

Article

The Bacterial Genomic Context of Highly Trimethoprim-Resistant DfrB Dihydrofolate Reductases Highlights an Emerging Threat to Public Health

Claudèle Lemay-St-Denis ^{1,2,3} , Sarah-Slim Diwan ^{1,2,3} and Joelle N. Pelletier ^{1,2,3,4,*}

- ¹ Department of Biochemistry and Molecular Medicine, Université de Montréal, Montréal, QC H3T 1J4, Canada; claudèle.lemay-st-denis@umontreal.ca (C.L.-S.-D.); sarah.slim.diwan@umontreal.ca (S.-S.D.)
- ² PROTEO, The Québec Network for Research on Protein, Function, Engineering and Applications, Québec, QC G1V 0A6, Canada
- ³ CGCC, Center in Green Chemistry and Catalysis, Montréal, QC H3A 0B8, Canada
- ⁴ Chemistry Department, Université de Montréal, Montréal, QC H2V 0B3, Canada
- * Correspondence: joelle.pelletier@umontreal.ca

Abstract: Type B dihydrofolate reductase (*dfrb*) genes were identified following the introduction of trimethoprim in the 1960s. Although they intrinsically confer resistance to trimethoprim (TMP) that is orders of magnitude greater than through other mechanisms, the distribution and prevalence of these short (237 bp) genes is unknown. Indeed, this knowledge has been hampered by systematic biases in search methodologies. Here, we investigate the genomic context of *dfrbs* to gain information on their current distribution in bacterial genomes. Upon searching publicly available databases, we identified 61 sequences containing *dfrbs* within an analyzable genomic context. The majority (70%) of those sequences also harbor virulence genes and 97% of the *dfrbs* are found near a mobile genetic element, representing a potential risk for antibiotic resistance genes. We further identified and confirmed the TMP-resistant phenotype of two new members of the family, *dfrb10* and *dfrb11*. *Dfrbs* are found both in *Betaproteobacteria* and *Gammaproteobacteria*, a majority (59%) being in *Pseudomonas aeruginosa*. Previously labelled as strictly plasmid-borne, we found 69% of *dfrbs* in the chromosome of pathogenic bacteria. Our results demonstrate that the intrinsically TMP-resistant *dfrbs* are a potential emerging threat to public health and justify closer surveillance of these genes.

Keywords: antibiotic resistance; resistome; type II dihydrofolate reductase; genomic context; mobile genetic elements



Citation: Lemay-St-Denis, C.; Diwan, S.-S.; Pelletier, J.N. The Bacterial Genomic Context of Highly Trimethoprim-Resistant DfrB Dihydrofolate Reductases Highlights an Emerging Threat to Public Health. *Antibiotics* **2021**, *10*, 433. <https://doi.org/10.3390/antibiotics10040433>

Academic Editor: Elena Perrin

Received: 23 March 2021

Accepted: 12 April 2021

Published: 13 April 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Trimethoprim (TMP) is a synthetic antimicrobial that is ranked as being highly important by the World Health Organization (WHO) [1]. TMP strongly and selectively inhibits a key enzyme in bacterial folate biosynthesis, chromosomal dihydrofolate reductase (K_I (*Escherichia coli* FoaA) = 20 pM [2]), thereby effectively abolishing bacterial proliferation. This antimicrobial was initially introduced for clinical application in 1968 in combination with sulfamethoxazole, also an inhibitor of folate biosynthesis, and later used alone to treat various infections [3,4]. TMP is widely prescribed to adults and to children, as well as to animals, worldwide [5–7]. In 2017, increasing concern over antibiotic resistance prompted the WHO to issue recommendations that include the reduction of TMP usage with food-producing animals as well as a complete restriction of its use with animals to promote growth and for preventive measures [8]. The goal of these recommendations is to lower the prevalence of antimicrobial resistance bacteria that could be transmitted to humans.

Multiple TMP resistance mechanisms have been reported [3]. The main mechanisms are the acquisition of type A (DfrA) or type B (DfrB) TMP-resistant dihydrofolate reductases.

These TMP-resistant enzymes are expressed in addition to the TMP-sensitive, chromosomal FoaA, allowing folate synthesis and bacterial survival. The DfrA family is homologous to FoaA. It includes nearly 40 members that are 150–190 amino acids in length, similar to the FoaA family [9]. The DfrB family currently consists of eight members; they are homotetrameric enzymes of 78 amino acids per protomer [10]. Contrary to DfrAs, they are phylogenetically and structurally unrelated to FoaA [10,11]. Their evolutionary origin is currently unknown. DfrBs maintain full activity at the clinical concentrations of TMP that fully inhibit FoaA [12]. They offer TMP resistance at concentrations at least 3 orders of magnitude greater than DfrAs [13], thus conferring TMP resistance that cannot be countered by administering TMP at higher concentrations.

DfrB1 is the first member of the DfrB family to have been reported. Over the past decades, DfrB1 has been characterized in great detail for its unique structure, its biophysical characteristics, its multimerization and its robustness [14–18]. In particular, DfrBs are distinguished from most enzymes in the fact that their single, central active site requires distinct contribution from each of the four identical protomers, creating an evolutionary conundrum [10,17]. Recently, other members of the DfrB family have been functionally characterized by our group, showing nearly indistinguishable dihydrofolate reductase activity, high resistance to TMP and similar inhibition by recently-reported inhibitors of distinct classes [19–22]. Nonetheless, at the outset of this study, none of the widely-used databases (CARD, ARDB and ARG-ANNOT) contained all eight known *dfrbs* sequences [23–25].

The prevalence of *dfrb* genes in clinical and environmental samples is currently unknown. The *dfrb* genes have rarely been reported in clinical samples [26,27]. Although this could be interpreted as the scarce presence of DfrBs in the collection of resistance genes in bacteria (resistome), it is important to note that the short 237-bp *dfrb* genes have not been routinely searched for. Until recently, gene-prediction algorithms used 300 nt as a cut-off to differentiate short non-protein-coding RNAs (ncRNA) from messenger RNAs (mRNA) [28]. Even now, sophisticated gene prediction algorithms such as the widely used Prodigal are unlikely to predict *dfrbs* as a result of their unusual codon usage and small size, both of which are penalized [29].

Experimental detection of TMP-resistant DfrB enzymes has also consistently failed to detect *dfrb* genes because of the prevalence of PCR-based methods: primers specific to *dfras* are used, with few or no primers specific to the unrelated *dfrbs* [30,31]. Fortunately, the advent of whole-genome sequencing now allows for large-scale computational screening of antimicrobial resistance gene databases beyond experimental biases.

According to Martínez et al., the greatest public health risk is observed when resistance genes to widely-used antibiotics are found on mobile genetic elements (MGE) of a human pathogen [32]. To date, a limited number of reports have found *dfrbs* near integrases and transposases, indicative of their genomic mobility [27,33]. Nonetheless, as *dfrbs* are rarely reported, their genomic context is essentially unexplored and the current public health risk that they represent is unknown.

Here, we searched publicly available databases to identify sequences containing *dfrbs*. We investigated the predicted pathogenicity of the organisms harbouring each sequence as well as the genomic context of *dfrbs*. We found that 70% of sequences containing *dfrbs* harbor virulence genes, mostly from *Pseudomonas*, a major cause of infection in humans that is difficult to treat because of its evolved resistance [34]. Overall, 97% of *dfrbs* are in proximity to a mobile genetic element, favoring their dissemination. Importantly, this investigation resulted in the identification of two new members of the *dfrb* family; we expressed both and confirmed their highly TMP-resistant phenotype in vitro. Our results demonstrate that the intrinsically TMP-resistant DfrBs can be found in a variety of contexts that are consistent with the transmission of multidrug resistance and justify closer surveillance of these genes.

2. Results

2.1. Expansion of the DfrB Family

Our first objective was to determine whether further DfrB homologues could be identified, to join the small but rapidly growing DfrB family. Using profile hidden Markov models (HMM) of the six functionally characterized DfrBs (DfrB1–5 and DfrB7) [19], we searched the TrEMBL database. In addition to confirming the presence of DfrB1–9, we also identified two genes displaying high homology to the conserved core of DfrBs yet sharing sequence identity of less than 95% to any known *dfrb* were identified. They were named *dfrb10* and *dfrb11* (Table 1).

Table 1. Information and MICs on the newly identified DfrB10 and DfrB11.

New Name	UniprotKB Accession Number	Genbank Accession Number	Closest Characterized DfrB (Protein Identity/DNA Identity) ^a	MIC (µg/mL)
DfrB10	A0A2Z1CLP9	ALZ46148.1	DfrB3 (92%/93%)	>600
DfrB11	A0A2N2TNN4	PKO69073.1	DfrB3 (90%/87%)	>600

^a Protein sequence identity of all members of the DfrB family are reported in Supplementary Materials Table S1.

DfrB10 was found on the p12969-DIM mega-plasmid (0.4 Mb) from a *Pseudomonas putida* strain isolated in China in 2013 from a patient suffering from pneumonia [35]. DfrB11 was identified in a groundwater sample at the Horonobe Underground Research laboratory in Japan in 2017, in a *Betaproteobacteria* sequence [36]. Both new DfrBs procure the same phenotype as the other DfrBs when overexpressed in *E. coli*: they confer resistance to 0.6 mg/mL TMP, the highest concentration of TMP that can be solubilized in 5% methanol (Table 1). This situates these genes amongst the most resistant dihydrofolate reductases known to date.

2.2. Identification of Bacterial Sequences

The DNA sequences of the eight previously reported *dfrbs* and the two new *dfrbs* were searched against publicly available genomic databases; we note that there is no DfrB8 (Supplementary Materials Table S2) [11]. Since our objective was to analyze the genomic context of *dfrbs*, we retained only genomic segments that include at least 10kb both upstream and downstream from a *dfrb*. A total of 110 sequences were collected, representing 16 different bacterial species. In some cases, multiple similar sequences originated from the same BioProject; in these instances, redundant sequences were excluded from further analysis, keeping one representative sequence.

The taxonomic summary for the 61 remaining sequences is presented in Table 2. All sequences but one came from *Gammaproteobacteria* and included three different orders: *Aeromonadales*, *Enterobacterales* and *Pseudomonadales*. The predominant species was the clinically-relevant *Pseudomonas aeruginosa*, accounting for 36 sequences (59%). The only *Betaproteobacteria* sequence was from *Burkholderia dolosa*, isolated from a cystic fibrosis patient in the United States of America [37].

When available, information on the isolation source of each sample and the country of origin was compiled. Most *dfrb*-containing strains were identified in samples collected in Asia (39%), followed by Europe (20%), America (17%) and Africa (5%). The majority of strains (62%) were found within humans; an additional 7% found in wastewater and 2% in hospital wastewater may also be of human origin. Surprisingly, despite intensive use of TMP for livestock, only 3% of strains from our dataset were isolated from animals. This could indicate a sampling bias from the databases. Finally, 2% of samples were identified as environmental.

Four species in the dataset (*Acinetobacter baumannii*, *Enterobacter hormaechei*, *Klebsiella pneumoniae* and *Pseudomonas aeruginosa*) are categorized as ESKAPE pathogens, accounting for 67% of the sequences. Overall, 79% of the 61 *dfrb*-containing sequences contain virulence genes enabling them to cause infection according to the VFDB database. Because our ge-

nomie sequence dataset included partial sequences, the fraction of sequences of pathogenic bacteria can be underestimated if virulence genes are outside of the sequenced region.

Table 2. Taxonomic classification of all strains identified that include at least one *dfrb*.

Class/Order/Family/Genus	Strain Count ^a
Betaproteobacteria	1
<i>Burkholderiales</i>	1
<i>Burkholderiaceae</i>	1
<i>Burkholderia</i>	1
Gammaproteobacteria	60 (110)
<i>Aeromonadales</i>	2
<i>Aeromonadaceae</i>	2
<i>Aeromonas</i>	2
<i>Enterobacteriales</i>	16 (17)
<i>Enterobacteriaceae</i>	14 (15)
<i>Citrobacter</i>	1
<i>Enterobacter</i>	1
<i>Escherichia</i>	4
<i>Klebsiella</i>	4 (5)
<i>Salmonella</i>	4
<i>Morganellaceae</i>	1
<i>Providencia</i>	1
<i>Yersiniaceae</i>	1
<i>Serratia</i>	1
<i>Pseudomonadales</i>	42 (91)
<i>Moraxellaceae</i>	1
<i>Acinetobacter</i>	1
<i>Pseudomonadaceae</i>	41 (90)
<i>Pseudomonas</i>	41 (90)

^a Values include sequences used in the analysis after exclusion of redundancy. Values in parentheses include redundant sequences.

2.3. Analysis of the Genomic Context

We investigated the genomic context within which the *dfrbs* were found. In particular, the presence of MGEs can inform us of the capacity of the *dfrbs* to transfer to other genomes. Ever since the initial discovery of *dfrb1* (R67) and *dfrb2* (R388) on plasmids, DfrBs have been systematically referred to as being plasmid-borne [38,39]. This is consistent with the importance of plasmids in acquired bacterial resistance [40]. The genomic context was determined by classifying the *dfrb*-containing genomes as either plasmidic or chromosomal using the PlasFlow classification tool [41]. This resulted in 18 sequences being labelled “plasmids” (30%) and 43 sequences labelled “chromosomes” (70%) (Figure 1). The only sequence identified in *Betaproteobacteria* was chromosomal. Among the *Gammaproteobacteria*, all 16 sequences identified in *Enterobacteriales* were plasmidic, whereas the two sequences in *Aeromonadales* were chromosomal. The *Pseudomonadales* sequences were chromosomal except for two sequences: one from *P. aeruginosa* and one from *P. putida*. Only one plasmidic sequence, that from *P. rettgeri*, was labelled as pathogenic.

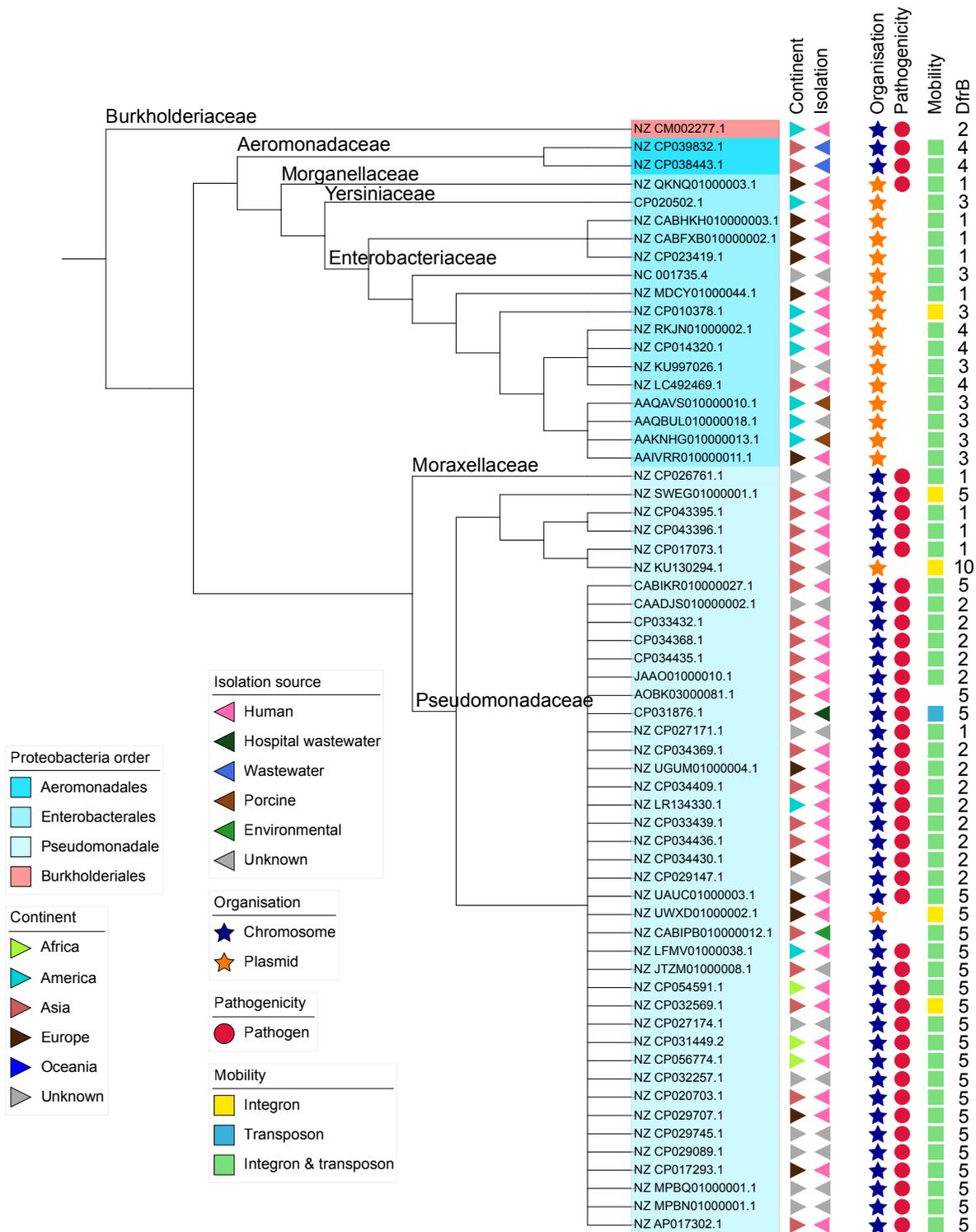


Figure 1. Annotated phylogenetic tree of species harboring a *dfrB*. Taxonomic classification of order and family is followed by categorization according to GenBank information on the strain’s isolation source and country of isolation. Sequences are further categorized as being located in a chromosome or a plasmid, pathogenicity of the host organism and information on mobile genetic elements. The *dfrB* gene member identified in each sequence is specified (i.e., “2” indicates *dfrB2*).

Next, we gained insight into the types of genes flanking the *dfrBs*. MGEs and other resistance genes near the *dfrBs* would define them as belonging to a multiresistance context. A blastx search was performed, using 20kb sequence segments containing *dfrBs* as queries, against a compiled antibiotic resistance gene database (see Materials and Methods),

keeping hits having at least 80% coverage and 60% identity. We first determined that the vast majority of *dfrbs* (89%) had both integrase and transposase within a 10kb window (Figure 1). Five sequences (NZ_CP010378.1, NZ_SWEG01000001.1, NZ_UWXD01000002.1, NZ_CP032569.1, NZ_KU130294.1) had only an integrase annotated nearby and one sequence (CP031876.1) had only a transposase annotated. Two sequences included neither; one (NZ_CM002277.1) was a chromosomal pathogenic sequence from *Burkholderia dolosa*, the only *Betaproteobacteria* identified. That sequence included no further genes related to antimicrobial resistance or genomic mobility within 10kb of its *dfrb* (Figure 2a). The other sequence including no integrase or transposase (AOBK03000081.1) was a chromosomal and pathogenic sequence from *P. aeruginosa*. It included only the rifampin-resistance gene *arr2* at a distance of 0.14kb from its *dfrb*.

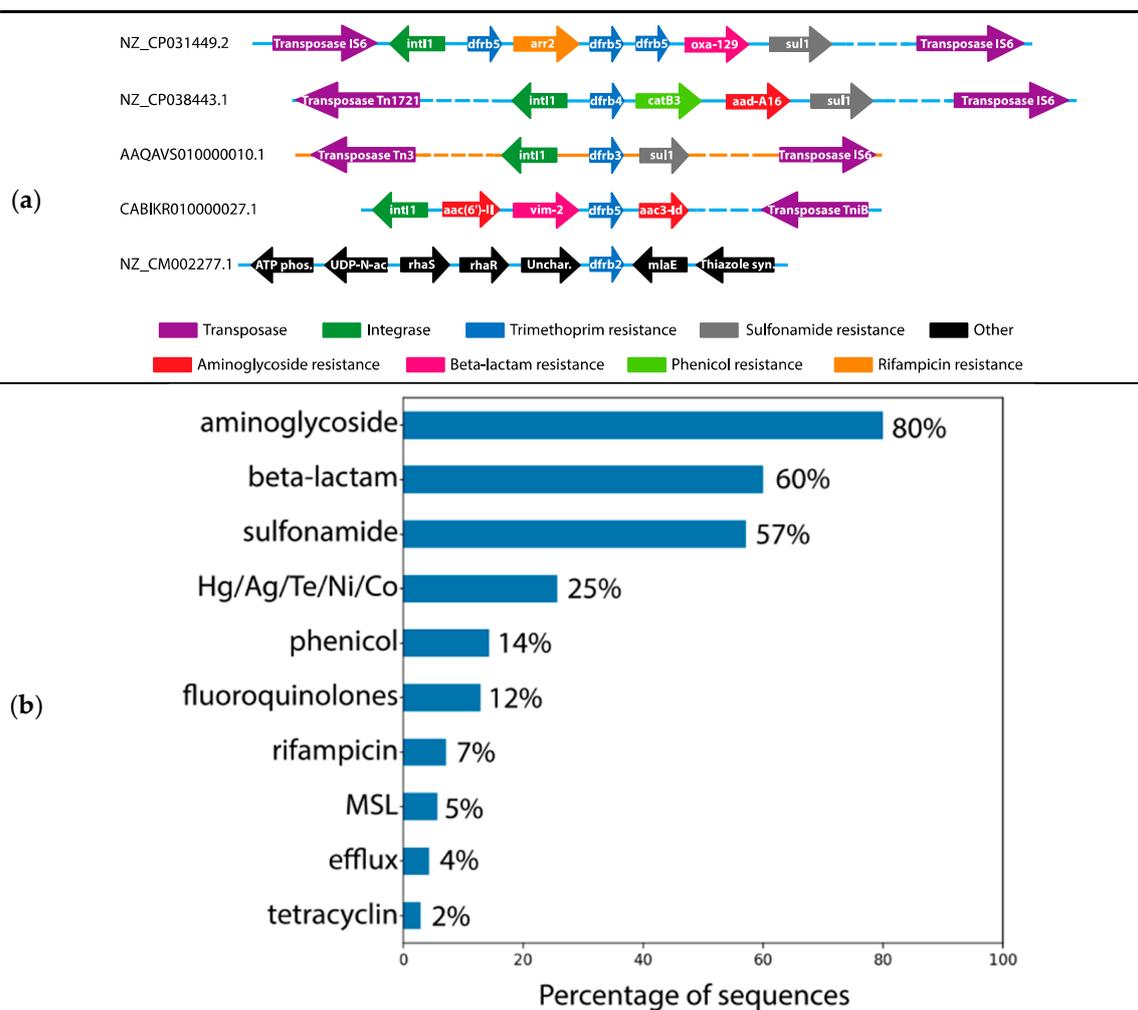


Figure 2. (a) Scheme of the genomic context of five representative sequences. Blue lines represent genomic sequences, whereas the orange line represents a plasmid sequence. Dashed lines contain regions that are not represented here. (b) Population of *dfrb* genes accompanied by a gene conferring resistance to another antimicrobial agent, expressed as percentage. MSL: macrolides, streptogramins, lincosamides.

The distance separating the *dfrbs* from the integrases and transposases was mapped (Figure 3). For *dfrb1*, *dfrb2* and *dfrb4*, we observed a large variability in those distances, suggesting diversity in their genomic context and thus diversity in the events of integration. In the cases of *dfrb3* and *dfrb5*, the same gene cassette was present in a few sequences, thus the same distance between elements was mapped. For example, all eight *dfrb3* genes were found in integrons directly upstream from the class 1 integrase, marking *dfrb3* as the gene

most recently integrated into the cassette. Interestingly, *dfrb3* was found only in plasmids, in *Enterobacteriales*. Six of these genes were found in the short *dfrb3/sul1* cassette. Similarly, multiple sequences held the same cassettes containing *dfrb5*. There were 15 sequences with the *aac(6′)-II/vim-2/dfrb5/aac(3)-Id* cassette and five with the *aac(6′)-II/dfrb5/aac(3)-Id* cassette; some other cassettes were found twice.

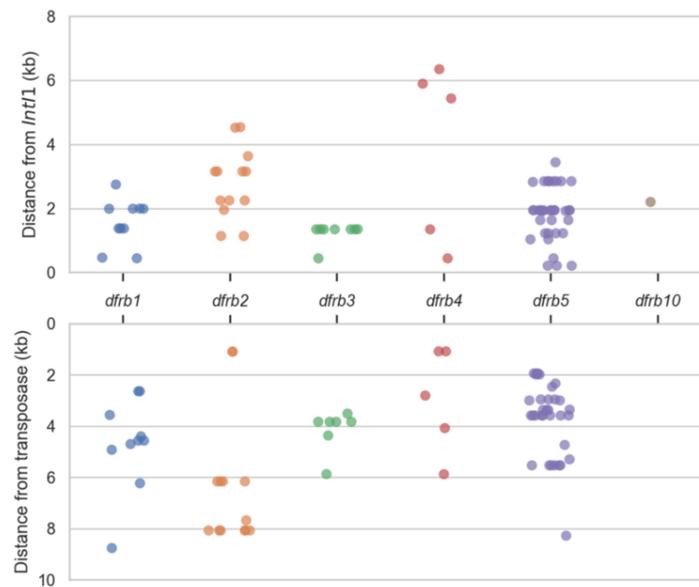


Figure 3. Distance between *dfrbs* and genes associated with genomic mobility. Top panel: distance between the *dfrbs* that are downstream a class 1 integrase. Bottom panel: distance between *dfrbs* and the closest transposase. Each dot represents one *dfrb* gene.

Finally, we mapped antimicrobial resistance genes annotated in a window 10kb on either side of the *dfrbs* and classified them according to the antimicrobial to which they conferred resistance (Figure 2a,b). The most prevalent phenotype was aminoglycoside resistance (80% of all sequences), followed by beta-lactamase genes (60%). Surprisingly, although TMP is often prescribed in combination with sulfonamide [42], only 57% of sequences included sulfonamide-resistance genes. Resistance to metals, phenicol, fluoroquinolones, rifampicin, macrolides, lincosamides, streptogramins, efflux genes and tetracyclin were observed at a lower frequency.

3. Discussion

The overwhelming majority of studies on clinical resistance to trimethoprim have focused exclusively on DfrAs. Recently, Sánchez-Osuna et al. reported two mechanisms of DfrA evolution [9]. One mechanism involves the mutation and mobilization of trimethoprim-sensitive *folA* genes (ex. *A. baumannii folA* mutated to *dfrA39* and *dfrA40*); the second relies on the mobilization of intrinsically trimethoprim-resistant *folA* genes. Trimethoprim resistance through DfrAs evolves readily, explaining the large number of DfrA genes: the recent addition of four members to the DfrA family has brought it to nearly 40 members [9].

Unlike DfrAs, the evolutionary origin of DfrBs has not been investigated. Little is known about these peculiar homotetrameric enzymes, including their prevalence and emergence in clinical and environmental samples [27]. We examine, for the first time, the genomic context of *dfrbs* by analyzing publicly available sequences containing *dfrbs*. By reporting the microorganisms that harbor them and determining whether they occur in the context of genetic mobility and/or resistance to multiple antibiotics, we provide insight into the risk they represent for public health.

Using a query-set consisting of the eight previously known *dfrb* genes and two newly identified and confirmed TMP-resistant *dfrb* genes, we identified 61 different genomic

sequences containing *dfrbs*. The country of origin of each sample illustrates that *dfrbs* are dispersed worldwide. The vast majority of *dfrbs* (74%) for which the source of isolation is known are related to human activities, whereas only one sequence comes from an environmental sample and 25% are from unknown sources. This observation could well reflect sampling biases due to overrepresentation of studies related to human activities relative to environmental studies in genomic databases. In Canada, TMP and sulfonamides are the fourth most highly prescribed antimicrobials for animals, representing 57,865 kg in 2018 [43], justifying the importance of increasing genomic analyses of animal samples to determine the prevalence of *dfrbs* in all relevant contexts.

Amongst the chromosomal sequences we identified, 93% contained virulence genes and included at least one MGE near the *dfrb* gene. These combined criteria define the highest risk that antimicrobial resistance genes can present [32]. The remaining sequences containing *dfrbs* included at least one of these two criteria. All plasmid-borne *dfrbs* were near an MGE, allowing them to spread easily among bacteria.

Since their discovery, DfrBs have been considered to be solely plasmid-borne [44,45]. Indeed, *dfrb1* was first observed in an *E. coli* strain where it is plasmid-borne, leading to the incorrect assumption that *dfrbs* are always plasmidic [38]. Nevertheless, we observed not only a few exceptions to this long-standing conjecture, but rather that only 29% of *dfrb*-containing sequences in our sample were plasmidic. The *Enterbacterales* species harbor *dfrbs* on plasmids, while *Aeromonadales* and *Pseudomonadales* species harbor *dfrbs* on their chromosome (Figure 1).

It is interesting to note that the only two *dfrb* genes that were not found near MGE were in chromosomal sequences (Figure 1). This suggests that *dfrbs* might have mobilized from a chromosome to a plasmid, and not have originated from plasmids. Further thorough examination of *dfrbs* and their mobilization context, both in plasmids and chromosomes, would allow the retracing of the early events of *dfrb* mobilization.

DfrB1 and DfrB2 were discovered in the 1970s, subsequent to the introduction of TMP. DfrB1, also named R67, dfrII and dfr2a, has been extensively characterized for its structure, assembly and catalytic activity [14,16,17,46–49]. Only recently have other members of the family been characterized [19] or even reported, such that it could have been thought that DfrB1 is the most widespread among DfrBs. However, *dfrb1* was identified in only 16% of the 61 sequences identified here. Unexpectedly, *dfrb5* was identified in 38% of the sequences, in various genomic contexts and geographical locations, suggesting that is more broadly disseminated than *dfrb1*. In addition, 23% of sequences contained *dfrb2*, 13% contained *dfrb3* while 8% contained *dfrb4*. The *dfrb10* gene, reported here for the first time, was identified in one sequence and none among the *dfrb6*, *dfrb7*, *dfrb9* or *dfrb11* genes were identified using our search criteria for genomic segments.

Although the sample sets included in the databases we searched represent only a fragmentary picture of gene dissemination, some members of the DfrB family are clearly more prevalent while others may not have yet emerged. Interestingly, the most prevalent *dfrb5* is closely related to the first-reported *dfrb1* (Table S1). Further investigation will be required to determine whether early events of *dfrb* mobilization in pathogenic bacteria involves either of these two genes. Close surveillance of the prevalence of these genes is needed to evaluate the spread of this family of genes.

All but one *dfrb* (NZ_CM002277.1) were found near at least one antibiotic resistance gene (ARG), the majority (83%) being in proximity (less than 10kb) to three other ARGs. Among these, the *sul1* sulfonamide-resistant gene is present in the vast majority of clinically relevant integrons [50]. We previously reported identification of *dfrb4* in a clinical class 1 integron within the *dfrb4/qacEΔ1/sul1* cassette, flanked by further resistance genes [27]. Since TMP is often prescribed in combination with sulfamethoxazole, it is noteworthy that only 57% of the *dfrbs* identified in integrons in this study were colocalized with *sul1*. The *dfrbs* were found in class 1 integrons as defined by the presence of a class 1 integrase. Multiple ARGs were observed within the same cassettes as the *dfrbs* (Figure 2a). Expression of some of these ARGs (e.g., β-lactam resistance *vim-2* and *oxa-10*, aminoglycoside

resistance *aadA1*, rifampin resistance *arr-2*) in other class 1 integrons has previously been reported [51–53]. Although this does not demonstrate gene expression in these genomic contexts, it is consistent with the hypothesis that the ARGs as well as the *dfrbs* are expressed in the class 1 integrons identified here.

No *dfras* were found in proximity to the *dfrbs*; this is expected since these genes procure the same phenotype. Nonetheless, duplication of *dfrb5* was observed in the integron of one genome (NZ_CP031449.2), where three copies of *dfrb5* were observed in the same integron (Figure 2a). In addition, duplication of similar integrons containing *dfrb5* in the same sequence was observed in one plasmid and four genomes, all from *P. aeruginosa*. Considering the incomplete nature of the sequences we analyzed, it is possible that a greater number of amplification events could be identified upon analysis of longer sequence segments.

Given the importance of TMP both in the clinic and with livestock, it is critical to monitor the emergence of resistance to this antimicrobial. TMP resistance has generally been associated with DfrAs. Here, we have demonstrated that monitoring the emergence and prevalence of DfrBs will provide important insights into global TMP resistance and thus contribute to policy making to contain the spread of antimicrobial resistance.

4. Materials and Methods

4.1. Identification of Putative Type B Dihydrofolate Reductases

The six DfrB sequences that were previously functionally characterized (DfrB1–DfrB5 and DfrB7) [19] were used to create a profile hidden Markov models (HMM) with HMMER version 3.3 (<http://hmmmer.org/> accessed on 10 January 2021). This profile was used as a query against the UniProtKB/TrEMBL database (22 April 2020 release, 184,998,855 sequences) [54]. Hits with E-value lower than 1×10^{-40} were considered and compared to known DfrB sequences. Predicted sequences having a protein sequence identity lower than 95% relative to any known DfrB sequence were considered as new genes.

4.2. Subcloning of *dfrb10* and *dfrb11*

The genetic sequences of *dfrb10* and *dfrb11* were obtained in pUC57 (BioBasic) according to the Genbank accession numbers in Table 1. The N-terminally His₆-tagged ORF sequences of *dfrb10* and *dfrb11* were subcloned into pET24 (Qiagen) downstream of the lactose operon repressor, following the lac operator sequence, using the *Nde*I and *Hind*III restriction sites. Both genes were amplified by PCR using the same forward primer 5'-GAAATAATTTTGTTTAACTTTAAGAAGGAGATATACATATGAGAGGATCTCACCATCAC-3' (*Nde*I site in bold) and a reverse primer that differs at one base (underlined) to maintain the native stop codon, *dfrb10*: 5'-GGTGGTGCTCGAGTGCGGCCGCAAGCTTTTAGGCCACGCG-3'; *dfrb11*: 5'-GGTGGTGCTCGAGTGCGGCCGCAAGCTTTCAGGCCACGCG-3' (*Hind*III site in bold). Phusion HF polymerase (ThermoFisher) was used according to the manufacturer's protocol, using 55 °C as the annealing temperature. Amplified genes, as well as pET24, were digested with *Hind*III (NEB) for 14 h and *Nde*I (NEB) for 2 h at 37 °C, followed by enzyme inactivation for 20 min at 80 °C. They were gel-extracted using the Monarch DNA gel extraction kit (NEB) and purified using the DNA Cleanup kit (NEB). Inserts were ligated into digested pET24 using a DNA ligation kit (Takara) according to the manufacturer's instructions. Briefly, the digested gene and pET24 vector were incubated at 16 °C for 3 h in Takara solution I. The ligation products were transformed into CaCl₂-competent *E. coli* DH5α prepared by the method of Inoue [55]. The DNA sequences of *dfrB10*-pET24 and *dfrB11*-pET24 were confirmed by DNA Sanger sequencing (Genome Quebec platform at Sainte-Justine Hospital). The final constructs yield N-terminally, His₆-tagged DfrB proteins. His₆-DfrB3 in pET24 was previously reported [19]. The negative control cTEM-19m, with an expressible β-lactamase insert instead of a *dfrb* insert, was previously described [56].

4.3. Minimal Inhibitory Concentration

MICs were determined according to Wiegand et al. [57] using the broth microdilution method. Briefly, *E. coli* BL21(DE3) cells expressing His₆-DfrB3 (positive control), His₆-DfrB10, His₆-DfrB11 or cTEM-19m (negative control) were propagated overnight in Luria-Bertani (LB) media with 50 µg/mL kanamycin. In 96-well plates, an inoculum of 10⁵ colony forming units (cfu) was inoculated in LB media, with 0.1 mM IPTG (ThermoFisher) and TMP (Sigma) in 2-fold concentration steps up to 600 µg/mL; the latter is the highest concentration of TMP soluble in 5% methanol. MICs were determined in triplicate.

4.4. Download of Genomes

The sequences of *dfrb1-dfrb7*, *dfrb9* and the newly identified *dfrb10* and *dfrb11* were used as queries for blastn 2.10.0 searches against four genomic databases (performed on the 2020.07.04): RefSeq (bacterial sequences), GenBank (bacterial sequences) and the Microbial Complete Genomes and Complete Plasmids databases found at <https://blast.ncbi.nlm.nih.gov/Blast.cgi> (accessed on 10 January 2021) [58–60]. Genomes containing at least one query sequence and having a sequence length of at least 10kb both upstream and downstream of the *dfrb* sequences were collated. In total, 110 sequences were identified and served for analysis.

4.5. Protein Database Constructions

The following protein databases were downloaded on 2020.07.07: Integrase, IntI1 and sul1 databases from the I-VIP pipeline [61], ARG-ANNOT [25], ICEberg 2.0 [62], CARD [23], BacMet [63] and UniProtKB/Swiss-Prot [64]. These databases were merged and redundant sequences were removed. Two genes, coding for 78 and 97 amino-acid products, were both named DfrB1; the shorter gene version was kept to match the consensus length of all other members of the family.

4.6. Annotation

The 110 sequences of 20 kb (10 kb upstream and downstream of a *dfrb*) were used as query sequences against the blast database for a blastx 2.10.0 search with the parameters of E-value lower than 1×10^{-10} and culling_limit of 1. Hits with coverage of $\geq 80\%$ and protein identity of $\geq 60\%$ were kept. Where multiple sequences from a same NCBI BioProject presented the same annotation, all but one were removed from the dataset. In total, 61 sequences served for analysis.

4.7. Classification of Sequences as Chromosomal or Plasmidic

The 110 sequences were classified as chromosomal or plasmidic using PlasFlow 1.1 with a threshold of 0.65 [41].

4.8. Identification of Pathogenic Hosts

For each sequence, a blastx 2.10.0 analysis was carried against virulence factor protein sequences from the core dataset of VFDB (last update on 17 July 2020) [65] using an e-value cutoff of 1×10^{-15} . Hits were filtered using an identity and coverage threshold of 60%. Sequences with one or more hit were labelled pathogenic.

4.9. Phylogenetic Tree

The phylogenetic tree was constructed with NGPhylogeny.fr [66] using the host species' 16S rRNA from the NCBI reference genome. The tools MUSCLE (with the Neighbor joining option), Noisy, PHyML + SMS, and Newick were used for tree construction. The tree was annotated with iTOL, where the GenBank accession of each sequence is displayed [67].

Supplementary Materials: The following are available online at <https://www.mdpi.com/article/10.3390/antibiotics10040433/s1>, Table S1: Protein sequence identity of the ten members of the DfrB family, Table S2: DfrB sequences.

Author Contributions: C.L.-S.-D. conceptualization, methodology, data analysis, writing—original draft preparation; S.-S.D. data analysis; J.N.P. conceptualization, writing—review and editing, funding acquisition. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by NSERC grant number RGPIN-2018-04686. C.L.S.D. and S.S.D. are grateful to NSERC, FQRNT and Université de Montréal for scholarships.

Data Availability Statement: Data is contained within the article or supplementary material. The accession numbers of analyzed sequences listed in Figure 1 are accessible in publicly available databases.

Acknowledgments: We thank Janine N. Copp for her insights and fruitful discussion, as well as Lorea Alejaldre and Étienne Lavallée for proofreading. We also thank William Huynh and Andrew McArthur at CARD and Daniel Haft at NCBI for their help navigating various databases.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. World Health Organization. *Critically Important Antimicrobials for Human Medicine*, 6th ed.; World Health Organization: Geneva, Switzerland, 2018.
2. Stone, S.R.; Morrison, J.F. Mechanism of Inhibition of Dihydrofolate Reductases from Bacterial and Vertebrate Sources by Various Classes of Folate Analogues. *Biochim. Biophys. Acta BBA Protein Struct. Mol. Enzymol.* **1986**, *869*, 275–285. [[CrossRef](#)]
3. Eliopoulos, G.M.; Huovinen, P. Resistance to Trimethoprim-Sulfamethoxazole. *Clin. Infect. Dis.* **2001**, *32*, 1608–1614. [[CrossRef](#)]
4. Caron, F.; Wehrle, V.; Etienne, M. The Comeback of Trimethoprim in France. *Med. Mal. Infect.* **2017**, *47*, 253–260. [[CrossRef](#)]
5. World Health Organization. *WHO Report on Surveillance of Antibiotic Consumption: 2016–2018 Early Implementation*; World Health Organization: Geneva, Switzerland, 2018.
6. Hsia, Y.; Lee, B.R.; Versporten, A.; Yang, Y.; Bielicki, J.; Jackson, C.; Newland, J.; Goossens, H.; Magrini, N.; Sharland, M.; et al. Use of the WHO Access, Watch, and Reserve Classification to Define Patterns of Hospital Antibiotic Use (AWaRe): An Analysis of Paediatric Survey Data from 56 Countries. *Lancet Glob. Health* **2019**, *7*, e861–e871. [[CrossRef](#)]
7. Cuong, N.V.; Padungtod, P.; Thwaites, G.; Carrique-Mas, J.J. Antimicrobial Usage in Animal Production: A Review of the Literature with a Focus on Low- and Middle-Income Countries. *Antibiotics* **2018**, *7*, 75. [[CrossRef](#)]
8. World Health Organization Department of Food Safety and Zoonoses; World Health Organization. *WHO Guidelines on Use of Medically Important Antimicrobials in Food-Producing Animals*; World Health Organization: Geneva, Switzerland, 2017; ISBN 978-92-4-155013-0.
9. Sánchez-Osuna, M.; Cortés, P.; Llagostera, M.; Barbé, J.; Erill, I. Exploration into the Origins and Mobilization of Di-Hydrofolate Reductase Genes and the Emergence of Clinical Resistance to Trimethoprim. *Microb. Genom.* **2020**, *6*. [[CrossRef](#)]
10. Howell, E.E. Searching Sequence Space: Two Different Approaches to Dihydrofolate Reductase Catalysis. *ChemBioChem* **2005**, *6*, 590–600. [[CrossRef](#)] [[PubMed](#)]
11. Faltyn, M.; Alcock, B.; McArthur, A. Evolution and Nomenclature of the Trimethoprim Resistant Dihydrofolate (Dfr) Reductases. *Preprints* **2019**. [[CrossRef](#)]
12. Kim, Y.B. Improved Trimethoprim-Resistance Cassette for Prokaryotic Selections. *J. Biosci. Bioeng.* **2009**, *108*, 441–445. [[CrossRef](#)]
13. Pattishall, K.H.; Acar, J.; Burchall, J.J.; Goldstein, F.W.; Harvey, R.J. Two Distinct Types of Trimethoprim-Resistant Dihydrofolate Reductase Specified by R-Plasmids of Different Compatibility Groups. *J. Biol. Chem.* **1977**, *252*, 2319–2323. [[CrossRef](#)]
14. Krahn, J.M.; Jackson, M.R.; DeRose, E.F.; Howell, E.E.; London, R.E. Crystal Structure of a Type II Dihydrofolate Reductase Catalytic Ternary Complex [†]. *Biochemistry* **2007**, *46*, 14878–14888. [[CrossRef](#)] [[PubMed](#)]
15. Bhojane, P.P.; Duff, M.R.; Bafna, K.; Agarwal, P.; Stanley, C.; Howell, E.E. Small Angle Neutron Scattering Studies of R67 Dihydrofolate Reductase, a Tetrameric Protein with Intrinsically Disordered N-Termini. *Biochemistry* **2017**, *56*, 5886–5899. [[CrossRef](#)] [[PubMed](#)]
16. Ebert, M.C.; Morley, K.L.; Volpato, J.P.; Schmitzer, A.R.; Pelletier, J.N. Asymmetric Mutations in the Tetrameric R67 Dihydrofolate Reductase Reveal High Tolerance to Active-Site Substitutions: Asymmetric Mutations in R67 Dihydrofolate Reductase. *Protein Sci.* **2015**, *24*, 495–507. [[CrossRef](#)] [[PubMed](#)]
17. Schmitzer, A.R.; Lépine, F.; Pelletier, J.N. Combinatorial Exploration of the Catalytic Site of a Drug-Resistant Dihydrofolate Reductase: Creating Alternative Functional Configurations. *Protein Eng. Des. Sel.* **2004**, *17*, 809–819. [[CrossRef](#)]
18. Martinez, M.A.; Pezo, V.; Marlière, P.; Wain-Hobson, S. Exploring the Functional Robustness of an Enzyme by In Vitro Evolution. *EMBO J.* **1996**, *15*, 1203–1210. [[CrossRef](#)] [[PubMed](#)]
19. Toulouse, J.L.; Shi, G.; Lemay-St-Denis, C.; Ebert, M.C.; Deon, D.; Gagnon, M.; Ruediger, E.; Saint-Jacques, K.; Forge, D.; Vanden Eynde, J.J.; et al. Dual-Target Inhibitors of the Folate Pathway Inhibit Intrinsically Trimethoprim-Resistant DfrB Dihydrofolate Reductases. *ACS Med. Chem. Lett.* **2020**. [[CrossRef](#)] [[PubMed](#)]
20. Toulouse, J.L.; Yachnin, B.J.; Ruediger, E.H.; Deon, D.; Gagnon, M.; Saint-Jacques, K.; Ebert, M.C.; Forge, D.; Bastien, D.; Colin, D.Y.; et al. Structure-Based Design of Dimeric Bisbenzimidazole Inhibitors to an Emergent Trimethoprim-Resistant Type II Dihydrofolate Reductase Guides the Design of Monomeric Analogues. *ACS Omega* **2019**, *4*, 10056–10069. [[CrossRef](#)]

21. Toulouse, J.L.; Abraham, S.M.J.; Kadnikova, N.; Bastien, D.; Gauchot, V.; Schmitzer, A.R.; Pelletier, J.N. Investigation of Classical Organic and Ionic Liquid Cosolvents for Early-Stage Screening in Fragment-Based Inhibitor Design with Unrelated Bacterial and Human Dihydrofolate Reductases. *Assay Drug Dev. Technol.* **2017**, *15*, 141–153. [[CrossRef](#)]
22. Bastien, D.; Ebert, M.C.; Forge, D.; Toulouse, J.; Kadnikova, N.; Perron, F.; Mayence, A.; Huang, T.L.; Vanden Eynde, J.J.; Pelletier, J.N. Fragment-Based Design of Symmetrical Bis-Benzimidazoles as Selective Inhibitors of the Trimethoprim-Resistant, Type II R67 Dihydrofolate Reductase. *J. Med. Chem.* **2012**, *55*, 3182–3192. [[CrossRef](#)]
23. Alcock, B.P.; Raphenya, A.R.; Lau, T.T.Y.; Tsang, K.K.; Boucharde, M.; Edalatmand, A.; Huynh, W.; Nguyen, A.-L.V.; Cheng, A.A.; Liu, S.; et al. CARD 2020: Antibiotic Resistance Surveillance with the Comprehensive Antibiotic Resistance Database. *Nucleic Acids Res.* **2019**, gkz935. [[CrossRef](#)]
24. Liu, B.; Pop, M. ARDB—Antibiotic Resistance Genes Database. *Nucleic Acids Res.* **2009**, *37*, D443–D447. [[CrossRef](#)]
25. Gupta, S.K.; Padmanabhan, B.R.; Diene, S.M.; Lopez-Rojas, R.; Kempf, M.; Landraud, L.; Rolain, J.-M. ARG-ANNOT, a New Bioinformatic Tool to Discover Antibiotic Resistance Genes in Bacterial Genomes. *Antimicrob. Agents Chemother.* **2014**, *58*, 212–220. [[CrossRef](#)]
26. Grape, M.; Farra, A.; Kronvall, G.; Sundström, L. Integrons and Gene Cassettes in Clinical Isolates of Co-Trimoxazole-Resistant Gram-Negative Bacteria. *Clin. Microbiol. Infect.* **2005**, *11*, 185–192. [[CrossRef](#)] [[PubMed](#)]
27. Toulouse, J.L.; Edens, T.J.; Alejandre, L.; Manges, A.R.; Pelletier, J.N. Integron-Associated DfrB4, a Previously Uncharacterized Member of the Trimethoprim-Resistant Dihydrofolate Reductase B Family, Is a Clinically Identified Emergent Source of Antibiotic Resistance. *Antimicrob. Agents Chemother.* **2017**, *61*, e02665-16. [[CrossRef](#)]
28. Dinger, M.E.; Pang, K.C.; Mercer, T.R.; Mattick, J.S. Differentiating Protein-Coding and Noncoding RNA: Challenges and Ambiguities. *PLoS Comput. Biol.* **2008**, *4*, e1000176. [[CrossRef](#)]
29. Hyatt, D.; Chen, G.-L.; LoCascio, P.F.; Land, M.L.; Larimer, F.W.; Hauser, L.J. Prodigal: Prokaryotic Gene Recognition and Translation Initiation Site Identification. *BMC Bioinform.* **2010**, *11*, 119. [[CrossRef](#)] [[PubMed](#)]
30. Kadlec, K.; Kehrenberg, C.; Schwarz, S. Molecular Basis of Resistance to Trimethoprim, Chloramphenicol and Sulphonamides in *Bordetella Bronchiseptica*. *J. Antimicrob. Chemother.* **2005**, *56*, 485–490. [[CrossRef](#)] [[PubMed](#)]
31. Grape, M.; Motakefi, A.; Pavuluri, S.; Kahlmeter, G. Standard and Real-Time Multiplex PCR Methods for Detection of Trimethoprim Resistance Dfr Genes in Large Collections of Bacteria. *Clin. Microbiol. Infect.* **2007**, *13*, 1112–1118. [[CrossRef](#)]
32. Martínez, J.L.; Coque, T.M.; Baquero, F. What Is a Resistance Gene? Ranking Risk in Resistomes. *Nat. Rev. Microbiol.* **2015**, *13*, 116–123. [[CrossRef](#)]
33. Jaillard, M.; van Belkum, A.; Cady, K.C.; Creely, D.; Shortridge, D.; Blanc, B.; Barbu, E.M.; Dunne, W.M.; Zambardi, G.; Enright, M.; et al. Correlation between Phenotypic Antibiotic Susceptibility and the Resistome in *Pseudomonas Aeruginosa*. *Int. J. Antimicrob. Agents* **2017**, *50*, 210–218. [[CrossRef](#)]
34. Azam, M.W.; Khan, A.U. Updates on the Pathogenicity Status of *Pseudomonas Aeruginosa*. *Drug Discov. Today* **2019**, *24*, 350–359. [[CrossRef](#)]
35. Sun, F.; Zhou, D.; Wang, Q.; Feng, J.; Feng, W.; Luo, W.; Liu, Y.; Qiu, X.; Yin, Z.; Xia, P. Genetic Characterization of a Novel *bla*_{DIM-2}-Carrying Megaplasmid P12969-DIM from Clinical *Pseudomonas Putida*. *J. Antimicrob. Chemother.* **2016**, *71*, 909–912. [[CrossRef](#)] [[PubMed](#)]
36. Hensdorf, A.W.; Amano, Y.; Miyakawa, K.; Ise, K.; Suzuki, Y.; Anantharaman, K.; Probst, A.; Burstein, D.; Thomas, B.C.; Banfield, J.F. Potential for Microbial H2 and Metal Transformations Associated with Novel Bacteria and Archaea in Deep Terrestrial Subsurface Sediments. *ISME J.* **2017**, *11*, 1915–1929. [[CrossRef](#)] [[PubMed](#)]
37. Workentine, M.L.; Surette, M.G.; Bernier, S.P. Draft Genome Sequence of Burkholderia Dolosa PC543 Isolated from Cystic Fibrosis Airways. *Genome Announc.* **2014**, *2*. [[CrossRef](#)]
38. Stone, D.; Smith, S.L. The Amino Acid Sequence of the Trimethoprim-Resistant Dihydrofolate Reductase Specified in *Escherichia Coli* by R-Plasmid R67. *J. Biol. Chem.* **1979**, *254*, 10857–10861. [[CrossRef](#)]
39. Amyes, S.G.B.; Smith, J.T. R-Factor Trimethoprim Resistance Mechanism: An Insusceptible Target Site. *Biochem. Biophys. Res. Commun.* **1974**, *58*, 412–418. [[CrossRef](#)]
40. San Millan, A. Evolution of Plasmid-Mediated Antibiotic Resistance in the Clinical Context. *Trends Microbiol.* **2018**, *26*, 978–985. [[CrossRef](#)] [[PubMed](#)]
41. Krawczyk, P.S.; Lipinski, L.; Dziembowski, A. PlasFlow: Predicting Plasmid Sequences in Metagenomic Data Using Genome Signatures. *Nucleic Acids Res.* **2018**, *46*, e35. [[CrossRef](#)]
42. Masters, P.A.; O'Bryan, T.A.; Zurlo, J.; Miller, D.Q.; Joshi, N. Trimethoprim-Sulfamethoxazole Revisited. *Arch. Intern. Med.* **2003**, *163*, 402–410. [[CrossRef](#)] [[PubMed](#)]
43. Public Health Agency of Canada. *Canadian Antimicrobial Resistance Surveillance System Report*; Public Health Agency of Canada: Ottawa, ON, Canada, 2020.
44. Swift, G.; McCarthy, B.J.; Heffron, F. DNA Sequence of a Plasmid-Encoded Dihydrofolate Reductase. *Mol. Gen. Genet. MGG* **1981**, *181*, 441–447. [[CrossRef](#)] [[PubMed](#)]
45. Brisson, N.; Hohn, T. Nucleotide Sequence of the Dihydrofolate-Reductase Gene Borne by the Plasmid R67 and Conferring Methotrexate Resistance. *Gene* **1984**, *28*, 271–274. [[CrossRef](#)]
46. Narayana, N.; Matthews, D.A.; Howell, E.E.; Xuong, N. A Plasmid-Encoded Dihydrofolate Reductase from Trimethoprim-Resistant Bacteria Has a Novel D2-Symmetric Active Site. *Nat. Struct. Mol. Biol.* **1995**, *2*, 1018–1025. [[CrossRef](#)]

47. Park, H.; Zhuang, P.; Nichols, R.; Howell, E.E. Mechanistic Studies of R67 Dihydrofolate Reductase. *J. Biol. Chem.* **1997**, *272*, 2252–2258. [[CrossRef](#)]
48. West, F.W.; Seo, H.-S.; Bradrick, T.D.; Howell, E.E. Effects of Single-Tryptophan Mutations on R67 Dihydrofolate Reductase †. *Biochemistry* **2000**, *39*, 3678–3689. [[CrossRef](#)]
49. Kamath, G.; Howell, E.E.; Agarwal, P.K. The Tail Wagging the Dog: Insights into Catalysis in R67 Dihydrofolate Reductase. *Biochemistry* **2010**, *49*, 9078–9088. [[CrossRef](#)]
50. Gillings, M.R. Class 1 Integrons as Invasive Species. *Curr. Opin. Microbiol.* **2017**, *38*, 10–15. [[CrossRef](#)]
51. Santos, C.; Caetano, T.; Ferreira, S.; Mendo, S. Tn5090-like Class 1 Integron Carrying BlaVIM-2 in a Pseudomonas Putida Strain from Portugal. *Clin. Microbiol. Infect.* **2010**, *16*, 1558–1561. [[CrossRef](#)]
52. Chiu, C.-H.; Lee, H.-Y.; Tseng, L.-Y.; Chen, C.-L.; Chia, J.-H.; Su, L.-H.; Liu, S.-Y. Mechanisms of Resistance to Ciprofloxacin, Ampicillin/Sulbactam and Imipenem in Acinetobacter Baumannii Clinical Isolates in Taiwan. *Int. J. Antimicrob. Agents* **2010**, *35*, 382–386. [[CrossRef](#)] [[PubMed](#)]
53. Naas, T.; Mikami, Y.; Imai, T.; Poirel, L.; Nordmann, P. Characterization of In53, a Class 1 Plasmid- and Composite Transposon-located Integron of Escherichia Coli Which Carries an Unusual Array of Gene Cassettes. *J. Bacteriol.* **2001**, *183*, 235–249. [[CrossRef](#)]
54. Boeckmann, B.; Bairoch, A.; Apweiler, R.; Blatter, M.-C.; Estreicher, A.; Gasteiger, E.; Martin, M.J.; Michoud, K.; O'Donovan, C.; Phan, I.; et al. The SWISS-PROT Protein Knowledgebase and Its Supplement TrEMBL in 2003. *Nucleic Acids Res.* **2003**, *31*, 365–370. [[CrossRef](#)]
55. Sambrook, J.; Russell, D.W. The Inoue Method for Preparation and Transformation of Competent E. Coli: "Ultra-Competent" Cells. *Cold Spring Harb. Protoc.* **2006**, *2006*, pdb.prot3944. [[CrossRef](#)]
56. Gobeil, S.M.C.; Gagné, D.; Doucet, N.; Pelletier, J.N. 15N, 13C and 1H Backbone Resonance Assignments of an Artificially Engineered TEM-1/PSE-4 Class A β -Lactamase Chimera and Its Deconvoluted Mutant. *Biomol. NMR Assign.* **2016**, *10*, 93–99. [[CrossRef](#)]
57. Wiegand, I.; Hilpert, K.; Hancock, R.E.W. Agar and Broth Dilution Methods to Determine the Minimal Inhibitory Concentration (MIC) of Antimicrobial Substances. *Nat. Protoc.* **2008**, *3*, 163–175. [[CrossRef](#)]
58. Sayers, E.W.; Agarwala, R.; Bolton, E.E.; Brister, J.R.; Canese, K.; Clark, K.; Connor, R.; Fiorini, N.; Funk, K.; Hefferon, T.; et al. Database Resources of the National Center for Biotechnology Information. *Nucleic Acids Res.* **2019**, *47*, D23–D28. [[CrossRef](#)] [[PubMed](#)]
59. Pruitt, K.D.; Tatusova, T.; Maglott, D.R. NCBI Reference Sequence (RefSeq): A Curated Non-Redundant Sequence Database of Genomes, Transcripts and Proteins. *Nucleic Acids Res.* **2005**, *33*, D501–D504. [[CrossRef](#)] [[PubMed](#)]
60. Benson, D.A.; Cavanaugh, M.; Clark, K.; Karsch-Mizrachi, I.; Lipman, D.J.; Ostell, J.; Sayers, E.W. GenBank. *Nucleic Acids Res.* **2017**, *45*, D37–D42. [[CrossRef](#)]
61. Zhang, A.N.; Li, L.-G.; Ma, L.; Gillings, M.R.; Tiedje, J.M.; Zhang, T. Conserved Phylogenetic Distribution and Limited Antibiotic Resistance of Class 1 Integrons Revealed by Assessing the Bacterial Genome and Plasmid Collection. *Microbiome* **2018**, *6*, 130. [[CrossRef](#)] [[PubMed](#)]
62. Liu, M.; Li, X.; Xie, Y.; Bi, D.; Sun, J.; Li, J.; Tai, C.; Deng, Z.; Ou, H.-Y. ICEberg 2.0: An Updated Database of Bacterial Integrative and Conjugative Elements. *Nucleic Acids Res.* **2019**, *47*, D660–D665. [[CrossRef](#)] [[PubMed](#)]
63. Pal, C.; Bengtsson-Palme, J.; Rensing, C.; Kristiansson, E.; Larsson, D.G.J. BacMet: Antibacterial Biocide and Metal Resistance Genes Database. *Nucleic Acids Res.* **2014**, *42*, D737–D743. [[CrossRef](#)] [[PubMed](#)]
64. Bairoch, A.; Boeckmann, B. The SWISS-PROT Protein Sequence Data Bank. *Nucleic Acids Res.* **1991**, *19*, 2247–2249. [[CrossRef](#)]
65. Liu, B.; Zheng, D.; Jin, Q.; Chen, L.; Yang, J. VFDB 2019: A Comparative Pathogenomic Platform with an Interactive Web Interface. *Nucleic Acids Res.* **2019**, *47*, D687–D692. [[CrossRef](#)] [[PubMed](#)]
66. Lemoine, F.; Correia, D.; Lefort, V.; Doppelt-Azeroual, O.; Mareuil, F.; Cohen-Boulakia, S.; Gascuel, O. NGPhylogeny.Fr: New Generation Phylogenetic Services for Non-Specialists. *Nucleic Acids Res.* **2019**, *47*, W260–W265. [[CrossRef](#)] [[PubMed](#)]
67. Letunic, I.; Bork, P. Interactive Tree of Life (ITOL) v3: An Online Tool for the Display and Annotation of Phylogenetic and Other Trees. *Nucleic Acids Res.* **2016**, *44*, W242–W245. [[CrossRef](#)] [[PubMed](#)]