Mosaic tetracycline resistance genes encoding ribosomal protection proteins

Philip J. Warburton1*, Nina Amodeo2 and Adam P. Roberts3

1School of Biomedical and Healthcare Sciences, Plymouth University Peninsular Schools of Medicine and Dentistry, School of Biomedical and Healthcare Sciences, Plymouth, UK; 2Department of Biomedical and Forensic Science, Faculty of Science and Technology, Anglia Ruskin University, Cambridge, UK; 3Department of Microbial Diseases, UCL Eastman Dental Institute, University College London, London, UK

*Corresponding author. Tel: +44-1752-584482; Fax: +44-1752-586788; E-mail: philip.warburton@plymouth.ac.uk

First reported in 2003, mosaic tetracycline resistance genes are a subgroup of the genes encoding ribosomal protection proteins (RPPs). They are formed when two or more RPP-encoding genes recombine resulting in a functional chimera. To date, the majority of mosaic genes are derived from sections of three RPP genes, tet(O), tet(W) and tet(32), with others comprising tet(M) and tet(S). In this first review of mosaic genes, we report on their structure, diversity and prevalence, and suggest that these genes may be responsible for an under-reported contribution to tetracycline resistance in bacteria.

Introduction

Tetracyclines bind to the A-site on the bacterial ribosome, resulting in steric blocking of the aminoacyl-tRNA binding site, which prevents protein synthesis.1 They are effective against both Gram-positive and Gram-negative bacteria and, due to the relative lack of major side effects and cheap cost, have been used extensively in the treatment of infections2 as well as growth promoters in animal husbandry.3

Bacterial resistance to tetracycline is often mediated through the acquisition of DNA encoding proteins that confer resistance by one of three main mechanisms: ATP-dependent efflux, enzymatic inactivation of tetracycline, or ribosomal protection.2 To date, a total of 60 different classes of tetracycline resistance gene, including oxytetracycline resistance genes, have been reported. These include 33 predicted or proven to encode active efflux pumps, 12 encoding ribosomal protection proteins (RPPs), 13 encoding inactivating enzymes and 1 reported to confer resistance via an as yet undetermined mechanism, designated tet(U) (a full list is periodically updated by Roberts4). Although it has yet to be assigned a mechanistic class, tet(U) has been identified in Enterococcus and Staphylococcus isolates.5,6 However, a study by Caryl et al.7 reported that when tet(U) was cloned and expressed in Escherichia coli, the transformants were not resistant to tetracycline.

To be considered a new class of tetracycline resistance gene, it must encode a protein <80% identical to known tetracycline resistance proteins.8 Determinants representing new classes were originally assigned a letter from the English alphabet.9 However, as all letters are used, they are now assigned an Arabic numeral,8 with new determinants referred to the Levy group (bonnie.marshall@tufts.edu) in order to obtain a designation prior to publication to avoid duplication and ensure taxonomic consistency.

RPPs

RPPs are a related group of proteins that, when bound to the ribosome, result in the release of tetracycline from the ribosome, enabling protein synthesis to proceed10 (reviewed by Thaker et al.11). Of the 12 classes of RPP gene currently reported (tet(M), (O), (Q), (S), (T), (W), (32), (36), (44), (BP), otr(A) and tet), tet(M) is considered the most prevalent due to its association with the broad host range Tn916/Tn1545 family of conjugative transposons.12 However, a subgroup of RPP genes has been identified that consist of regions of different, already characterized RPP genes that appear to have undergone recombination forming a mosaic gene. It must be stressed here that the progenitors of mosaic genes are assumed based purely on the order in which they were discovered and we cannot be sure of the directionality of mosaic gene formation.

Mosaic RPP genes

In 2003, Stanton and Humphery13 reported two RPP genes in Megasphaera elsdenii that encoded predicted proteins showing 89.1% and 91.9% identity to Tet(W) (accession number AJ222769) from Butyrivibrio fibrisolvens. As this was above the <80% cut-off, they did not qualify as a new resistance class.
under the nomenclature system. However, further analysis of
the amino acid sequence revealed variability in the percentage
identity to Tet(W) across its length. The large central section in
either sequences showed 98.1% identity to Tet(W), while
small sections at the N- and C-terminal ends were found to have
a lower amino acid sequence identity to Tet(W) [between
66.6% and 75.3%]. However, these same N- and C-terminal sections were shown to have between 99.3% and 100%
amiacid identity to Tet(O) (accession number M18896), des-
pite the central section showing identity to Tet(W). Given the
evidence, this suggested recombination had occurred, creating a
mosaic determinant with a central Tet(W) region flanked by
two Tet(O) regions. Although never before observed between
two different RPP classes, recombination resulting in functional
genes has previously been reported between different phylo-
types of tet(M) as well as in other antibiotic resistance
genes, such as penA and pbp2x, which confer resistance to peni-
cillin. Furthermore, in vitro experiments have successfully
recombined teta and tetc to create mosaics that confer
resistance to tetracycline at levels comparable to the non-
mosaic tetc. The guideline for determining a new resistance gene class was
established prior to the discovery of these mosaic RPP genes and
none of the mosaic genes qualified as a new class when analysed
as one single continuous sequence. It was clear, however, that
these mosaic genes were different from their non-mosaic coun-
terparts and that the current classification did not adequately
reflect the evolutionary background of these genes. There-
fore, an expansion of the nomenclature system was sug-
gested whereby the mosaic gene would receive a designation
that reflected the structural order and class of the genes they
comprised, better reflecting their variable nature. For
example, the two resistance genes reported in M. elsdenii, which
comprised a central tet(W) region flanked by two tet(O) regions,
were designated tet(O/W/O).13

Although Stanton and Humphrey were the first to report
mosaic RPP genes, Melville et al. had unknowingly reported a
mosaic gene 2 years previously. This resistance gene, found in Clostridium saccharolyticum K10, encoded a predicted protein that showed 76% amino acid identity to Tet(O) (accession number Y07780). As per the original nomenclature guidelines, it was given the new designation Tet(32). However, subsequent re-examination of the sequence found that only the central section showed <80% identity to known proteins, while the N- and C-terminal regions flanking the central section shared 100% and 97.7% identity, respectively, to Tet(O) (accession number M18896). The central region was still thought to represent a section of a new tet(O/W/O) class and therefore the determinant was reclassified Tet(O/W/O).18 Subsequently, the proposed full, non-mosaic sequences of Tet(32) have been reported in several isolates identified from the human oral cavity, with the Tet(O/W/O) mosaic determinant now showing 89% amino acid identity to these.

Similarly, the previously reported tet(S) allele (accession num-
ber AY534326) on the conjugative transposon Tn9165 has been reclassified as a result of in silico analysis. The amino acid sequence shows identity to Tet(S) across 595 amino acids (1–595 inclusive), with the final 61 amino acids at the C-terminus end identical to Tet(M) (accession number U09422), resulting in it being reclassified as Tet(S/M).24

Mosaic gene diversity

To date, a total of 30 mosaic genes have been reported in the lit-
erature, of which 26 currently have sequences deposited in
GenBank (Table 1). Some studies have reported multiple occur-
rences of known genes; however, many of these have been char-
acterized by PCR amplification only. Structurally, these chimeric
genes currently comprise either two [e.g. tet(O/W)], three [e.g.
tet(O/W/O)], four [e.g. tet(O/W/32/O)] or six [e.g. tet(O/W/32/O/ W/O)] different regions (Figure 1), with tet(O), tet(W) and tet(32)
being the predominant RPP genes reported to form mosaic genes,
comprising all but two of the reported variants, and tet(M) and
tet(S) forming the remaining two.24,25 Given the prevalence of
tet(M) in certain samples, and the previous reports of self-
recombination, it is surprising that there are so few reports of
mosaic genes containing tet(M). Furthermore, alignment of 12 representative RPP gene sequences shows tet(M) sharing 75% and 70% identity, respectively, to tet(O) and tet(44), which is higher than the percentage identity observed between the more commonly reported RPP mosaic genes comprising tet(O), (W) and (32) (Table 2). However, mosaic genes comprising tet(M) and any other gene, with the exception of tet(S), have yet to be reported. It is entirely possible that this may be due to a lack of investigation rather than an absence of recombination followed by fixation of the recombinant allele in the bacterial population. Alternatively, it is possible that there is little selective pressure for tet(M)-based mosaic genes if the resultant protein is no more efficient than Tet(M) itself and/or there is no indirect select-
ive pressure for mosaicism. A similar situation may exist for other proteins, such as Tet(S). Stanton et al. reported that the protein encoded by the tet(O/W/O) mosaic genes in M. elsdenii conferred a higher level of resistance to tetracycline than their non-mosaic counterparts, but similar investigations are still to be reported for other RPP genes. Therefore, the prevalence of certain mosaic gene variants could suggest that they are in some way more beneficial to the host than the non-mosaic genes they comprise.

PCR-based analysis

PCR-based assays have been developed to help researchers
detect specific mosaic genes. Stanton and Humphrey described an assay that distinguished between the non-mosaic genes
tet(O) and tet(W) and the mosaic tet(O/W/O) from Megaspheora
strains, enabling them to detect tet(O/W/O) variants in six addi-
tional M. elsdenii strains. Patterson et al. investigated the pres-
ence of mosaic genes using various specific oligonucleotide sets that either bound within the resistance genes or flanked them. Amplicons specific to tet(O/W), tet(O/32) and tet(W/32) were detected in faecal samples, with tet(O/32) being the most common of these mosaic amplicons; it was amplified in all 12 pig faec-
cal samples and 6 of 7 human faecal samples tested. In contrast, the faecal samples from cows and sheep, as well as human saliva samples, failed to produce any amplicons for these mosaic genes, suggesting they were not present at detectable levels.

Chen et al. also used a oligonucleotide primer set that annealed outside tet(O) to determine the presence of tetracycline
resistance genes in two Streptococcus suis isolates. Although no
tmplicon was produced using internal, tet(O)-specific primers, the primers binding to flanking DNA yielded an amplicon, indicat-
ing the presence of mosaic genes (identified as tet(O/32/O) and
Table 1. A summary of the mosaic tetracycline genes reported to date

<table>
<thead>
<tr>
<th>Gene</th>
<th>Organism</th>
<th>Source(s)</th>
<th>Accession number</th>
<th>Reference(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>tet(O/W)</td>
<td><em>Bifidobacterium thermophilum</em> B0219</td>
<td>environmental (pig slaughterhouse) sample</td>
<td>AM889118</td>
<td>32</td>
</tr>
<tr>
<td>tet(O/W)</td>
<td><em>B. thermophilum</em> B0241</td>
<td>pig faeces</td>
<td>AM889119</td>
<td>32</td>
</tr>
<tr>
<td>tet(O/W)</td>
<td><em>B. thermophilum</em> B0242</td>
<td>pig faeces</td>
<td>AM889120</td>
<td>32</td>
</tr>
<tr>
<td>tet(O/W)</td>
<td><em>B. thermophilum</em> B0253</td>
<td>pig faeces</td>
<td>AM889121</td>
<td>32</td>
</tr>
<tr>
<td>tet(O/W)</td>
<td><em>B. thermophilum</em> B0256</td>
<td>pig faeces</td>
<td>AM889122</td>
<td>32</td>
</tr>
<tr>
<td>tet(O/W)-2</td>
<td><em>Megasphaera elsdenii</em> 25-51</td>
<td>swine faeces</td>
<td>AM889119</td>
<td>32</td>
</tr>
<tr>
<td>tet(O/W)-1</td>
<td><em>M. elsdenii</em> 25-51</td>
<td>swine faeces</td>
<td>AM889119</td>
<td>32</td>
</tr>
<tr>
<td>tet(O/W)-4</td>
<td>unculture bacterial clone</td>
<td>pig faeces</td>
<td>AM889119</td>
<td>32</td>
</tr>
<tr>
<td>tet(O/W)-3</td>
<td><em>M. elsdenii</em> 14-14</td>
<td>pig faeces</td>
<td>AM889120</td>
<td>32</td>
</tr>
<tr>
<td>tet(O/W)-2</td>
<td><em>M. elsdenii</em> 7-11</td>
<td>pig faeces</td>
<td>AM889120</td>
<td>32</td>
</tr>
<tr>
<td>tet(O/W)/2-1</td>
<td>unculture bacterial clone</td>
<td>human faeces</td>
<td>AM889119</td>
<td>32</td>
</tr>
<tr>
<td>tet(O/W)/2-2</td>
<td><em>S. suis</em></td>
<td>human faeces</td>
<td>AM889119</td>
<td>32</td>
</tr>
<tr>
<td>tet(O/W)/2-3</td>
<td><em>S. suis</em></td>
<td>human faeces</td>
<td>AM889119</td>
<td>32</td>
</tr>
<tr>
<td>tet(O/W)/2-4</td>
<td>unculture bacterial clone</td>
<td>human faeces</td>
<td>AM889119</td>
<td>32</td>
</tr>
<tr>
<td>tet(O/W)/2-5</td>
<td>unculture bacterial clone</td>
<td>human faeces</td>
<td>AM889119</td>
<td>32</td>
</tr>
<tr>
<td>tet(O/W)/2-6</td>
<td><em>Dorea longicatena</em> AGR2136</td>
<td>human faeces</td>
<td>AM889119</td>
<td>32</td>
</tr>
<tr>
<td>tet(O/W)/2-7</td>
<td><em>C. coli</em></td>
<td>human faeces</td>
<td>AM889119</td>
<td>32</td>
</tr>
<tr>
<td>tet(O/W)/2-8</td>
<td><em>C. jejuni</em> subspecies jejuni</td>
<td>human faeces</td>
<td>AM889119</td>
<td>32</td>
</tr>
<tr>
<td>tet(O/W)/2-9</td>
<td><em>R. intestinalis</em> X8684</td>
<td>human faeces</td>
<td>AM889119</td>
<td>32</td>
</tr>
<tr>
<td>tet(S/M)</td>
<td><em>S. equinus</em></td>
<td>food</td>
<td>AM889119</td>
<td>32</td>
</tr>
<tr>
<td>tet(S/M)</td>
<td><em>S. intermedius</em></td>
<td>human isolate</td>
<td>AM889119</td>
<td>32</td>
</tr>
</tbody>
</table>
This full-length oligonucleotide primer set does aid the identification of mosaic genes; however, it is only specific for those with regions homologous to tet(O) flanking sequences. Since PCR strategies aimed at identifying resistance genes require knowledge of the sequence of the target, mosaic RPP genes are likely to be largely undetected and under-reported by PCR-based studies.

Reflecting the findings by Patterson et al.,21 almost all the mosaic genes reported to date have originated from faecal samples, with the majority identified from a porcine origin and less commonly from humans (Table 1). The gut houses a complex and diverse bacterial community with potential for widespread horizontal gene transfer, and the mosaic genes found in faecal samples are likely to reflect the pool of non-mosaic genes present within the gut microbiota. Genes such as tet(W) and tet(O) are commonly reported from these types of samples,29 but the prevalence of tet(32)-containing mosaic genes suggests that tet(32) may be more common than initially thought. In fact, tet(O/32/O) was found to be the most common mosaic gene in both the human and pig faecal samples tested and was present in almost as many samples tested as the non-mosaic tet(O) and tet(W) genes.21 In contrast, mosaic genes have not yet been reported in faecal samples from bovine and ovine origin or in human saliva.21 Why they are predominantly found in pigs while as yet unreported in other animals is not immediately clear, though the extensive use of tetracyclines in the swine industry3,30,31 may have contributed to their selection.

### Draft genome analysis

The advent of high-throughput genomic sequencing has led to an increase in the number of genomes being deposited in sequence databases. Many contain tetracycline resistance genes that are generically labelled simply as ‘tetracycline resistance protein’ or as ‘tet(M)-like’, the designation of which may be a result of automated annotation pipelines. A preliminary search of the NCBI nucleotide database, using tet(O) (accession number Y07780) as the query, found that some of these generically labelled tetracycline resistance genes gave a partial match to tet(O). Further examination indicates that some are as yet uncharacterized and unreported mosaic genes, which have been further defined for this review using the nucleotide sequence to determine the crossover points. For example, the tet(M)-like gene (accession number NZ_AUJS01000017, location 41626–43545 bp) in the draft genome of Dorea longicatena AGR2136 from a human faecal sample appears to be a previously unreported variant of tet(O/32/O) (Figure 1).

Furthermore, the tetracycline resistance genes present in Campylobacter jejuni subspecies jejuni 2008-B94, Campylobacter coli 202/04, C. coli 317/04 (accession numbers AI001000025, A1NH01000038 and NZ_AIJ01000054, respectively) and Roseburia intestinalis XB6B4 (accession number FP929050) are also structurally novel variants of tet(O/32/O) (Figure 1). The three mosaic genes present in the Campylobacter spp. are identical to each other, while that in R. intestinalis is different. Taking into account these newly defined genes, the total number of mosaic genes reported increases from 30 to 35 (not including those identified via PCR amplification only; Table 1) and suggests that other generically labelled tetracycline resistance genes
Figure 1. Schematic representation of reported mosaic tetracycline RPP genes. The coded bars indicate sequences of high identity to specific RPP genes: vertical line bars for tet(M), white bars for tet(O), grey bars for tet(S), black bars for tet(W) and checked bars for tet(32). The number above the bar indicates the reported crossover point. aIndicates those sequences that are incomplete or absent in GenBank, with the crossover points taken from the publication. bIndicates sequences that have been analysed in this study due to no specific crossover point(s) reported.
present in the database [e.g. those labelled as tet(M)-like] could be further classified, helping to understand mosaic gene proliferation and diversity.

Conclusions

Our knowledge of the mosaic RPP gene group is steadily increasing since their discovery in 2003, with the majority derived from tet(O), tet(W) and tet(32) and others deriving from tet(M) and tet(S). It is clear that these genes are being under-reported both in terms of experimental detection and also within genomic data. Further work and increased attention on mosaic RPP genes is important if we are to understand the evolutionary selective pressures driving their fixation in bacterial populations and the subsequent effects on resistance and mobile genetic element evolution within their host.

Acknowledgements

P. J. W. initiated this work whilst at Anglia Ruskin University.

Transparency declarations

None to declare.

References


Table 2. Sequence identity matrix showing the percentage nucleotide identity between representatives of all 12 RPP gene classes, in descending order, compared with tet(M)

<table>
<thead>
<tr>
<th>RPP gene</th>
<th>tet(M)</th>
<th>tet(S)</th>
<th>tet(O)</th>
<th>tet(44)</th>
<th>tet(32)</th>
<th>tet(W)</th>
<th>tet(T)</th>
<th>tet(36)</th>
<th>tet(Q)</th>
<th>tetB(P)</th>
<th>otr(A)</th>
<th>tet</th>
</tr>
</thead>
<tbody>
<tr>
<td>tet(M)</td>
<td>100</td>
<td>78</td>
<td>75</td>
<td>70</td>
<td>69</td>
<td>64</td>
<td>57</td>
<td>49</td>
<td>46</td>
<td>23</td>
<td>11</td>
<td>11</td>
</tr>
<tr>
<td>tet(S)</td>
<td>100</td>
<td>70</td>
<td>69</td>
<td>67</td>
<td>62</td>
<td>56</td>
<td>56</td>
<td>48</td>
<td>48</td>
<td>15</td>
<td>12</td>
<td>11</td>
</tr>
<tr>
<td>tet(O)</td>
<td>100</td>
<td>69</td>
<td>69</td>
<td>69</td>
<td>65</td>
<td>56</td>
<td>56</td>
<td>49</td>
<td>48</td>
<td>15</td>
<td>12</td>
<td>11</td>
</tr>
<tr>
<td>tet(44)</td>
<td>100</td>
<td>71</td>
<td>67</td>
<td>67</td>
<td>65</td>
<td>55</td>
<td>49</td>
<td>47</td>
<td>11</td>
<td>12</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>tet(32)</td>
<td>100</td>
<td>74</td>
<td>71</td>
<td>71</td>
<td>64</td>
<td>50</td>
<td>58</td>
<td>46</td>
<td>46</td>
<td>15</td>
<td>11</td>
<td>10</td>
</tr>
<tr>
<td>tet(W)</td>
<td>100</td>
<td>77</td>
<td>70</td>
<td>70</td>
<td>64</td>
<td>50</td>
<td>58</td>
<td>46</td>
<td>47</td>
<td>11</td>
<td>12</td>
<td>10</td>
</tr>
<tr>
<td>tet(T)</td>
<td>100</td>
<td>12</td>
<td>12</td>
<td>12</td>
<td>12</td>
<td>12</td>
<td>12</td>
<td>12</td>
<td>12</td>
<td>12</td>
<td>12</td>
<td>12</td>
</tr>
<tr>
<td>tet(36)</td>
<td>100</td>
<td>13</td>
<td>13</td>
<td>13</td>
<td>13</td>
<td>13</td>
<td>13</td>
<td>13</td>
<td>13</td>
<td>13</td>
<td>13</td>
<td>13</td>
</tr>
<tr>
<td>tet(Q)</td>
<td>100</td>
<td>14</td>
<td>14</td>
<td>14</td>
<td>14</td>
<td>14</td>
<td>14</td>
<td>14</td>
<td>14</td>
<td>14</td>
<td>14</td>
<td>14</td>
</tr>
<tr>
<td>tetB(P)</td>
<td>100</td>
<td>15</td>
<td>15</td>
<td>15</td>
<td>15</td>
<td>15</td>
<td>15</td>
<td>15</td>
<td>15</td>
<td>15</td>
<td>15</td>
<td>15</td>
</tr>
<tr>
<td>otr(A)</td>
<td>100</td>
<td>16</td>
<td>16</td>
<td>16</td>
<td>16</td>
<td>16</td>
<td>16</td>
<td>16</td>
<td>16</td>
<td>16</td>
<td>16</td>
<td>16</td>
</tr>
<tr>
<td>tet</td>
<td>100</td>
<td>17</td>
<td>17</td>
<td>17</td>
<td>17</td>
<td>17</td>
<td>17</td>
<td>17</td>
<td>17</td>
<td>17</td>
<td>17</td>
<td>17</td>
</tr>
</tbody>
</table>

Accession numbers of representative genes included in the matrix: tet(M), U09422; tet(O), Y07780; tetB(P), AE001437; tet(Q), X58717; tet(S), X92946; tet(T), L42544; tet(W), AJ222769; tet(32), DQ647324; tet(36), AJ514254; tet(44), FN594949; otr(A), X53401; tet, AL939106.

Shaded boxes represent those genes currently reported to comprise mosaic genes.
17 Rubin RA, Levy SB. Interdomain hybrid Tet proteins confer tetracycline resistance only when they are derived from closely related members of the tet gene family. J Bacteriol 1990; 172: 2303–12.