNA at known concentrations (from 0.39ng to 100ng) and by plotting the logarithm of the initial quantity of template (along the x axis) versus the respective cycle threshold (Ct) values (along the y axis). Thereafter for each experimental sample the amount of $bla_{KPC}$ and *rpoB* was calculated from the appropriate standard curve. Expression levels of $bla_{KPC}$ were first normalized relative to the *rpoB* gene in *E. coli* and fold difference in $bla_{KPC}$ expression was calculated: (normalized $bla_{KPC}$ expression *E. coli* transformant Tn*4401*a or h) / (normalized $bla_{KPC}$ expression *E. coli* transformant Tn*4401*b).

**PCR and cloning**

Online software ClustalW (http://www.ebi.ac.uk/Tools/msa/clustalw2/) was utilized to align promoter regions of isoforms Tn*4401*a, b and h. Primers (F-ISKpn7, R-prom) containing restriction sites for *Bam*HI and *Eco*RI (New England Biolabs, Ipswich, MA) were used to amplify putative promoter sequences from purified plasmids carrying Tn*4401*a, Tn*4401*b and Tn*4401*h (Supplementary Table 1). Primers were also used to amplify putative promoter P1 (P1Eco-F, R-prom); P1 + intervening sequence- (IVS; intervening sequence between P1 and P2 that is missing in Tn*4401*h) (P1+IVS Eco-F, R-prom); P2 (F-ISKpn7, P2 Bam –R); P2 + IVS which is in (F-ISKpn7, P2+IVS Bam –R), in order to assess the relative activity of putative promoters P1, P2 alone and in addition to the IVS, (Figure 2, 4 Supplementary Table 1). All amplified PCR products were ligated into the *lacZ* reporter plasmid pRS551 (24) after restriction digestion using *Bam*HI and *Eco*RI . The ligated constructs and the *lacZ* reporter plasmid pRS551 were transformed into *E. coli* DH10B calcium-competent cells and then grown on ampicillin (100µg/ml) selective LB plates. Plasmids isolated from transformants were screened by PCR and

all 7 ligated constructs were sequenced using primer (pRS551-R Sequencing; Supplementary Table 1) to confirm proper orientation and nucleotide sequence of the inserted regions.

**β-galactosidase assay**

The protocol for assay of β-galactosidase activity was adapted from Griffith and Wolf (38). Five µl of the overnight culture of the DH10B *E. coli* transformants were added to 3ml LB broth containing ampicillin (50µg/ml) and the cultures were incubated on a platform shaker at 37°C until the culture reached an optical density of 0.4-0.5 at 600nm.

For cell permeabilization to release the enzyme, 1 ml of Z buffer (60 mM Na2HPO4· 7H2O, 40 mM NaH2PO4·z H2O, 10 mM KCl, 1 mM MgSO4 · 7H2O, 50 mM β-mercaptoethanol) was added along with 20µl of freshly prepared 0.1% SDS and 40µl of chloroform to100µl of bacterial culture. Permeabilization was achieved by vortexing test tubes containing the cell-chloroform/SDS mixture for 30 secs (one at a time) followed by a 10 minute settling period.

One hundred µl of supernatant from each tube were aliquoted in quadruplicate into a flat bottom 96-well plate. At zero time, the assay was initiated by using a multichannel pipette to add 20µl of ortho-Nitrophenyl-β-galactoside (ONPG, 4 mg/ml). Once the yellow chromophore was visualized the reaction was terminated by the addition of 50µl of 1 M $Na_2CO_3$ and then absorbance at $A_{420}$ and $A_{550}$ was measured. β-galactosidase activity was calculated in Miller units from the absorbance data (Miller Units = 1000 x $[(A_{420} – (1.75 \times A_{550}))]/[T \times V \times A_{600}]$; $A_{420}$ and $A_{550}$ are read from the reaction mixture, $A_{600}$ reflects cell density in the mid-logarithmic phase, T

= time of the reaction in minutes before adding 1 M $Na_2CO_3$, V = volume of culture used in the assay in ml) (38).

**Statistical analysis**

A paired-samples student t-test was conducted to compare $bla_{KPC}$ expression between *E.coli* transformants containing parent plasmids Tn*4401*a, Tn*4401*b and Tn*4401*h. A paired-samples student t-test was also used to compare β-galactosidase activitiy between lacZ fusion constructs containing: Tn*4401*a, Tn*4401*b, Tn*4401*h, P1, P1+IVS, P2 and P2+IVS.

## FUNDING INFORMATION

This research was conducted with support from the NIH T32 training grant (NIH grant 3185T32AI007 046-37). EW was supported by the Summer Research Internship Program (SRIP) and the University of Virginia School of Medicine. This work was also supported by National Institute for Health Research [NIHR grant G0800778]); NIHR Oxford Biomedical Research Centre (NS, AS and DC).

## ACKNOWLEDGMENTS

We thank the Clinical Microbiology lab at UVaHS for their outstanding contributions to this study.

335 **REFERENCES**

336


337 1. Centers for Disease Control and Prevention (CDC). 2013. *Antibiotic resistance threats in*
338 *the United States*. (CDC), Atlanta, GA USA.

339 2. Mathers AJ, Peirano G, Pitout JD. 2015. The role of epidemic resistance plasmids and
340 international high-risk clones in the spread of multidrug-resistant Enterobacteriaceae.
341 Clin Microbiol Rev 28:565-91.

342 3. Naas T, Cuzon G, Villegas MV, Lartigue MF, Quinn JP, Nordmann P. 2008. Genetic
343 structures at the origin of acquisition of the beta-lactamase bla KPC gene. Antimicrob
344 Agents Chemother 52:1257-63.

345 4. Bowers JR, Kitchel B, Driebe EM, MacCannell DR, Roe C, Lemmer D, de Man T,
346 Rasheed JK, Engelthaler DM, Keim P, Limbago BM. 2015. Genomic Analysis of the
347 Emergence and Rapid Global Dissemination of the Clonal Group 258 Klebsiella
348 pneumoniae Pandemic. PLoS One 10:e0133727.

349 5. Chmelnitsky I, Shklyar M, Leavitt A, Sadovsky E, Navon-Venezia S, Ben Dalak M,
350 Edgar R, Carmeli Y. 2014. Mix and match of KPC-2 encoding plasmids in
351 Enterobacteriaceae-comparative genomics. Diagn Microbiol Infect Dis 79:255-60.

352 6. Sheppard AE, Stoesser N, Wilson DJ, Sebra R, Kasarskis A, Anson LW, Giess A,
353 Pankhurst LJ, Vaughan A, Grim CJ, Cox HL, Yeh AJ, Sifri CD, Walker AS, Peto TE,
354 Crook DW, Mathers AJ, Group MMMMI. 2016. Nested Russian Doll-Like Genetic
355 Mobility Drives Rapid Dissemination of the Carbapenem Resistance Gene blaKPC.
356 Antimicrob Agents Chemother 60:3767-78.

7. Luo Y, Yang J, Ye L, Guo L, Zhao Q, Chen R, Chen Y, Han X, Zhao J, Tian S, Han L. 2014. Characterization of KPC-2-producing Escherichia coli, Citrobacter freundii, Enterobacter cloacae, Enterobacter aerogenes, and Klebsiella oxytoca isolates from a Chinese Hospital. Microb Drug Resist 20:264-9.

8. Cuzon G, Naas T, Nordmann P. 2011. Functional characterization of Tn4401, a Tn3-based transposon involved in blaKPC gene mobilization. Antimicrob Agents Chemother 55:5370-3.

9. Chen L, Chavda KD, Mediavilla JR, Jacobs MR, Levi MH, Bonomo RA, Kreiswirth BN. 2012. Partial excision of blaKPC from Tn4401 in carbapenem-resistant Klebsiella pneumoniae. Antimicrob Agents Chemother 56:1635-8.

10. Bryant KA, Van Schooneveld TC, Thapa I, Bastola D, Williams LO, Safranek TJ, Hinrichs SH, Rupp ME, Fey PD. 2013. KPC-4 Is encoded within a truncated Tn4401 in an IncL/M plasmid, pNE1280, isolated from Enterobacter cloacae and Serratia marcescens. Antimicrob Agents Chemother 57:37-41.

11. Pecora ND, Li N, Allard M, Li C, Albano E, Delaney M, Dubois A, Onderdonk AB, Bry L. 2015. Genomically Informed Surveillance for Carbapenem-Resistant Enterobacteriaceae in a Health Care System. MBio 6:e01030.

12. Baraniak A, Izdebski R, Fiett J, Herda M, Derde LP, Bonten MJ, Adler A, Carmeli Y, Goossens H, Hryniewicz W, Brun-Buisson C, Gniadkowski M, MOSAR WP2 WP, and WP5 Study Groups. 2016. KPC-Like Carbapenemase-Producing Enterobacteriaceae Colonizing Patients in Europe and Israel. Antimicrob Agents Chemother 60:1912-7.

13. Munoz-Price LS, Poirel L, Bonomo RA, Schwaber MJ, Daikos GL, Cormican M, Cornaglia G, Garau J, Gniadkowski M, Hayden MK, Kumarasamy K, Livermore DM,

Maya JJ, Nordmann P, Patel JB, Paterson DL, Pitout J, Villegas MV, Wang H, Woodford N, Quinn JP. 2013. Clinical epidemiology of the global expansion of Klebsiella pneumoniae carbapenemases. Lancet Infect Dis 13:785-96.

14. Curiao T, Morosini MI, Ruiz-Garbajosa P, Robustillo A, Baquero F, Coque TM, Canton R. 2010. Emergence of bla KPC-3-Tn4401a associated with a pKPN3/4-like plasmid within ST384 and ST388 Klebsiella pneumoniae clones in Spain. J Antimicrob Chemother 65:1608-14.

15. Gootz TD, Lescoe MK, Dib-Hajj F, Dougherty BA, He W, Della-Latta P, Huard RC. 2009. Genetic organization of transposase regions surrounding blaKPC carbapenemase genes on plasmids from Klebsiella strains isolated in a New York City hospital. Antimicrob Agents Chemother 53:1998-2004.

16. Cuzon G, Naas T, Truong H, Villegas MV, Wisell KT, Carmeli Y, Gales AC, Venezia SN, Quinn JP, Nordmann P. 2010. Worldwide diversity of Klebsiella pneumoniae that produce beta-lactamase blaKPC-2 gene. Emerg Infect Dis 16:1349-56.

17. Kitchel B, Rasheed JK, Patel JB, Srinivasan A, Navon-Venezia S, Carmeli Y, Brolund A, Giske CG. 2009. Molecular epidemiology of KPC-producing Klebsiella pneumoniae isolates in the United States: clonal expansion of multilocus sequence type 258. Antimicrob Agents Chemother 53:3365-70.

18. Chen L, Chavda KD, Fraimow HS, Mediavilla JR, Melano RG, Jacobs MR, Bonomo RA, Kreiswirth BN. 2013. Complete nucleotide sequences of blaKPC-4- and blaKPC-5-harboring IncN and IncX plasmids from Klebsiella pneumoniae strains isolated in New Jersey. Antimicrob Agents Chemother 57:269-76.

19. Martinez T MI, Vazquez GJ, Aquino EE and Robledo IE. 2016. Genetic environment of the KPC gene in Acinetobacter baumanii ST2 clone from Puerto Rico and genomic insights into its drug resistance. Journal of Medical Microbiology 65:784-792.

20. Naas T, Cuzon G, Truong HV, Nordmann P. 2012. Role of ISKpn7 and deletions in blaKPC gene expression. Antimicrob Agents Chemother 56:4753-9.

21. Kos VN, Déraspe M, McLaughlin RE, Whiteaker JD, Roy PH, Alm RA, Corbeil J, Gardner H. 2015. The resistome of Pseudomonas aeruginosa in relationship to phenotypic susceptibility. Antimicrob Agents Chemother 59:427-36.

22. CLSI. 2015. Performance standards for antimicrobial susceptibility testing. M100-S19. Clinical and Laboratory Standards Institute, Wayne, PA.

23. Mathers AJ, Carroll J, Sifri CD, Hazen KC. 2013. Modified Hodge test versus indirect carbapenemase test: prospective evaluation of a phenotypic assay for detection of Klebsiella pneumoniae carbapenemase (KPC) in Enterobacteriaceae. J Clin Microbiol 51:1291-3.

24. Simons RW, Houman F, Kleckner N. 1987. Improved single and multicopy lac-based cloning vectors for protein and operon fusions. Gene 53:85-96.

25. Mathers AJ, Stoesser N, Sheppard AE, Pankhurst L, Giess A, Yeh AJ, Didelot X, Turner SD, Sebra R, Kasarskis A, Peto T, Crook D, Sifri CD. 2015. Klebsiella pneumoniae carbapenemase (KPC)-producing K. pneumoniae at a single institution: insights into endemicity from whole-genome sequencing. Antimicrob Agents Chemother 59:1656-63.

26. Kitchel B, Rasheed JK, Endimiani A, Hujer AM, Anderson KF, Bonomo RA, Patel JB. 2010. Genetic factors associated with elevated carbapenem resistance in KPC-producing Klebsiella pneumoniae. Antimicrob Agents Chemother 54:4201-7.

425 27. Roth AL, Kurpiel PM, Lister PD, Hanson ND. 2011. bla(KPC) RNA expression
426 correlates with two transcriptional start sites but not always with gene copy number in
427 four genera of Gram-negative pathogens. Antimicrob Agents Chemother 55:3936-8.
428 28. Tavares CP, Pereira PS, Marques Ede A, Faria C, Jr., de Souza Mda P, de Almeida R,
429 Alves Cde F, Asensi MD, Carvalho-Assef AP. 2015. Molecular epidemiology of KPC-2-
430 producing Enterobacteriaceae (non-Klebsiella pneumoniae) isolated from Brazil. Diagn
431 Microbiol Infect Dis 82:326-30.
432 29. Shen P, Wei Z, Jiang Y, Du X, Ji S, Yu Y, Li L. 2009. Novel genetic environment of the
433 carbapenem-hydrolyzing beta-lactamase KPC-2 among Enterobacteriaceae in China.
434 Antimicrob Agents Chemother 53:4333-8.
435 30. Queenan AM, Bush K. 2007. Carbapenemases: the versatile beta-lactamases. Clin
436 Microbiol Rev 20:440-58, table of contents.
437 31. Ellington MJ, Findlay J, Hopkins KL, Meunier D, Alvarez-Buylla A, Horner C, McEwan
438 A, Guiver M, McCrae LX, Woodford N, Hawkey P. 2016. Multicentre evaluation of a
439 real-time PCR assay to detect genes encoding clinically relevant carbapenemases in
440 cultured bacteria. Int J Antimicrob Agents 47:151-4.
441 32. Chen L, Mediavilla JR, Endimiani A, Rosenthal ME, Zhao Y, Bonomo RA, Kreiswirth
442 BN. 2011. Multiplex real-time PCR assay for detection and classification of Klebsiella
443 pneumoniae carbapenemase gene (bla KPC) variants. J Clin Microbiol 49:579-85.
444 33. Poirel L, Walsh TR, Cuvillier V, Nordmann P. 2011. Multiplex PCR for detection of
445 acquired carbapenemase genes. Diagn Microbiol Infect Dis 70:119-23.
446 34. Adler A, Khabra E, Chmelnitsky I, Giakkoupi P, Vatopoulos A, Mathers AJ, Yeh AJ,
447 Sifri CD, De Angelis G, Tacconelli E, Villegas MV, Quinn J, Carmeli Y. 2014.

Development and validation of a multiplex PCR assay for identification of the epidemic ST-258/512 KPC-producing Klebsiella pneumoniae clone. Diagn Microbiol Infect Dis 78:12-5.

35. Dallenne C, Da Costa A, Decre D, Favier C, Arlet G. 2010. Development of a set of multiplex PCR assays for the detection of genes encoding important beta-lactamases in Enterobacteriaceae. J Antimicrob Chemother 65:490-5.

36. Carattoli A, Bertini A, Villa L, Falbo V, Hopkins KL, Threlfall EJ. 2005. Identification of plasmids by PCR-based replicon typing. J Microbiol Methods 63:219-28.

37. Mathers AJ, Cox HL, Kitchel B, Bonatti H, Brassinga AK, Carroll J, Scheld WM, Hazen KC, Sifri CD. 2011. Molecular dissection of an outbreak of carbapenem-resistant enterobacteriaceae reveals Intergenus KPC carbapenemase transmission through a promiscuous plasmid. MBio 2:e00204-11.

38. Griffith KL, Wolf RE, Jr. 2002. Measuring beta-galactosidase activity in bacteria: cell growth, permeabilization, and enzyme assays in 96-well arrays. Biochem Biophys Res Commun 290:397-402.

**Table 1.** Minimal inhibitory concentrations (MIC) for parent strains CAV1016, CAV1746, CAV1438 and *E. coli* transformants AMGH-1, ACGH, ACGH2 containing plasmids bearing Tn*4401*b, Tn*4401*a and Tn*4401*h respectively from the parent strains. MICs were determined using the VITEK2 automated susceptibility testing platform and broth microdilution.

| | | | | **VITEK2 Result** | | | **Broth Microdilution** | |
|---|---|---|---|---|---|---|---|---|
| Tn*4401* variant | Strain | Plasmid size, PCR replicon type and putative plasmid backbone | Species | Meropenem MIC (µg/ml) | Ertapenem MIC (µg/ml) | Cefepime MIC (µg/ml) | Meropenem MIC (µg/ml) | Cefepime MIC (µg/ml) |
| Tn*4401*b | CAV1016 | | *K. pneumoniae* | ≥16 | 4 | 2 | 256 | 64 |
| Tn*4401*a | CAV1746 | | *K. pneumoniae* | ≥16 | ≥8 | ≥64 | 1024 | 512 |
| Tn*4401*h | CAV1438 | | *E. cloacae* | ≥16 | ≥8 | ≥64 | 1024 | 512 |
| | ***E.coli* transformed strain (parent strain)** | | | | | | | |
| Tn*4401*b | AMGH-1 (CAV1016) | 43 kb, non-typeable, confirmed pKPC_UVA01 (CP017937.1) (25) | *E. coli* | 0.5 | ≤0.5 | ≤1 | 1 | 8 |
| Tn*4401*a | ACGH (CAV1746) | >60 kb, IncR, putatively resembles pKPC-484 (CP008798.1) | *E. coli* | ≥16 | ≥8 | ≥64 | 16 | 128 |
| Tn*4401*h | ACGH2 (CAV1438) | >70 kb, IncR, putatively resembles pKPC_CAV1176 (CP011661.1)(6) | *E. coli* | ≥16 | 4 | 4 | 4 | 64 |

478 **Figure Legends**

479 Figure 1. Structural variation in published isoforms of the Tn*4401* transposons. Highlighted
480 yellow sequence denotes the reference Tn*4401*b (EU176013). Gray sequence bars denote areas
481 of 100% identity to the reference; black sequence bars denote areas of nucleotide variation.
482

483 Figure 2. Alignment of the promoter regions of $bla_{KPC}$ isoforms with putative promoters as
484 defined in Naas et al. -35 regions and -10 regions are highlighted in grey. +1 indicates the P2
485 transcription start site previously identified by RACE (Rapid Amplification of cDNA Ends). The
486 isoforms a and h have deletion differences in the intervening sequence between P2 and P1. The
487 extent of the intervening sequence that is deleted in isoform h is underlined and denoted IVS.
488 The start codon of $bla_{KPC}$ is highlighted in grey. The arrowhead indicates the site of fusions to the
489 *lacZ* reporter.

490 Figure 3. Differences in $bla_{KPC}$ expression normalized to *rpoB* and relative to Tn*4401*b. Tn*4401*a
491 had significantly higher β-galactosidase activity compared with Tn*4401*b and Tn*4401*h ($p<0.05$),
492 and the Tn*4401*h putative promoter sequence had significantly higher β-galactosidase activity
493 compared with Tn*4401*b ($p<0.05$).
494

495 Figure 4 (A and B). 4A. Schematic to explain constructs showing the promoter regions cloned in
496 *lacZ* fusions. The arrow indicates the P2 transcription start site. 4B. β-galactosidase activity level
497 associated with putative promoter regions of Tn*4401*a, Tn*4401*b and Tn*4401*h isoforms; putative
498 promoters P1, P1+IVS, P2, P+IVS. The *lacZ* reporter plasmid pRS551(vector) served as a
499 negative control. Reporter constructs: P2 had significantly higher β-galactosidase activity

compared with vector, P1, P1+IVS ($p<0.05$); P2+IVS had significantly higher β-galactosidase activity relative to P2, P1, P1+IVS (student t-test: $p<0.05$); promoter sequence Tn*4401*a had significantly higher β-galactosidase activity compared with P2+IVS ($p<0.05$).

Figure 1. Structural variation in published isoforms of the Tn*4401* transposons. Highlighted yellow sequence denotes the reference Tn*4401*b (EU176013). Gray sequence bars denote areas of 100% identity to the reference; black sequence bars denote areas of nucleotide variation.

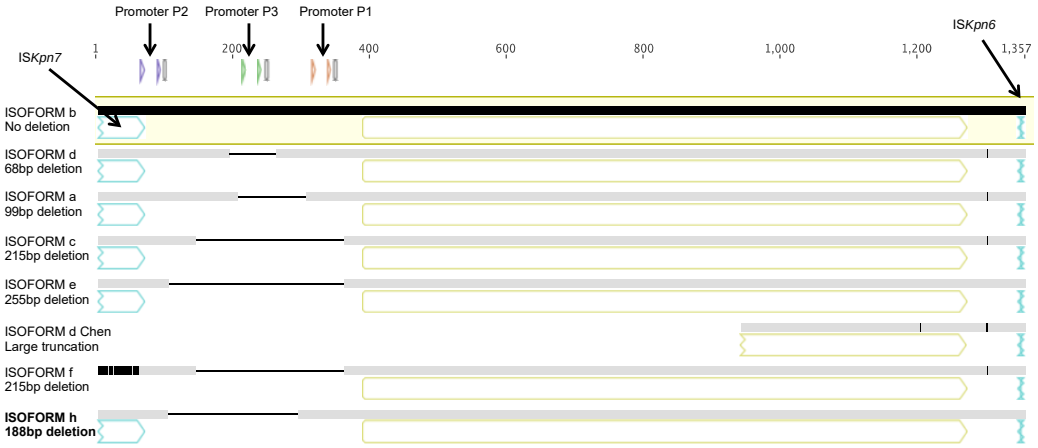


Figure 2. Alignment of the promoter regions of *bla*$_{KPC}$ isoforms with putative promoters as defined in Naas et al. -35 regions and -10 regions are highlighted in grey. +1 indicates the P2 transcription start site previously identified by RACE (Rapid Amplification of cDNA Ends). The isoforms a and h have deletion differences in the intervening sequence between P2 and P1. The extent of the intervening sequence that is deleted in isoform h is underlined and denoted IVS. The start codon of *bla*$_{KPC}$ is highlighted in grey. The arrowhead indicates the site of fusions to the *lacZ* reporter.

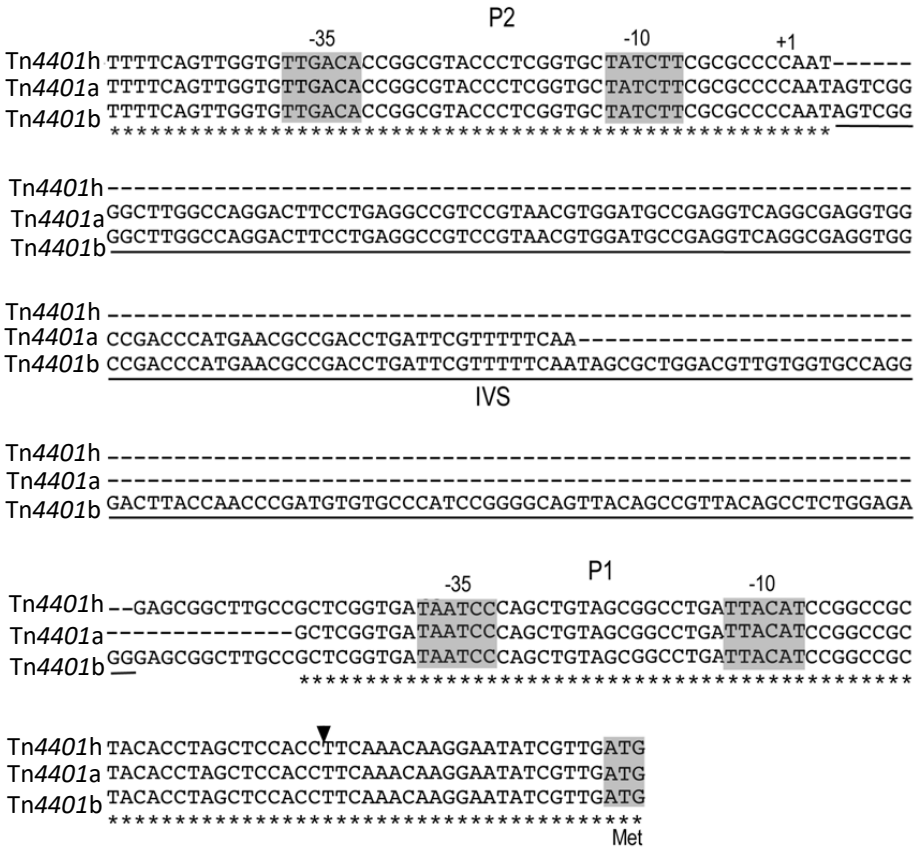


Figure 3. Differences in $bla_{KPC}$ expression normalized to *rpoB* and relative to Tn*4401*b. Tn*4401*a had significantly higher β-galactosidase activity compared with Tn*4401*b and Tn*4401*h ($p<0.05$), and the Tn*4401*h putative promoter sequence had significantly higher β-galactosidase activity compared with Tn*4401*b ($p<0.05$).

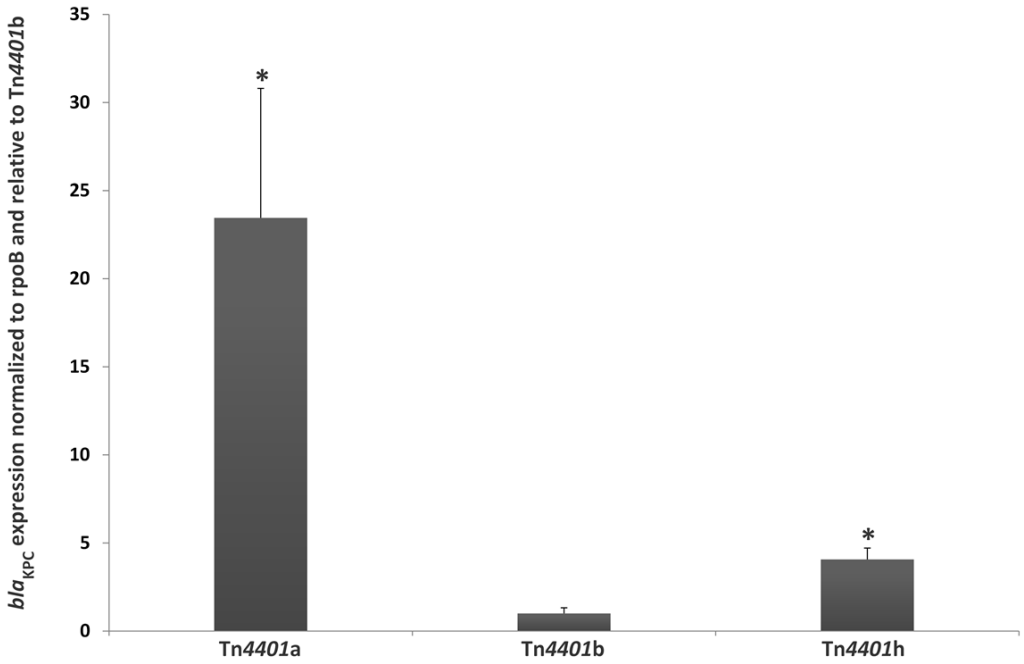
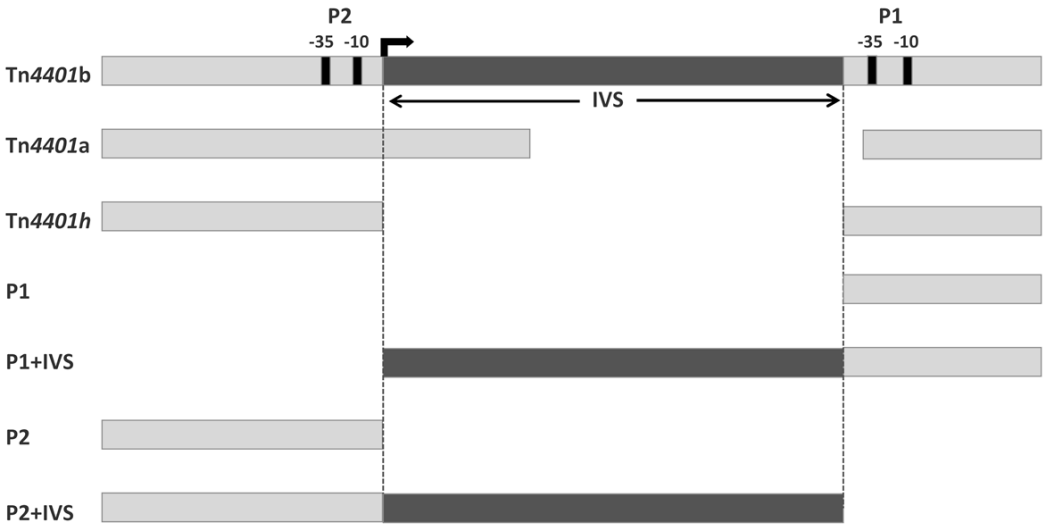


Figure 4 (A and B). 4A. Schematic to explain constructs showing the promoter regions cloned in lacZ fusions. The arrow indicates the P2 transcription start site. 4B. β-galactosidase activity level associated with putative promoter regions of Tn*4401*a, Tn*4401*b and Tn*4401*h isoforms; putative promoters P1, P1+IVS, P2, P2+IVS. The *lacZ* reporter plasmid pRS551(vector) served as a negative control. Reporter constructs: P2 had significantly higher β-galactosidase activity compared with vector, P1, P1+IVS ($p<0.05$); P2+IVS had significantly higher β-galactosidase activity relative to P2, P1, P1+IVS (student t-test: $p<0.05$); promoter sequence Tn*4401*a had significantly higher β-galactosidase activity compared with P2+IVS ($p<0.05$).

4A.

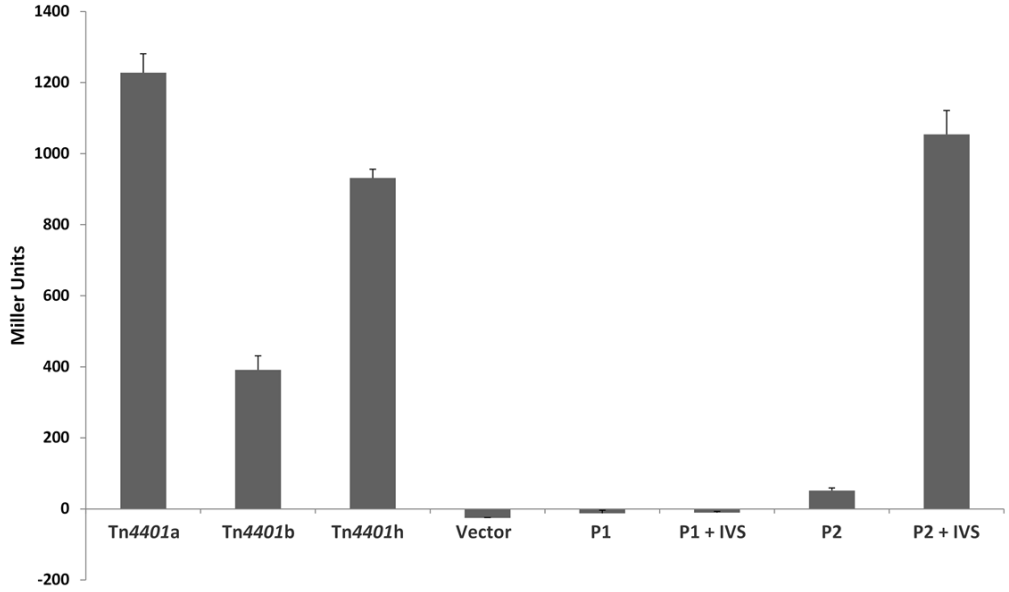


542

543

544

545 4B.




546

547

548

549

550