

RESEARCH ARTICLE

Open Access

# The frequency of tetracycline resistance genes co-detected with respiratory pathogens: a database mining study uncovering descriptive trends throughout the United States

Matthew D Huff<sup>1\*</sup>, David Weisman<sup>2</sup>, John Adams<sup>3</sup>, Song Li<sup>4</sup>, Jessica Green<sup>1</sup>, Leslie L Malone<sup>1</sup> and Scott Clemmons<sup>1</sup>

## Abstract

**Background:** The Center for Disease Control and Prevention (CDC) indicates that one of the largest problems threatening healthcare includes antibiotic resistance. Tetracycline, an effective antibiotic that has been in use for many years, is becoming less successful in treating certain pathogens. To better understand the temporal patterns in the growth of antibiotic resistance, patient diagnostic test records can be analyzed.

**Methods:** Data mining methods including frequent item set mining and association rules via the Apriori algorithm were used to analyze results from 80,241 Target Enriched Multiplex-PCR (TEM-PCR) reference laboratory tests. From the data mining results, five common respiratory pathogens and their co-detection rates with tetracycline resistance genes (TRG) were further analyzed and organized according to year, patient age, and geography.

**Results:** From 2010, all five pathogens were associated with at least a 24% rise in co-detection rate for TRGs. Patients from 0–2 years old exhibited the lowest rate of TRG co-detection, while patients between 13–50 years old displayed the highest frequency of TRG co-detection. The Northeastern region of the United States recorded the highest rate of patients co-detected with a TRG and a respiratory pathogen. Along the East–west gradient, the relative frequency of co-detection between TRGs and respiratory pathogens decreased dramatically.

**Conclusions:** Significant trends were uncovered regarding the co-detection frequencies of TRGs and respiratory pathogens over time. It is valuable for the field of public health to monitor trends regarding the spread of resistant infectious disease, especially since tetracycline continues to be utilized a treatment for various microbial infections. Analyzing large datasets containing TEM-PCR results for co-detections provides valuable insights into trends of antibiotic resistance gene expression so that the effectiveness of first-line treatments can be continuously monitored.

## Background

In 1948, tetracycline was one of the first antibiotics discovered that disrupted ribosomal subunit assembly in bacterial pathogens. Tetracycline was advantageous over existing agents because the medication was inexpensive and had broad-spectrum response, but with a lower level of allergenic properties [1]. As a result, physicians began to prescribe the antibiotic to patients for many bacterial infections [2]. Over time, tetracycline appeared in agriculture, as farmers learned to fortify animal feeds with

the antibiotic as it promoted growth and decreased the spread of bacterial pathogens [3-6]. Tetracycline was also sprayed on crops to halt the transmission of plant diseases [7]. By 1990, the rampant utilization of tetracycline had resulted in an increase in the number of tetracycline resistant infections.

The United States Centers for Disease Control and Prevention (CDC) has identified antibiotic resistance as one of the largest problems threatening global health. Antibiotic resistance directly costs healthcare over \$20 billion annually; after including lost productivity, the total costs exceed \$35 billion [8]. These enormous costs of tetracycline resistance may be largely unnecessary, as over 50% of the antibiotics prescribed are unnecessary

\* Correspondence: matthew.huff@diatherix.com

<sup>1</sup>Diatherix Laboratories Inc., 601 Genome Way, Suite 2100, Huntsville, AL 35806, USA

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

[8]. The CDC, The United States Food and Drug Administration (FDA), and The United States Department of Agriculture are trying to modify policies regarding the prescription of medical antibiotics and their usage in agriculture to attempt to halt the growth of antibiotic resistance [9]. However, agriculturalists are reluctant to follow these policy changes because of a lack of factual evidence linking excess usage to rising levels of resistance. Even if this aspect of tetracycline resistance is widely disputed, the CDC strongly argues that tracking trends in existing antibiotic infections is a key weapon in preventing new antibiotic resistant infections.

Target Enriched Multiplex-Polymerase Chain Reaction (TEM-PCR™) (Diatherix Laboratories Inc. Huntsville, AL) has the ability to simultaneously detect Methicillin Susceptible *Staphylococcus aureus* (*S. aureus*), Methicillin Resistant *Staphylococcus aureus* (MRSA), *Streptococcus pneumoniae* (*S. pneumoniae*), *Haemophilus influenzae* (*H. influenzae*), *Moraxella catarrhalis* (*M. catarrhalis*), and tetracycline resistance genes (TRG) tet (K) and tet (M) [10-14]. TEM-PCR is a nucleic acid multiplex technology that includes two main amplification steps: target enrichment and subsequent exponential amplification [10-14]. For the target enrichment step, two pairs of nested gene-specific primers (*Fo*, forward out; *Fi*, forward in; *Ri* reverse in, *Ro* reverse out) are designed for each target. Low concentrations of the nested primers are used to amplify the target DNA, which will be exponentially amplified in the next step. The *Fi* and *Ri* primers contain a unique tag sequence that can be recognized by a universal set of forward and biotin-labeled reverse SuperPrimers. In the exponential amplification step, the SuperPrimers are present at a concentration that allows for asymmetric PCR amplification of the biotin-labeled single-stranded DNA targeted by the reverse SuperPrimer. Target-specific detection probes are then coupled to carboxyl groups on a bead-based substrate, and probe fluorescence is read using a bead-based detection platform.

Previous literature has documented the rise in tetracycline resistant pathogens and warned of the possible medical consequences regarding antimicrobial immunity [1,2,15-17]. This descriptive study analyzed a dataset of 80,421 patient TEM-PCR test results, and explored the co-detections between tetracycline resistance genes and the respiratory pathogens. By elucidating these co-detection patterns, the study contributes to the knowledge surrounding tetracycline resistance and aids in accomplishing the ultimate goals of optimizing patient care and reducing unnecessary healthcare costs.

## Methods

### TEM-PCR

Nucleic acid extraction for TEM-PCR was performed using a KingFisher™ system (Thermo Scientific). The

amplification steps were completed using a GeneAmp® PCR System 9700 thermocycler (Applied Biosystems). Probe fluorescence was read using a Luminex detection platform. TEM-PCR is a proprietary testing assay with detailed procedures and SOPs under the ownership of Diatherix Laboratories Inc.

### Data mining methods

The two techniques used to elucidate relationships in these data are *frequent item sets* and *association rules* via the Apriori algorithm [18-22]. Briefly, the Apriori algorithm is a widely used algorithm that is commonly applied to transactional datasets. Frequent item sets are groups of DNA markers that occur more frequently than a minimum *support* level in the database. For example, in this dataset, TRG and *S. pneumoniae* were co-detected in 19.8% of the tests that identified at least one marker, which is above the minimum support count of 50 positive tests. The Apriori algorithm assembles frequent item sets into association rules that indicate conditional frequency. For example, the rule {A, B} => {C} indicates that when A and B are found, C is also frequently found. The *confidence* of an association rule indicates its conditional probability in the dataset, that is  $P(C | A, B)$ . A rule's interestingness measure *lift* indicates the degree of independence between the antecedent and the consequent, with values greater than 1 indicating an interesting rule.

Full details of support count, support, confidence, and lift can be found in publications [23,24]. Data mining methods described here were recently used in many biomedical publications [19,23-26]. One publication that clearly explains the methods applied in this study includes a paper by Naulaerts et al. [27]. The confidence, support, lift, and support count values from the Apriori algorithm promoted further co-detections analysis of the 5 selected respiratory pathogens and TRGs.

### Retrospective co-detection analysis of reference laboratory data

Using reference laboratory TEM-PCR data, 80,241 patient tests results were analyzed for the detection of *S. aureus*, MRSA, *S. pneumoniae*, *H. influenzae*, *M. catarrhalis*, and TRGs. 49,840 (62.1%) tests were detected for one pathogen or TRG. The co-detection rates between each pathogen and TRG were further analyzed by year, patient age by year, and geography. In all categorical subsets, results were analyzed for all respiratory samples from November 9, 2009-August 7, 2013. The analysis of *M. catarrhalis* began on September 30, 2011 after genetic targets for this pathogen were added to the multiplexing assay. Data reflecting multiple samples assayed from a given patient on a single date were removed from the analysis. The sample size for each year was within the range of 6,396 (2009) < X <

26,841 (2012). Patient age was organized into 4 groups (0 (exclusive)-2 (inclusive)], (2-13], (13-50], and (50-100]. 91.1% of patients aged (0-2], 94.7% of patients aged (2-13], 92.4% patients aged (13-50], and 76.6% of patients aged (50-100] had specimen collected by nasal, nasalpharangeal, or throat swab. 8.7% of patients age (50-100] had sputum samples collected. Other specimen sources which represented a very small percentage of patients included aspirate, mucus, or bronchial or nasal washes. Results were included from each state where samples were collected within the sample size range of 60 (Maryland) < X < 16,985 (Tennessee). Patient sample size for each categorical subset including year, age group, and geographical state can be seen in Additional file 1: Tables S1-S7. Data analyzed comes from the frequency of patients that tested positive for the simultaneous presence of one pathogen and a TRG. Further molecular testing or bacterial plating by culture was not done to verify if these pathogens were the source of detected TRGs and further traditional testing was out of the scope of this study. Inclusion criteria were defined as submission for respiratory pathogen testing, without regard to age, demographic categories, or other medical conditions.

All of the database information was archived and organized using Microsoft SQL Server Management Studio 2012. Graphical analysis and statistics including scatter plots with loess smoothing, confidence intervals, variance, and heat maps were performed in the R language, with ggplot2 [28,29]. Confidence intervals were at 95% confidence and constructed via the Jefferys prior [30]. Tables were generated in Microsoft Excel 2012. All reference laboratory tests were conducted at Diatherix Laboratories Inc. in Huntsville, AL.

## Results

The Apriori algorithm created a list of the 25 most frequent item sets found throughout the respiratory results database (Table 1). The largest measurements of support included TRGs (40.5%), *S. pneumoniae* (36.3%), *H. influenzae* (28.7%), and *M. catarrhalis* (25.1%). *S. pneumoniae* and *H. influenzae* were co-detected with TRGs 19.8% and 15.1% of all positive tests results. MRSA and *S. aureus* were detected 12.4% and 10.0% of all positive results, respectively. 6.5% of all positive tests included co-detection for MRSA and TRGs and 4.1% included simultaneous detection between *S. aureus* and TRGs.

From the frequent item set results, the Apriori algorithm identified the 25 association rules with the highest levels of support (Table 2). The top of the table includes the rules {*S. pneumoniae*} => {Tetracycline}, {*H. influenzae*} => {Tetracycline}, and {*H. influenzae*, *M. catarrhalis*} => {*S. pneumoniae*} with 19.8%, 15.1%, and 7.2% support, respectively. Each of the highlighted association rules

**Table 1 Frequent item sets**

Item sets	Support count	Support
<b>{Tetracycline}</b>	<b>20,207</b>	<b>40.54%</b>
<b>{<i>S. pneumoniae</i>}</b>	<b>18,113</b>	<b>36.34%</b>
<b>{<i>H. influenzae</i>}</b>	<b>14,281</b>	<b>28.65%</b>
<b>{<i>M. catarrhalis</i>}</b>	<b>12,491</b>	<b>25.06%</b>
<b>{<i>S. pneumoniae</i>, Tetracycline}</b>	<b>9,850</b>	<b>19.76%</b>
<b>{<i>H. influenzae</i>, Tetracycline}</b>	<b>7,543</b>	<b>15.13%</b>
<b>{<i>H. influenzae</i>, <i>S. pneumoniae</i>}</b>	<b>6,231</b>	<b>12.50%</b>
<b>{MRSA}</b>	<b>6,175</b>	<b>12.39%</b>
<b>{<i>S. pneumoniae</i>, <i>M. catarrhalis</i>}</b>	<b>5,864</b>	<b>11.77%</b>
<b>{Tetracycline, <i>M. catarrhalis</i>}</b>	<b>5,864</b>	<b>11.04%</b>
{CoxEchovirus1}	5,027	10.09%
<b>{<i>S. aureus</i>}</b>	<b>4,980</b>	<b>9.99%</b>
{Rhinovirus}	4,151	8.33%
<b>{<i>H. influenzae</i>, <i>M. catarrhalis</i>}</b>	<b>4,089</b>	<b>8.20%</b>
<b>{<i>H. influenzae</i>, <i>S. pneumoniae</i>, Tetracycline}</b>	<b>3,603</b>	<b>7.23%</b>
<b>{MRSA, Tetracycline}</b>	<b>3,066</b>	<b>6.15%</b>
{Parainfluenza}	3,027	6.07%
<b>{<i>S. pneumoniae</i>, Tetracycline, <i>M. catarrhalis</i>}</b>	<b>2,987</b>	<b>5.99%</b>
<b>{<i>H. influenzae</i>, <i>S. pneumoniae</i>, <i>M. catarrhalis</i>}</b>	<b>2,486</b>	<b>4.99%</b>
{ <i>K.pneumoniae</i> }	2,160	4.33%
<b>{<i>S. aureus</i>, Tetracycline}</b>	<b>2,025</b>	<b>4.06%</b>
<b>{<i>H. influenzae</i>, Tetracycline, <i>M. catarrhalis</i>}</b>	<b>1,966</b>	<b>3.94%</b>
{Human Metapneumovirus}	1,955	3.92%
{ <i>A.baumannii</i> }	1,801	3.61%

The first 25 frequent items with the highest support and support count are listed. The boldface rows indicate items that contain relationships between 1 or more of the pathogens or TRGs in this study.

indicates a pathogen of interest or TRG that all have confidence measurements over 50.9% and lift values over 1.3. This confidence threshold establishes that, given the antecedent of a rule, the consequent is found in at least 50.9% of the cases. Since each rule had a lift measurement over 1, they were all considered to be rules of interest. From these results, and from literature associating TRGs with specific gram-positive bacteria, we sought to elucidate temporal co-detection patterns to better understand how patient age and geography relate to TRG co-detection.

Figure 1 displays a scatterplot including the frequencies of respiratory pathogens co-detected with TRGs by year. It is clear that the frequency of co-detection between each pathogen and TRG increased with time. From January 2010- August 2013, the percentage of gram-positive bacteria *S. aureus*, MRSA, and *S. pneumoniae* co-detected with TRGs rose 43%, 26%, and 27%, respectively. From January 2010- August 2013, the frequency of gram-negative bacteria *H. influenzae* and *M. catarrhalis* co-

**Table 2 Association rules**

Association rules	Support count	Support	Confidence	Lift
{ <i>S. pneumoniae</i> }= > {Tetracycline}	9,850	19.76%	54.38%	1.34
{ <i>H. influenzae</i> }= > {Tetracycline}	7,543	15.13%	52.82%	1.30
{ <i>H. influenzae, S. pneumoniae</i> }= > {Tetracycline}	3,603	7.23%	57.82%	1.43
{ <i>S. pneumoniae, M. catarrhalis</i> }= > {Tetracycline}	2,987	5.99%	50.94%	1.26
{Tetracycline, <i>M. catarrhalis</i> }= > { <i>S. pneumoniae</i> }	2,987	5.99%	54.27%	1.49
{ <i>H. influenzae, M. catarrhalis</i> }= > { <i>S. pneumoniae</i> }	2,486	4.99%	60.80%	1.67
{ <i>K. pneumoniae</i> }= > {Tetracycline}	1,518	3.05%	70.28%	1.73
{ <i>H. influenzae, S. pneumoniae, M. catarrhalis</i> }= > {Tetracycline}	1,311	2.63%	52.74%	1.30
{ <i>H. influenzae, Tetracycline, M. catarrhalis</i> }= > { <i>S. pneumoniae</i> }	1,311	2.63%	66.68%	1.83
{ <i>A.baumannii</i> }= > {Tetracycline}	957	1.92%	53.14%	1.31
{CoxEchovirus1, <i>S.pneumoniae</i> }= > {Tetracycline}	888	1.78%	55.64%	1.37
{CoxEchovirus1, <i>S.pneumoniae</i> }= > { <i>S.pneumoniae</i> }	888	1.78%	55.88%	1.54
{ <i>P.aeruginosa</i> }= > {Tetracycline}	812	1.63%	55.43%	1.36
{PVLgene}= > {MRSA}	805	1.62%	93.71%	7.56
{MRSA, <i>S. pneumoniae</i> }= > {Tetracycline}	761	1.53%	61.56%	1.52
{ <i>S.pyogenes</i> }= > {Tetracycline}	689	1.38%	71.84%	1.77
{ <i>S. aureus, S. pneumoniae</i> }= > {Tetracycline}	651	1.31%	62.59%	1.54
{CoxEchovirus1, <i>H.influenzae</i> }= > {Tetracycline}	646	1.30%	52.48%	1.29
{CoxEchovirus1, <i>H.influenzae</i> }= > { <i>S.pneumoniae</i> }	634	1.27%	51.50%	1.42
{CoxEchovirus1, <i>M.catarrhalis</i> }= > { <i>S.pneumoniae</i> }	627	1.26%	50.64%	1.39
{Rhinovirus, <i>S.pneumoniae</i> }= > {Tetracycline}	594	1.19%	51.65%	1.27
{Rhinovirus, Tetracycline}= > { <i>S.pneumoniae</i> }	594	1.19%	53.66%	1.48
{MRSA, <i>M. catarrhalis</i> }= > {Tetracycline}	574	1.15%	55.84%	1.37
{MRSA, <i>H. influenzae</i> }= > {Tetracycline}	556	1.12%	60.30%	1.49
{PVLgene}= > {Tetracycline}	516	1.04%	60.00%	1.48

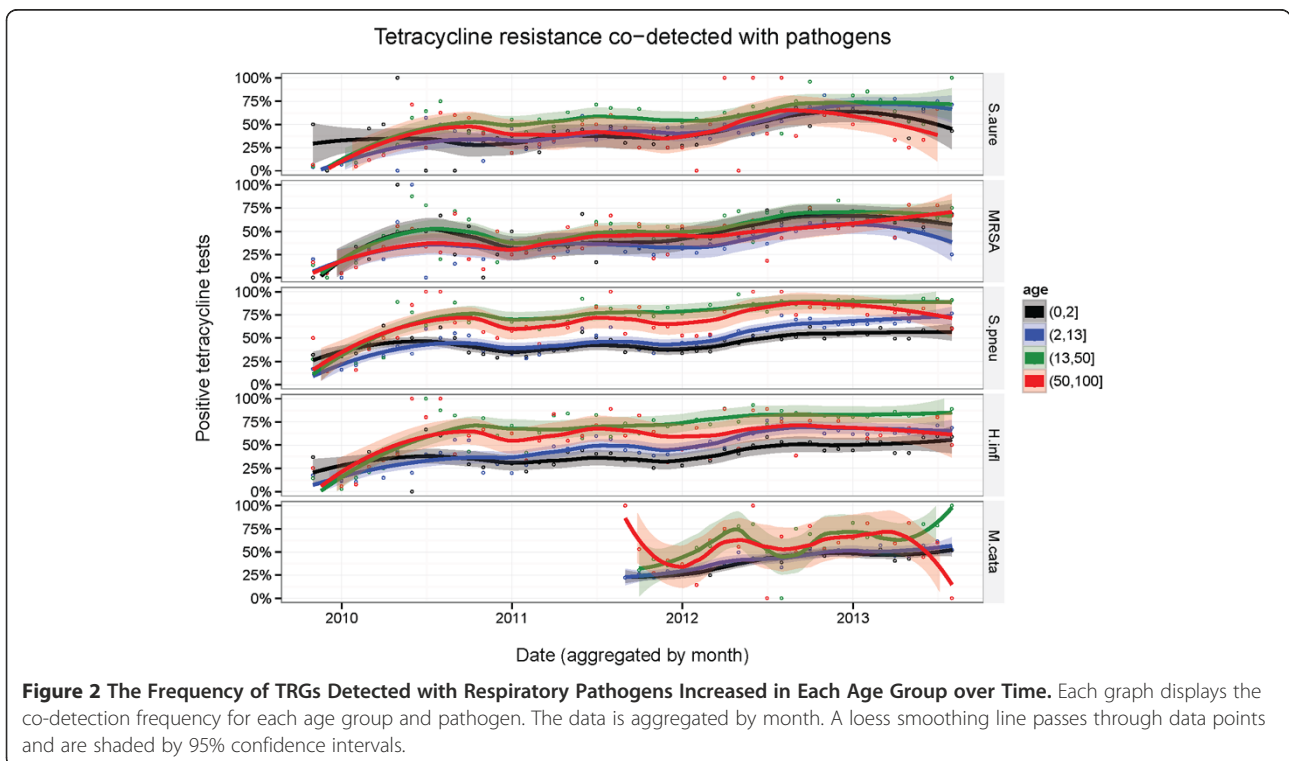
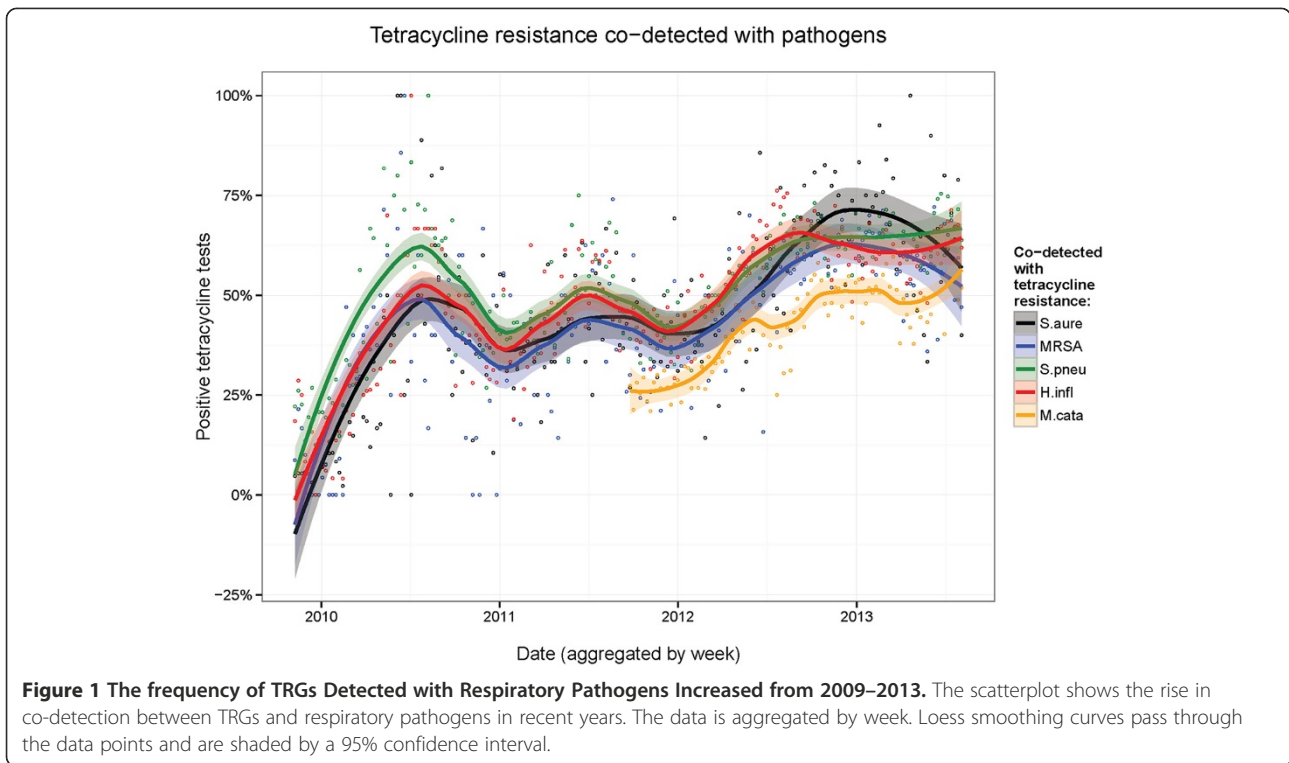
The first 25 association rules with the highest support and support count are listed. Other calculations include confidence and lift which have measurements over 52.8% and 1.3, respectively for all boldfaced rows.

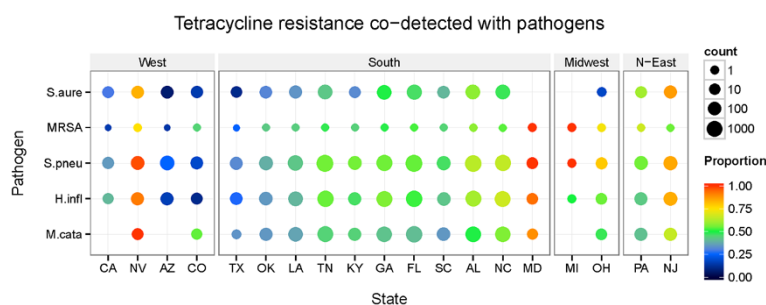
detected with TRGs rose 32% and 24%, respectively. Over each year, the rate of TRGs co-detected with these five potential pathogens rose a minimum of 3.9%.

Figure 2 displays the time-series co-detection frequency of TRGs with each pathogen segmented by 4 age groups. The patient sample size and total number of positive tests for each pathogen are included in Additional file 1: Tables S1-S7. In each age group, the rates of co-detection between TRGs and each pathogen increased with time. Older patients showed a larger co-detection rate than younger patients. For *H. influenzae*, *M. catarrhalis*, and *S. pneumoniae*, patients (0–2] had the lowest rate of co-detection, followed by patients (2–13] years of age, then patients (50–100] years of age, concluded by patients (13–50] years of age, which had the highest rate of co-detection. In 2013, there was an average 33%, 22%, and 26% difference in co-detection rate from children aged (0–2] to patients age (13–50] for *H. influenzae*, *M. catarrhalis*, and *S. pneumoniae*, respectively. In 2013, an alarming 82% patients of age (13–50] had TRGs co-detected

with *S. pneumoniae*. For *S. aureus* and MRSA, the age group (13–50] had the highest rate of co-detection between TRGs and the designated pathogen, but all other age groups had similar and co-detection frequencies over time.

Figure 3 includes a heatmap indicating the co-detection frequencies of TRGs and each of the 5 respiratory pathogens separated by state from November 2009- August 2013. Additional file 1: Table S9 displays the number of positive tests recorded from each state and the frequency that the pathogen was co-detected with a TRG. The highest frequency of co-detection occurred in the Northeast, especially Maryland and New Jersey. In the Southeast, Florida, Alabama, and Georgia accounted for the largest percentage of TRG co-detection. Along a westward gradient, the percentages of TRG co-detection dropped dramatically. Louisiana, Oklahoma, Texas, and Colorado all had co-detection frequencies under 53%. The level of TRGs co-detected from patients in New Jersey for the five pathogens was between 54.4% and 86.9%. Moving westward, the frequencies of TRGs co-detected was lower in





**Figure 3 Heatmap Displaying the Longitudinal Effect of TRG Co-detection across the United States.** The heat map is ordered by state on the bottom where the eastern states are on the right and the western states are on the left side of the map. The circles which are most red are the states that have the highest levels of tetracycline resistance co-detection regarding each pathogen, while states that are the most blue have the lowest rates. Size of the circles correspond to sample size of the data.

Tennessee with levels between 42.7% and 55.3% for the five pathogens. In Texas, the rate of co-detection decreased even further to between 33.3% and 10.7%. Since Eastern states reported a much higher frequency of co-detection over Western states, a longitudinal trend was established.

### Discussion

Figure 1 indicates that the frequency of co-detection between the 5 respiratory pathogens and TRGs continues to rise over time. One trend observed in Figure 1 is the periodicity, where the co-detection rates of TRGs and each pathogen reach a maximum during the middle of each year. The frequency of co-detection is the highest in the summer and the lowest in the winter of each year. Various literature report a seasonal incidence of antibiotic resistance in various pathogens [31-33]. It is possible that this trend occurs from fluctuation in the number of tests physicians administer to patients during different times of the year. Alternatively, the seasonal trend here may be occurring because of the different levels of antibiotics prescribed to patients throughout the each season, especially during the flu season [33]. The CDC considers the respiratory illness/flu season in the United States to begin in October and end in May, peaking in February [34]. As a result of the rise in number of antibiotics prescribed during February, an increased number of pathogens activate the production of TRGs [35]. Throughout spring months, TRG up-regulation and transmission continues to increase until months after flu season subsides. As a result, the rise in TRG co-detection frequency seen each summer could be a delayed effect from the increase in tetracycline exposure from the rise in number of antibiotic prescriptions that are administered during early winter/spring of each year. Moreover, these seasonal data could also have a clinical impact as drug therapy could be periodically altered and antibiotic effectiveness could be increased. Physicians may have greater success prescribing tetracycline in the

winter when co-detection frequency is lower, than in the summer when co-detection is at maximum levels. Seasonal trends could be occurring with other first-line antibiotics as well, but more studies are needed to test this hypothesis.

Figure 2 illustrates that all 4 age groups rose in co-detection frequency over time, but the frequencies of co-detection follow different trends for the age groups. Across the 5 pathogens, patients age (0-2) display the lowest frequency of co-detection of TRGs and respiratory pathogens, while patients age (13-50) demonstrate the highest rate of co-detection. These data suggest that tetracycline usage can be optimized in different age groups over the yearly cycle. Even though tetracycline has the highest chance of being effective in patients age (0-2), physicians typically avoid prescribing tetracycline to younger children, as the treatment can cause staining of permanent teeth [36,37]. Once dental maturity is reached, the staining of teeth from tetracycline use is no longer an issue. The avoidance of prescribing tetracycline to infant patients could be an underlying reason for the lower co-detection measurement in the youngest age bracket. Alternatively, one approved use of tetracycline is to combat dermatological pathogens and is a therapeutic option for treating young adults with acne [31,34]. The increased utilization of tetracycline in teenage patients could be an underlying cause for the higher TRG co-detection frequency observed in the (13-50) age group. In addition, environmental factors such as diet may have an effect on the large co-detection frequency of both (13-50) and (50-100) age groups. Since over half the antibiotics produced for use in the United States are utilized for agricultural purposes, it may be plausible to hypothesize that the longer a person ingests tetracycline-infused foods, the greater likelihood that normal microbial flora build an immunity to the antibiotic. Then, when a respiratory pathogen infects a patient, these flora can pass tetracycline genes to the foreign pathogen via conjugation [38].

There are two possible explanations for why a longitudinal effect is observed in Figure 3. Physicians in the Northeast could be prescribing tetracycline to patients at higher rates than in other areas of the country. Alternatively, various diets and varieties of food eaten across the United States could have an impact on different frequencies of TRG and respiratory pathogen co-detection. An underlying reason for the varying frequencies of TRG co-detection throughout different regions of the United States could be because of geographical differences regarding tetracycline application in agriculture. The only outlier in the Western states is Nevada, which has one of the largest rates of co-detection, but is surrounded by states that have the lowest levels of co-detection of TRGs and respiratory pathogens. There is no apparent hypothesis for why Nevada has a large co-detection frequency. Taken together, these data suggest an environmental influence, but a subsequent analysis of data estimating per-state tetracycline loads would be needed to correlate these data to co-detection frequencies.

There are two common causes for increased tetracycline exposure. The first includes excess prescriptions being provided to patients by physicians and the other involves overexposure in agriculture and livestock feed. It is likely that these pathogens contained genes and pathways that predated the antibiotic era that could neutralize tetracycline [39,40]. Increased utilization in the fields of agriculture and medicine allowed these time-weighted genetic determinants such as synthesis of inactivation enzymes, efflux pumps, or ribosomal protection proteins to be up-regulated and passed to other pathogens via plasmid transfer [39-42]. *S. aureus*, MRSA, *S. pneumoniae*, *H. influenzae*, and *M. catarrhalis* are all found in proximity in the upper respiratory tract of most humans, therefore, it is plausible that these pathogens exchange genetic information frequently. The genomes of many pathogens contain TRGs with the three major genes including tet (K), tet (M), and tet (O) [1,15,16]. In this study, only genetic targets for tet (K) and tet (M) were included in testing. With the exclusion of tet (O) and other tet genes, it can be speculated that the actual rate of tetracycline resistance in patients is even higher than seen here.

A strong hypothesis for the high frequencies of TRG co-detection is that the high tetracycline use in agriculture facilitates tetracycline resistant gene activations, which is then passed to other pathogens via conjugation. The rampant use of antibiotics is a more pressing issue in the United States than in other countries because over half the antibiotics that are produced in the United States are used for agricultural purposes, an extremely high percentage [3,4]. For many years it has been disputed whether or not antibiotic use in agriculture is responsible for the rise in the number of antibiotic resistant pathogens. The FDA argues antibiotic levels recommended for

agricultural use have been tested in a laboratory setting on bacteria in culture, with very low frequencies of pathogens acquiring resistance [43]. However, some literature demonstrates that sub-inhibitory antibiotic exposure can cause development of resistance over time [44-47]. Recently, the FDA has begun to regulate the amount of antibiotics utilized in the agricultural and medical industries [9]. Playing close attention to the levels of pathogens that are becoming resistant to antibiotics is imperative because throughout the past 15 years only 9 new antibiotics have been approved to combat certain illnesses [8]. Since there are a limited number of new antibiotic medications in the developmental pipeline, it is important to monitor levels of tetracycline resistant pathogens so that the medication can continue to be an effective treatment for acne and uncomplicated community-acquired MRSA [48-50].

## Conclusion

After a data mining analysis, it is clear that there is a relationship between *S. aureus*, MRSA, *S. pneumoniae*, *H. influenzae*, and *M. catarrhalis* and the co-detection of TRGs. Since 2010, there has been a rise in TRG co-detection by at least 24% for each of the five potential respiratory pathogens studied and co-detection frequency is more widespread in summer months than in winter months. Patients aged (13–50] show the highest rate of TRG co-detection with respiratory pathogens, while patients aged (0–2] show the lowest frequency. These findings could correlate with the deleterious effect of tetracycline prescriptions given to younger children because of medication side-effects or the large amount of teenagers who are prescribed tetracycline for the effective treatment of acne. Partitioned by geography, the Northeastern United States demonstrated the largest percentage of TRG co-detection, while co-detection rates decreased longitudinally westward.

The CDC has realized that antibiotic resistance is the largest threat to future global health and these data further underscore the need for careful management of antibiotics. By establishing certain trends regarding tetracycline resistance, we can begin to uncover underlying causes for the rise in TRG co-detection frequency. These data are consistent with the hypothesis that environmental pressures in medicine and agriculture effect levels of TRGs. As a result, many people who have been educated on the overuse of antibiotics in society have switched to strict organic diets so that the risk of coming in contact with TRGs is minimized. Moreover, the FDA, CDC, and Department of Agriculture are beginning to revise regulations regarding antibiotic usage. The overuse of antibiotics in society is an ongoing controversial topic. However, the continuous monitoring of co-detection rates between pathogens and antibiotic resistance genes is imperative so

that physicians can assess the effectiveness of an antibiotic at the time of infection and provide optimal patient care.

Further studies are needed to generalize these descriptive findings to a larger population and to associate environmental factors with the rise in tetracycline resistance. We hope that this study fosters a rich discussion of this vitally important medical topic.

#### Ethics statement

This study was approved by Chesapeake Internal Review Board (IRB) (Columbia, MD) which has full accreditation with the Association for the Accreditation of Human Research Protection Programs (AAHRPP) (Protocol #: Pro00010357). Using the Department of Health and Human Services regulations at 45 CFR 46, the IRB determined that this study does not constitute human subject research and, therefore, does not require IRB oversight. All procedures employed followed the guidelines in the Declaration of Helsinki and HIPAA. This study analyzed a dataset containing patient age, gender, U.S. state, test date, and test results. Patient identity was not available to the researchers in this study.

#### Additional file

**Additional file 1: Table S1.** Percentages of Positive Tests Detected with TRGs aggregated by Year. The table includes sample size, number of tests, number of positive tests co-detected with TRGs, and % of positive tests co-detected with TRGs for each year and pathogen. **Table S2.** Percentage of Tests Positive for *S. aureus* and TRGs Categorized by Age and Year. **Table S3.** Percentage of Tests Positive for MRSA and TRGs Categorized by Age and Year. **Table S4.** Percentage of Tests Positive for *S. pneumoniae* and TRGs Categorized by Age and Year. **Table S5.** Percentage of Tests Positive for *H. influenzae* and TRGs Categorized by Age and Year. **Table S6.** Percentage of Tests Positive for *M. catarrhalis* and TRGs Categorized by Age and Year. Each table (Tables S2-S6) includes the sample size, number of positive tests, number of tests co-detected with a TRG, and % of tests co-detected for a TRG for each pathogen in this study that is categorized by age group and year. **Table S7.** Percentage of Positive Tests detected with TRGs Categorized by State. The table includes the sample size, number of positive tests, number of tests co-detected with a TRG, and % of tests co-detected for a TRG for each pathogen in this study organized by state.

#### Abbreviations

CDC: Center for disease control and prevention; FDA: Food and drug administration; RES: Respiratory panel; TEM-PCR: Target enriched multiplex polymerase chain reaction; TRGs: Tetracycline resistance genes.

#### Competing interests

MDH, JG, LLM, SC are all full time employees of Diatherix Laboratories Inc. DW is a consultant paid by Diatherix Laboratories Inc. JA and Diatherix collaborate in conducting case studies. SL is an employee of The Hudson Alpha Institute of Biotechnology.

#### Authors' contributions

MDH generated the hypothesis for this paper, conducted initial data mining, wrote, organized, and revised manuscript and figures, and did majority of research. DW provided knowledge of data mining methods for biological applications, conducted data mining, and initially generated all figures. JA provided his medical knowledge and expertise of infectious disease in helping revise manuscript drafts. SL provided assistance in draft revisions. JG lead a team of molecular technologists during the TEM-PCR testing process

for this study. LLM researched and developed the RES panels used in this study. SC innovated and developed all computer programs used at Diatherix Laboratories Inc. that organized patient data in the database used for this study. All authors read and approved the final manuscript.

#### Acknowledgements

We would like to thank all of the Diatherix clinical laboratory employees who successfully conducted over 80,000 TEM-PCR respiratory panel tests for this study. Without their work, we would not have any data saved to conduct meaningful research. Also, we offer a special thanks to Donna Hockman who was one of the pioneer designers for our current respiratory panel. Without her hard work, we would not have the panels that were used in this manuscript to diagnose patients for respiratory pathogens or TRGs. All funding for this study was provided by Diatherix Laboratories Inc.

#### Author details

<sup>1</sup>Diatherix Laboratories Inc., 601 Genome Way, Suite 2100, Huntsville, AL 35806, USA. <sup>2</sup>Department of Biology, University of Massachusetts Boston, 100 Morrissey Blvd., Boston, MA 02125-3393, USA. <sup>3</sup>Knoxville Infectious Disease Consultants, P.C., 2210 Sutherland Ave., Suite 110, Knoxville, TN 37919, USA. <sup>4</sup>Hudson Alpha Institute of Biotechnology, 601 Genome Way, Huntsville, AL 35806, USA.

Received: 30 January 2014 Accepted: 18 August 2014

Published: 25 August 2014

#### References

1. Chopra I, Roberts M: Tetracycline antibiotics: mode of action, applications, molecular biology, and epidemiology of bacterial resistance. *Microbiol Mol Biol Rev* 2001, **62**:232–260.
2. Speer B, Shoemaker N, Salyeers A: Bacterial resistance to tetracycline: mechanisms, transfer, and clinical significance. *Clin Microbiol Rev* 1992, **5**:387–399.
3. Lipsitch M, Singer R, Levin B: Antibiotics in agriculture: When is it time to close the barn door? In *Proceedings of the National Academy of Sciences of the United States of America*. *PNAS* 2002, **99**:5752–5754.
4. Soo K: USDA invites comments on Use of tetracycline, other substances in organic agriculture. [http://www.ams.usda.gov]
5. Schmitt H, Stood K, Hamscher G, Smit K, Seinen W: Tetracyclines and tetracycline resistance in agricultural soils: microcosm and field studies. *Microb Ecol* 2006, **51**:267–76.
6. Stanford JC, Aminov RI, Krapac IJ, Garrigues-Jeanjean N, Mackie RL: Occurrence and diversity of tetracycline resistance genes in lagoons and groundwater underlying two swine production facilities. *Appl Environ Microbiol* 2001, **64**:1494–1502.
7. Mcmanus P, Stockwell V, Sundin G, Jones A: Antibiotic use in plant agriculture. *Phytopathology* 2002, **40**:443–465.
8. CDC: Antibiotic resistance threats in the United States. [http://www.cdc.gov/drugresistance/threat-report-2013/pdf/ar-threats-2013-508.pdf]
9. Johnson SR: FDA moves to curb use of medically important antibiotics in livestock. *Modern Health* 2011, [http://www.modernhealthcare.com/article/20131211/NEWS/312119951/fda-moves-to-curb-use-of-medically-important-antibiotics-in]
10. Han J, Swan DC, Smith SJ, Lum SH, Sefers SE, Unger ER, Tang Y: Simultaneous amplification and identification of 25 human Papillomavirus types with Plexplex technology. *J Clin Microbiol* 2006, **44**(11):4157–4162.
11. Brunstein J, Thomas E: Direct screening of clinical specimens for multiple respiratory pathogens using Genaco respiratory panels 1 and 2. *Diag Mol Pathol* 2006, **15**:169–73.
12. Li H, McCormac MA, Estes RW, Sefers SE, Dare RK, Chapell JD: Simultaneous detection and high throughput identification of a panel of RNA viruses causing respiratory tract infections. *J Clin Microbiol* 2007, **45**:2105–2109.
13. Zou S, Han J, Wen L, Liu Y, Cronin K, Lum SH: Human influenza A virus (H5N1) detection by a novel multiplex PCR typing method. *J Clin Microbiol* 2007, **45**:1889–92.
14. Benson R, Tondella ML, Bhatnagar J, Carvalho MS, Sampson JS, Talkington DF: Development and evaluation of a novel multiplex PCR technology for molecular differential detection of bacterial respiratory disease pathogens. *J Clin Microbiol* 2008, **46**:2074–7.



15. Aires J, Butel MJ: **Tetracycline resistance mediated by tet (W), tet (M), and tet (O) genes of *Bifidobacterium* isolates from humans.** *Applied Environmental Microbiology* 2007, **73**:2751–2754.
16. Li W, Atkinson G, Thakor N, Allas U, Chuao-chao L, Kwok-Yan C, Tenson T, Schulten K, Wilson K, Hauryliuk V, Frank J: **Mechanism of tetracycline resistance by ribosomal protection protein Tet(O).** *Nature Communications* 2013, **4**:1447.
17. Schmitz F, Krey A, Sadurski R, Verhoef R, Milatovic D, Fluit A: **Resistance to tetracycline and distribution of tetracycline resistance genes in European *Staphylococcus aureus* isolates.** *J Antimicrob Chemother* 2001, **47**:239–240.
18. Charikar M, Chen K, Farach-Colton M: **Finding Frequent Items in Data Streams.** In *Proceedings of the 29<sup>th</sup> International Colloquium on Automata, Languages, and Programming ICALP*. 2002:693–703.
19. Park J, Chen M, Yu P: **An effective hash-based algorithm for mining association rules.** In *Proceedings of the 1995 ACM SIGMOD international conference on Management of data ACM SIGMOD*. 1995:175–186.
20. Savasere A, Omiecinski E, Navathe S: **An efficient algorithm for mining association rules in large databases.** In *Proceedings of the 21<sup>st</sup> International Conference on Very Large Databases VLDB*, 444. 1995:432–444.
21. Toivonen H: **Sampling Large Databases for Association Rules.** In *Proceedings of the 22<sup>nd</sup> International Conference on Very Large Databases VLDB*. 1996:134–145.
22. Agrawal R, Srikant R: **Fast algorithms for mining association rules in large databases.** In *Proceedings of the 20th International Conference on Very Large Data Bases*. Santiago, Chile: VLDB; 1994:487–499.
23. Harpaz R, Chase H, Friedman C: **Mining multi-item drug adverse effect associations in spontaneous reporting systems.** *BMC Bioinformatics* 2010, **11**(Suppl 9):S7.
24. Lamma E, Mello P, Nanetti A, Riguzzi F, Storari S, Valastro G: **Artificial intelligence techniques for monitoring dangerous infections.** *IEEE Trans Inf Technol Biomed* 2006, **10**:143–155.
25. Brossette S, Sprague A, Hardin M, Waites K, Jones W, Moser S: **Association rules and data mining in hospital infection control and public health surveillance.** *J Am Med Inform Assoc* 1998, **5**:373–381.
26. Wright A, Chen E, Maloney F: **An automated technique for identifying association between medications, laboratory results and problems.** *J Biomed Inform* 2010, **43**:891–901.
27. Naulaerts S, Meysman P, Bittremieux W, Vu T, Berghe W, Goethals B, Laukens K: **A primer to frequent itemset mining for bioinformatics.** *Briefings in Bioinformatics* 2013, bbt074+.
28. Wickham H: **ggplot2.** [http://ggplot2.org/]
29. Cleveland WS, Grosse E, Shyu WM: **Local regression models.** In *Chapter 8 of Statistical Models in S*. Edited by Chambers JM, Hastie TJ. Wadsworth & Brooks/Cole.
30. Brown LD, Cai TT, DasGupta A: **Interval estimation for a binomial proportion.** *Statistical Science* 2001, **16**:101–133.
31. Goossens H, Ferech M, Stichele RV, Elseviers M: **Outpatient antibiotic use in Europe and associated with resistance: a cross-national database study.** *The Lancet* 2005, **365**(9459):579–587.
32. Hassani L, Imzilil B, Boussaid A, Gauthier J: **Seasonal incidence of and antibiotic resistance among *Aeromonas* species isolated from domestic wastewater before and after treatment in stabilization ponds.** *Microb Ecol* 1992, **23**(3):227–237.
33. Pathak SP, Bhattacharjee JW, Ray PK: **Seasonal variation in survival and antibiotic resistance among various bacterial population in a tropic river.** *J Gen Appl Microbiol* 1993, **39**(1):47–56.
34. CDC: **The Flu Season.** http://www.cdc.gov/flu/about/season/flu-season.htm.
35. University of Illinois at Chicago: **How some bacteria survive antibiotics.** *Science Daily* 2008. http://www.sciencedaily.com/releases/2008/04/080430154945.htm.
36. Chiappinelli J, Walton R: **Tooth discoloration resulting from long-term tetracycline therapy: a case report.** *Quintessence Int: Esthetic Dentistry* 1992, **23**:539–541.
37. Sanchez AR, Rodgers RS, Sheridan PJ: **Tetracycline and other tetracycline-derivative staining of the teeth and oral cavity.** *Int J Dermatol* 2004, **43**:709–15.
38. Davies J: **Inactivation of antibiotics and the dissemination of resistance genes.** *Science* 1994, **264**(5157):375–382.
39. Finley RL, Collignon P, Larsson DGJ, McEwen SA, Li XZ, Gaze WH, Reid-Smith R, Timinouni M, Graham DW, Edward T: **The scourge of antibiotic resistance: the important role of the environment.** *Clinical Infectious Diseases*. [Online May 2 2013]. doi:10.1093/cid/cit355.
40. D'Costa VM, King CE, Kalan L, Morar M, Sung WWL, Schwarz C, Froese D, Zazula G, Calmels F, Debruyne R, Golding GB, Poinar HN, Wright GD: **Antibiotic resistance is ancient.** *Nature* 2011, **477**(7365):457–461.
41. Bahl M, Sorensen S, Hansen L, Licht T: **Effect of tetracycline of transfer and establishment of the tetracycline-inducible conjugative transposon Tn916 in the guts of gnotobiotic rats.** *Applied Environmental Microbiology* 2004, **70**:758–764.
42. Wang Y, Rotman ER, Shoemaker NB, Salyers AA: **Translational control of tetracycline resistance and conjugation in the Bacteroides conjugative transposon CTnDOT.** *J Bacteriol* 2005, **187**:2673–80.
43. Connell S, Tracz D, Nierhaus K, Taylor D: **Ribosomal protection proteins and their mechanism of tetracycline resistance.** *Antimicrob Agents Chemother* 2003, **47**:3675–3681.
44. Deneve C, Bouttier S, Dupuy B, Barbut F, Collignon A, Janoir C: **Effects of subinhibitory concentrations of antibiotics on colonization factor expression by maxifloxacin-susceptible and moxifloxacin-resistance *Clostridium difficile* strains.** *Antimicrob Agents Chemother* 2009, **53**:5155–5162.
45. Drayson R, Leggat T, Wood M: **Increased antibiotic resistance post-exposure to sub-inhibitory concentrations is independent of capsular polysaccharide production in *E. coli*.** *Journal of Experimental Microbiology and Immunology* 2011, **15**:36–42.
46. Henderson-Begg S, Livermore D, Hall L: **Effect of subinhibitory concentrations of antibiotics on mutation frequency in *Streptococcus pneumoniae*.** *J Antimicrob Chemother* 2006, **57**:849–854.
47. Wu YL, Scott EM, Po A, Tariq N: **Development of resistance and cross-resistance in *Pseudomonas aeruginosa* exposed to subinhibitory antibiotic concentrations.** *APMIS* 1999, **107**:585–592.
48. Bhambri S, Kim G: **Use of Oral Doxycycline for Community-acquired Methicillin-resistant *Staphylococcus aureus* (CA-MRSA) infections.** *J Clin Aesthet Dermatol* 2009, **2**(Suppl 4):45–50.
49. Amichai B, Grunwald MH: **Tetracyclines and their uses in dermatology.** *J Dermatolog Treat* 1998, **9**:221–229.
50. Eliopoulos G, Roberts M: **Tetracycline therapy: update.** *Clin Infect Dis* 2003, **36**:462–467.

doi:10.1186/1471-2334-14-460

**Cite this article as:** Huff et al.: The frequency of tetracycline resistance genes co-detected with respiratory pathogens: a database mining study uncovering descriptive trends throughout the United States. *BMC Infectious Diseases* 2014 **14**:460.

**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

