RESEARCH ARTICLE

Primer set 2.0 for highly parallel qPCR array targeting antibiotic resistance genes and mobile genetic elements

Robert D. Stedtfeld[1,†], Xueping Guo[2,3,4,†], Tiffany M. Stedtfeld[1], Hongjie Sheng[3,4,6], Maggie R. Williams[1], Kristin Hauschild[4], Santosh Gunturu[4], Leo Tift[4], Fang Wang[3,4,6], Adina Howe[5], Benli Chai[4], Daqiang Yin[2], James R. Cole[3,4], James M. Tiedje[3,4] and Syed A. Hashsham[1,3,4,*]

[1]Department of Civil and Environmental Engineering, Michigan State University, East Lansing, Michigan 48824, USA, [2]State Key Laboratory of Pollution Control and Resources Reuse, College of Environmental Science and Engineering, Tongji University, Shanghai, China, [3]Department of Plant, Soil and Microbial Sciences, Michigan State University, East Lansing, Michigan 48824, USA, [4]Center for Microbial Ecology, Michigan State University, East Lansing, Michigan 48824, USA, [5]Department of Agricultural and Biosystems Engineering, Iowa State University, Ames, Iowa 50010, USA and [6]Key Laboratory of Soil Environment and Pollution Remediation, Institute of Soil Science, Chinese Academy of Sciences, Nanjing 210008, China

*Corresponding author: Department of Civil and Environmental Engineering, Center for Microbial Ecology, Michigan State University, 1449 Engineering Research Court, East Lansing, MI 48824, USA. E-mail: hashsham@egr.msu.edu

One sentence summary: Updated qPCR array targeting antibiotic resistance and mobile genetic elements.

†Equal contributions.

Editor: Tong Zhang

ABSTRACT

The high-throughput antibiotic resistance gene (ARG) qPCR array, initially published in 2012, is increasingly used to quantify resistance and mobile determinants in environmental matrices. Continued utility of the array; however, necessitates improvements such as removing or redesigning questionable primer sets, updating targeted genes and coverage of available sequences. Towards this goal, a new primer design tool (EcoFunPrimer) was used to aid in identification of conserved regions of diverse genes. The total number of assays used for diverse genes was reduced from 91 old primer sets to 52 new primer sets, with only a 10% loss in sequence coverage. While the old and new array both contain 384 primer sets, a reduction in old primer sets permitted 147 additional ARGs and mobile genetic elements to be targeted. Results of validating the updated array with a mock community of strains resulted in over 98% of tested instances incurring true positive/negative calls. Common queries related to sensitivity, quantification and conventional data analysis (e.g. Ct cutoff value, and estimated genomic copies without standard curves) were also explored. A combined list of new and previously used primer sets is provided with a recommended set based on redesign of primer sets and results of validation.

INTRODUCTION

Antibiotic resistance is considered an emerging pollutant due to the threat of acquired resistance in human and animal pathogens (Alanis 2005). Antibiotic resistance genes (ARGs) have been found in numerous environments such as water, soil, manure and air (Zhang *et al.* 2009; Hu *et al.* 2016; Pal *et al.* 2016; Xu *et al.* 2016), necessitating comprehensive tools that can be used to quantify dissemination and provide ecological risk assessment. One increasingly used tool is the high-throughput qPCR array, which was originally published over 5 years ago for simultaneous detection of hundreds of mobile genetic elements (MGEs) and ARGs (Looft *et al.*, 2012). Combined with commercially available tools such as the Takara (previously Wafergen) SmartChip, which can amplify up to 5184 qPCR assays per chip within 3–4 h, up to 384 primer sets can be analyzed in parallel.

Highly parallel qPCR studies targeting ARGs have typically used 296 or 384 primer set formats (Zhu *et al.* 2013; Wang et al., 2014, 2016; Karkman *et al.* 2016; Muziasari *et al.* 2016; Muurinen *et al.* 2017; Stedtfeld et al. 2017a, 2017b) with minimal changes to original assays (termed ARG array 1.0). However, multiple ARG types (e.g. NDM, mcr-1 genes) have since been discovered (Kumarasamy *et al.* 2010; Liu *et al.* 2016), as have new sequences for previously targeted genes in various environments and host-associated conditions (Forsberg *et al.* 2012; Hu *et al.* 2013). Information regarding mobile potential of ARGs (e.g. horizontal gene transfer) is also increasingly available (Courvalin 2008; Martinez, Coque and Baquero 2015; Hu *et al.* 2016).

Numerous studies using the ARG qPCR array have also provided insight into questionable primer sets in terms of specificity; as hundreds of samples from environmental (Stedtfeld *et al.* 2016; Zhu *et al.* 2017), wastewater (Karkman *et al.* 2016), and fecal matrices (Do *et al.* 2018; Qian *et al.* 2018) have now been analyzed using the qPCR ARG array. Primers that rarely or often amplify are thought to provide limited information in quantifying selective pressure in environmental samples. With questionable utility and specificity, primer sets that rarely or often amplify were redesigned or removed.

To make additional room for genes that were previously not targeted on the old array, genes requiring multiple primer sets for coverage were redesigned using a novel tool developed by the Ribosomal Database Project (RDP) at Michigan State University (https://github.com/rdpstaff/EcoFunPrimer), which reduces subjective bias of design caused by manual searches for conserved regions in diverse genes. In total, advancements of the ARG qPCR array included: (i) reducing number of primer sets required for high coverage of divergent genes, (ii) redesign or removal of questionable primer sets, and (iii) targeting additional ARGs and MGEs with an emphasis on mobility. Results include criteria used to select genes for primer redesign or removal, validation, and a comprehensive list of old and new primer sets with recommended assays for the updated ARG qPCR array (referred as ARG qPCR Array 2.0).

MATERIALS AND METHODS

Reference sequences collection and primer design

Selection of primer sets and gene targets that were either redesigned or removed from the array are described in more detail below. Removal of primer sets provided space for primer sets for previously untargeted genes. Selection of previously untargeted genes on the array was based on the analysis of ARGs observed on mobile elements via whole genome analysis of 23,435 bacterial genomes as described (Hu *et al.* 2016), or genes listed as experimentally confirmed plasmid-borne genes in the Antibacterial and Biocide and Metal Resistance Genes Database (Pal *et al.*, 2014). Reference sequences for selected target genes were assembled using the RDP FunGene Pipeline (http://fungene.cme.msu.edu/), which can be used to download, align, and trim sequences for a given gene. For genes not currently listed in FunGene, reference sequences from the latest version of ARG-ANNOT AA V3 March 2017 (Antibiotic Resistance Gene-ANNOTation) were used to gather additional sequences of high similarity via NCBI.

Following sequence collection, primer sets were designed for individual gene targets using the RDP EcoFunPrimer design tool (https://github.com/rdpstaff/EcoFunPrimer). The tool highlights conserved regions of a gene with thermodynamically stable primer pairs for qPCR, and automatically suggests primers with or without degeneracy as specified by the user. In addition, the tool evaluates conserved genomic regions, which can be used for manual selection of forward and reverse positions with desired amplicon length and coverage. For primer sets in ARG qPCR array 2.0, the following specifications were used with RDP EcoFunPrimer: sample select command with sliding scale, a theoretical melting temperature of 60°C, and zero degeneracy. A weighting system was implemented to select primers that cover sequences which improve the diversity of target genes.

Validation of the updated ARG qPCR array

Specificity, sensitivity and amplification efficiency of primer sets were tested using a dilution series of gDNA extracted from multiple organisms (28 strains) obtained from ATCC (Table S1, Supporting Information). Two SmartChip validation experiments were run, in that the first chip was validated with a mixture of type strains that targeted 17 old and 35 new primer sets, and the second chip had a mixture of strains that targeted 55 old and 68 new primer sets. Validation of the first plate was performed as an initial test of primer sets generated using EcoFunPrimer and contained only a subset of new primer sets. The second plate was a more complete update of the ARG qPCR array, which contained a subset of the new primer sets used on the first chip, and multiple old primer sets (retained from the original array). One hundred ng of gDNA from each bacterial strain was mixed for the first validation chip, which provided a total concentration of 16.6 ng per μl to 1.6×10^{-4} ng per μl, corresponding to 0.166 ng to 1.6×10^{-6} ng per reaction well on the SmartChip. Several additional strains were included with the second validation mixture, yielding a total of 14.6 ng per μl, which was diluted down to 1.6×10^{-4} ng per μl. All analysis regarding sensitivity and quantification described herein were based on results of the second and more complete validation of the updated array, while specificity analysis includes both validations.

A third validation plate was also performed with 10 environmental samples (Table S2, Supporting Information) that had previously been run with the old array. This validation plate was

performed to ensure that new primer sets would amplify environmental samples. This plate was also run to examine rate of amplification of primer sets that were redesigned due to questionable specificity. Genomic DNA was extracted from environmental samples with the PowerSoil® DNA isolation kit (MO BIO, Carlsbad, CA, USA), and DNA concentrations were measured using the Qubit Fluorometer (Life Technologies, OR, USA).

Conventional qPCR

Five of the primer sets were selected for comparison using a conventional real time cycler (Realplex, Eppendorf). Genes targeted for the comparison included: *blaTEM*, *tetM*, *aac(3)-xa*, *IS26* and *IS6* gene (primer sets 1512, 1513, 1545, 1546, 1560 as listed in Table S4, Supporting Information). Reactions on the conventional cycler consisted of 25 μl volume with 500 nM forward and reverse primers, four dilutions starting with 1.46 ng of DNA per reaction from the second validation mixture (described above), and suggested reagents from the Power SYBR Green PCR Master Mix (Life Technologies, 4367659). Real-time cycling conditions included a 10 min enzyme activation at 95°C followed by 40 cycles of 95°C for 15 sec and 60°C for 60 sec. A no template control was included for all assays, and all reactions were run in triplicate vials.

Data analysis

To compare coverage of new and replaced assays, an in-silico analysis was performed (Dec 2017), to generate the number of non-redundant targeted sequences for each primer set. Requirements for coverage included 100% perfect matches between primer sequences and known sequences in NCBI database with both forward and reverse primers.

An in-silico analysis of genomes and plasmids from strains included in the mock community validation mixture was used to verify targeted primer sets. To determine sensitivity of targeted assays, the number of genomic copies per reaction was estimated based on results of in-silico analysis (between strains and primer sets), the mass of gDNA per reaction per known strain, and the genome size of each strain. Only amplification events in two or more of three technical replicates were included in the analysis. A threshold cycle cutoff of 28 was used for analysis of targeted genes in environmental samples. For validation of known strain mixtures, a sliding threshold cycle cutoff was used to examine the rate of true/false positive results. A previously described equation (Looft *et al.* 2012) with a Ct cutoff of 28 or 31 was used to estimate genomic copies for comparison with actual genomic copies per reaction. Amplification efficiency was calculated using the equation $-1 + 10^{-1/\text{slope}}$ as previously described (Stedtfeld *et al.*, 2008). Figures plotting amplification efficiency were generated by subtracting efficiency by one to show absolute difference of slope from an ideal 3.33 threshold cycles per 10-fold dilution.

RESULTS

Summary of updated or removed primer sets

To assess questionable primer sets, data were summarized from 580 samples that have been run on the old array, which previously had 384 primer sets targeting 264 unique ARGs and MGEs (Fig. S1B, Supporting Information). Samples used for this evaluation included wastewater, farm animal manures, soil and surface water in US, Eurozone, China, and Antarctica (Table S3,

Supporting Information). Assays that rarely amplified or amplified often were categorized as questionable primers sets (Fig. 1A). For example, 36 primers sets did not amplify in any of the 580 samples and 41 additional primer sets amplified in less than 1% of samples. Primer sets that rarely amplified (less than five out of 580 samples) were separated into two additional categories, those that have less than five target sequences in NCBI and thus may be considered rare, and those that have more than five target sequences and may be false negatives. To help with this differentiation, we also examined the correlation between frequency of amplification and number of targeted sequences available in NCBI (Fig. S1A, Supporting Information). Primer sets that rarely amplified but had a higher number of target sequences were redesigned (n = 19). Primer sets that rarely amplified and had a low number of target sequences were removed (n = 14).

Conversely, primer sets that amplified in a majority of samples were also deemed as questionable akin to false positive and were redesigned (n = 22). Genes that required multiple primer sets for high coverage (i.e. divergent ARG sequences found in multiple species) on the old ARG array were also redesigned using EcoFunPrimer. For instances in which newly designed primer sets did not improve coverage, analysis was performed to differentiate and retain one older primer set for the given gene.

Following redesign and validation (described below), the updated qPCR ARG array contains 209 new and 175 old primer sets (retained from the old ARG qPCR array). Both arrays had 384 primer sets; however, the updated array targets 372 unique genes including 315 and 57 primer sets targeting ARGs and MGEs, respectively (Table S4, Supporting Information; Fig. 1B). For genes that were not previously targeted, emphasis was placed on selection of genes that are mobile as previously described (Pal *et al.* 2014; Hu *et al.* 2016). In total, primer sets for 147 previously untargeted genes were added to the updated array. Compared to the old array (Fig. S1B, Supporting Information), the number of primer sets targeting trimethoprim resistance (n = 17 new), fluoroquinolone resistance (n = 10 new), aminoglycoside resistance (n = 37 new), β-lactamase (n = 15 new) and phenicol (n = 12 new) increased significantly, among others. Compared to previous version of the array, a smaller number of primers were used for some categories of resistance including β-lactamases and tetracycline.). However, except for the chromosomal multiple drug resistance category, the total number of resistance genes targeted within each category, increased. The same universal bacteria primers targeting the 16S rRNA gene were also included similar to previous version of the array.

One goal of updating the qPCR ARG array was to capture diverse genes with the fewest number of primer sets, with the intent of maintaining coverage in terms of number of target sequences for a given gene. Thus, the number of target sequences was compared for all new and replaced primer sets. Results of gene coverage varied with the primer set; however, the total number of primer sets used for diverse genes was reduced from 91 to 52, and only 10% of total coverage (e.g. number of gene target sequences) was lost. For example, EcoFunPrimer captured 155 sequences targeting the *tetPB* gene with a single primer set, whereas four primer sets were used on the old array to target 133 sequences.

For 10 genes, EcoFunPrimer was only able to capture the same level of coverage as previously designed primer sets. In these instances, the older primer set with the highest level of coverage for a particular gene was retained and all additional primer sets for a given gene were removed. For this occurrence,
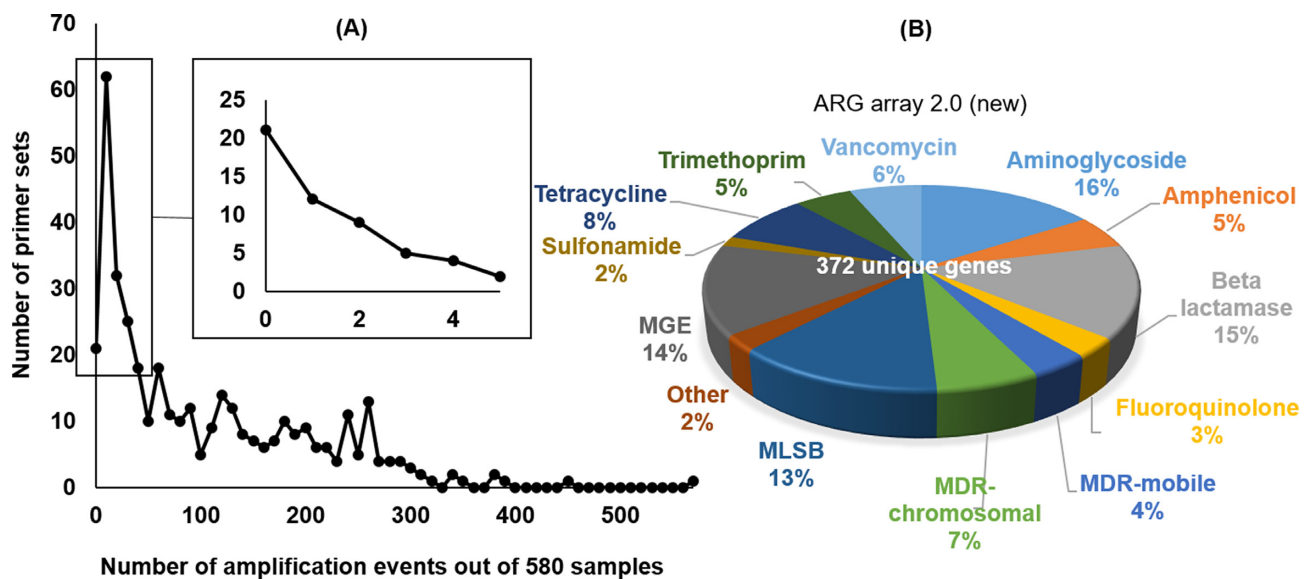
Figure 1. (**A**) Frequency of primer sets that amplified in 580 samples tested on the old array (e.g. ~10 primer sets amplified in 1 out of 580 samples), and (**B**) summarized distribution of target gene categories on the updated ARG qPCR array.

the number of assays was reduced from 21 to 10 primer sets, with only 17% loss of coverage. Additional primers sets targeting housekeeping genes were also removed.

Experimental validation of primer sets with mock communities

The specificity, sensitivity and quantitative capacity of the updated ARG qPCR array was experimentally validated with two separate mock communities containing mixtures of gDNA from sequenced bacterial strains (Table S1, Supporting Information) and 10 environmental samples (Table S2, Supporting Information). Validation of new primer sets using the first mock community resulted in 35 true positives, 140 true negatives, one false negative, and one false positive. The second mock community (targeting 123 primer sets) resulted in 121 true positive, five false positives, two false negative, and 254 true negative events. Assays that did not behave as expected were flagged and are not suggested for further use on the updated ARG array (Table S5, Supporting Information). The high rate of specificity observed with mock community mixtures further demonstrate utility of the new primer sets and EcoFunPrimer.

Analytical sensitivity of both new and old primer sets selected for the updated ARG array were also examined (Fig. 2A) and compared with varying threshold cycle (Ct) cutoff values (Fig. 2B). In detail, 76% of retained primer sets and 71% of new primer sets targeted by the mock community mixture amplified with 1 to 10 copies per reaction. With 10 to 100 copies per reaction, the percentage of targeted primer sets that amplified increased to 93% of old and 91% of new primer sets. These numbers are based on a Ct cutoff of 28. Using a higher Ct cutoff of 31, which has also been described (Wang et al., 2014, 2016), 86% of targeted primer sets amplified with 1 to 10 copies per reaction. In opposition, a higher Ct cutoff also influenced false positive calls (Fig. 2B).

The estimated number of copies was also influenced by the selected Ct cutoff value (Fig. S2A, Supporting Information). A linear best-fit line based on estimated (with Ct cutoff of 28) vs actual copies follows a 1:1 trend; however, estimated copies (with Ct cutoff of 31) overestimates the average amount of genes

10-fold. Considering all true positive instances for the tested dilutions and targeted primer sets, true positive amplification with Ct above 28 only occurred 11% of the time. Thus, a Ct cutoff of 28 should be routinely adopted for analysis of the ARG 2.0 array.

The influence of amplicon length on linearity of standard curves, amplification efficiency, and sensitivity was also examined (Fig. 3). Approximately 50% of the tested primer sets had an amplification efficiency between 90% and 110%. Fifty eight out of 112 primers had slightly lower slopes (amplification efficiency >110%), which may be due to one of the dilutions being near quantitative limits. Primer sets that generated amplicons greater than 150 bp had a negative influence on amplification efficiency (e.g. deviation of slope from an ideal 3.33 threshold cycles per 10-fold dilution) and sensitivity (Fig. 3B and C) as previously described (Martin et al. 2013; Debode et al. 2017). For example, 76% of targeted primers sets generating amplicons less than 100 bp amplified with 1 to 10 copies per reaction, while 65% of targeted primers sets generating amplicons greater than 200 bp amplified with 1 to 10 copies per reaction. Certainly, other parameters may also influence amplification efficiency and sensitivity such as GC content of primer and target as previously described (Stedtfeld et al. 2008; Bustin and Huggett 2017).

Experimental validation of primer sets with environmental samples

Ten environmental samples (Table S2, Supporting Information) were also analyzed using a subset of the updated array including clinical isolates, one dairy cow manure sample, two pig manure, two soil samples and two zoo animal fecal samples (bongo and tree shrew). These samples were also analyzed using the old 384 primer set ARG array, in which the same DNA concentrations was used for comparison. Between zero and 42 genes that were not previously targeted were detected using the updated array with the environmental samples; representing a 0 to 26% increase in detected genes (Fig. 4A).

A high Pearson correlation ($R^2 = 0.62$) of estimated copies between old and replaced primers sets was also observed (Fig. 4B). This correlation was skewed by lower abundance targets
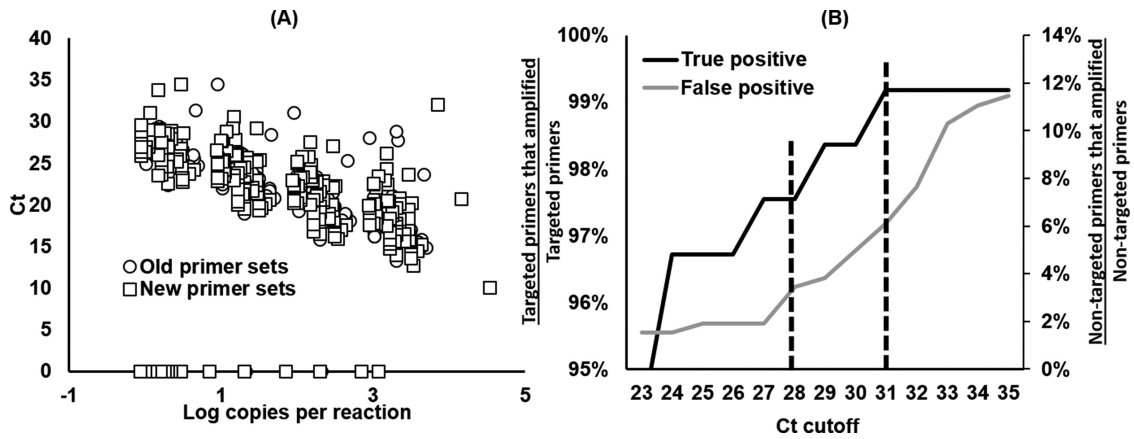
Figure 2. Specificity and sensitivity of primer sets targeted within a mock community of strains. (**A**) Sensitivity of retained old and new primer sets on the updated ARG array, and (**B**) percentage of targeted (black line, primary y-axis) and non-targeted (gray line, secondary y-axis) primer sets that are deemed positive amplification events based on the threshold cycle (Ct) cutoff.
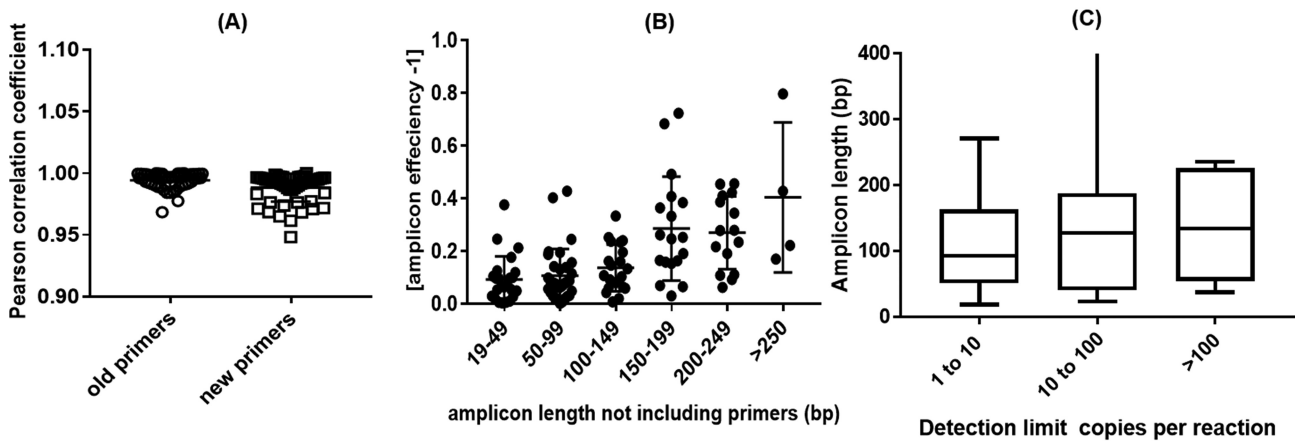

Figure 3. Influence of generated amplicon length on quantification and sensitivity of primer sets. (**A**) Pearson correlation coefficient based on standard curves generated with targeted primer sets and dilutions of strains in a mock community sample. (**B**) Influence of amplicon length on amplification efficiency, and (**C**) sensitivity.
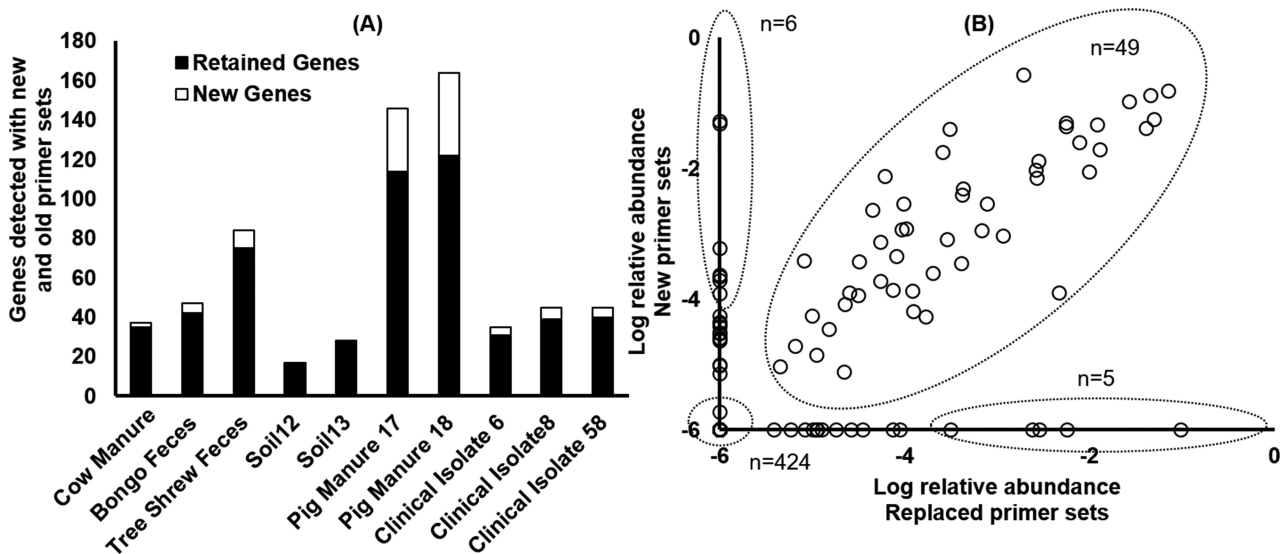

Figure 4. Testing 10 environmental samples on new and old primer sets used on the ARG 2.0 array. (**A**) Number of amplified genes with retained (old primer sets used on the new array) and new primer sets targeting genes that were not previously on the old array. (**B**) Comparing relative abundance of new primer sets (used to replace questionable primer sets) and the replaced primer sets (not used on the new array) within the 10 environmental samples.

that amplified in one primer set and not the other. Discrepancies between old and new primers are also expected for old primer sets (that were replaced) due to questionable specificity. For example, the old *fox5* gene primer set, which had been observed in 292 of 580 samples, including all Antarctic soil samples (Wang *et al.* 2016); amplified in six of the ten environmental samples. However, no amplification was observed in the same 10 environmental samples using a newly designed *fox5* gene primer set. Conversely, genes that had never amplified in 580 samples such as the *spcN* gene, which confers resistance to aminoglycosides, amplified with a new primer set in one of the pig manure samples.

DISCUSSION

A list of old and new assays with suggested primer sets based on validation, and results of experiments with mock communities is provided (Tables S4 and S5, Supporting Information). Primer sets listed as 'old retained,' 'new target,' and 'new primer replace old primer' are suggested for prospective analysis. The updated ARG qPCR array is expected to yield greater comparative analysis due to the elimination of primer sets that rarely or often amplify. As observed in Fig. 4, updated primer sets correlate well with old primers; thus comparisons can still be made for samples analyzed using the old or new arrays. In addition, there may be value in using a common subset of primer sets (e.g. *tetM*, *intI1*, *qacEdelta* genes) in all samples so that data are broadly comparable. Notably, the described number of 372 and 265 uniquely targeted genes on the updated and old array does not fully include the total number of targeted genes. For example, 21 fluoroquinolone resistance genes are targeted with nine primer sets, which is reported as nine unique genes. Thus, the total number of uniquely targeted genes is underestimated. Similar instances are also expected with β-lactamase genes (e.g. *blaOXA and blaCMY*).

Reducing number of primers required for diverse genes

While multiple primer design tools have been described and reviewed (Thornton and Basu 2015; Kim *et al.* 2016), the RDP Eco-FunPrimer design tool provided a more automated mechanism for capturing the highest possible level of diversity for target genes of interest. In this study, EcoFunPrimer was used to reduce the number of primers required for high coverage of a divergent gene, providing additional space on the array.

A common inquiry with the old ARG qPCR array involves analysis of assays in which multiple primer sets are used to target the same gene (e.g. nine primer sets targeting the *ampC* gene). While these primer sets were initially designed to target different sub-groups of the same gene, manual design of primer sets incurred some overlap, questioning the ability to simply sum results or treat each primer set as an individual group. For example, the nine primers used to target 808 non-redundant *ampC* gene sequences contained 30% overlap between primer sets, thus the summed abundance is not an accurate means of quantification. Redesign of a single primer set targeting the *ampC* gene, only provided 57% coverage obtained with the previously used nine primers, but removes issues related to quantitative analysis of multiple primer sets for the same gene. Thus, the ARG qPCR array was updated to remove occurrence of multiple primer sets for the same gene.

ARG assays on various platforms

Quantitative values observed with the ARG array should be comparable to qPCR on a conventional cycler in that major differences are merely volume and material used to enclose the reaction. For example, Morrison and coauthors observed similar quantitative accuracy and precision between reactions run in 384 well plates (2 μl) and a qPCR array with 33 nl reactions (Morrison *et al.*, 2006). A comparison of five primer sets run via conventional qPCR and on the ARG array also showed similar results (Table S6, Supporting Information). Limited volume and mass of DNA per reaction can reduce sensitivity of high-throughput qPCR arrays. For instance, conventional qPCR reactions can be run with DNA amounts in excess of 10 ng per reaction, while a maximum mass of 0.5 ng of DNA per reaction is suggested with SmartChip.

While a majority of studies using the ARG qPCR array have been performed on the Takara SmartChip (previously Wafergen) platform, primer sets are expected to behave similarly with different platforms. We evaluated the previous ARG qPCR assays on three separate platforms—Access Array from Fluidigm, OpenArray from BioTrove (now part of Life Technologies), and SmartChip and all were comparable in terms of throughput. In addition, SmartChip and Fluidigm platforms allow harvesting of amplicons for sequencing (Johnson *et al.* 2016). The SmartChip platform provided greater analytical sensitivity (due to its larger 100 nl reaction well volumes) and flexibility in terms of customizing the number of assays vs. the number of samples matrix. The smaller reaction volume on the Fluidigm system required pre-amplification, which may impact the accuracy and sensitivity. The OpenArray platform is less flexible in terms of run-to-run customization of assays. Use of TaqMan probes instead of SYBR-green based assays on the OpenArray has made primer design for large group of similar sequences more difficult and increased consumable costs.

High-throughput ARG qPCR array and shotgun metagenomics

Numerous approaches may be employed to obtain information about antimicrobial resistance genes. These include: (i) high throughput qPCR, (ii) multiplexed qPCR, (iii) high throughput or multiplexed qPCR combined with sequencing, (iv) targeted next-generation sequencing, and (v) shotgun metagenomics. They vary in terms of their depth of information, ease of data analysis, need for bioinformatics pipeline, cost per sample, and the number of samples that must be analyzed to observe the economy of scale. No single approach is good with respect to all parameters. The high throughput qPCR approach has simpler analysis protocols and less expensive, but it does not provide sequence-related information. Depending upon the primer design approach (e.g. for conserved regions), it does have the potential to capture novel sequences. For instance, an *in-silico* analysis of primer sets for all sequences available in public databases in 2015 and 2017 reveals coverage of multiple sequences that were previously uncharacterized (Fig. S2B, Supporting Information). The next-generation sequencing-based approaches and specially targeted sequencing provides sequence-related information and may become cost-effective when a larger number of samples are analyzed together. However, it does require somewhat longer time for sample and library preparation as well as ARG-focused data analysis pipeline. Thus, the type of information desired and cost may determine the choice of approach.

While shotgun metagenomics has been widely adopted to profile known functional and taxonomic genes in a given sample , the qPCR array may potentially be better suited for hypothesis driven studies of targeted genes with benefits of cost, sensitivity, and quantification. Unlike qPCR, shotgun metagenomics will also provide unlimited coverage of known genes, with exploration of novel genes dependent on means of analysis. QPCR primer sets designed from conserved regions of known genes may also allow for detection of unknown sequences. For instance, an *in silico* analysis of primer sets for all sequences available in public databases in 2015 and 2017 reveals coverage of multiple sequences that were previously uncharacterized (Fig. S2B, Supporting Information).

Further considerations

The updated array is expected to provide greater utility and specificity, with over 98% of validated instances incurring true positive/negative calls. Notably, not all primer sets were validated, and absolute quantification is limited without standard curves. Targets from different genera may also influence specificity and sensitivity, limiting utility of standard curves. Thus, sample sets that include controls (e.g. samples from similar matrices that have not been influenced by anthropogenic selective pressure) are recommended, allowing for quantitative comparison of genes via $2^{-\triangle\triangle Ct}$ method (Livak and Schmittgen 2001; Zhu *et al.* 2013).

Some primers may also target genes that are not easily distinguished from common class relatives present in most organisms, such as the *vanR* and *vanS* gene, a phosphorylase and two component regulator, and *bacA* gene. The previous *bacA* and *vanSB* gene primer sets (primer numbers 310 and 158, respectively) amplified in 36% and 57% of the 580 previously tested samples, indicating that they may target functions unrelated to antibiotic resistance. The inclusion of a control sample within a given sample set (e.g. samples that have not been influenced by anthropogenic selective pressure) should aid in differentiation of common class relatives and resistance genes.

SUPPLEMENTARY DATA

Supplementary data are available at *FEMSEC* online.

REFERENCES

Alanis A. Resistance to antibiotics: are we in the post-antibiotic era ? *Arch Med Res* 2005;**36**:697–705.

Bustin S, Huggett J. qPCR primer design revisited. *Biomol Detect Quantif* 2017;**14**:19–28.

Courvalin P. Predictable and unpredictable evolution of antibiotic resistance. *J Intern Med* 2008;**264**:4–16.

Debode F, Marien A, Janssen E *et al*. The influence of amplicon length on real-time PCR results. *Biotechnol Agron Soc Env* 2017;**21**:3–11.

Do T, Tamames J, Stedtfeld R *et al*. Antibiotic resistance gene detection in the microbiome context. *Microb Drug Resist* 2018;**24**:542–6.

Forsberg K, Reyes A, Bin W *et al*. The shared antibiotic resistome of soil bacteria and human pathogens. *Science* 2012;**337**:1107–11.

Hu Y, Yang X, Li J *et al*. The bacterial mobile resistome transfer network connecting the animal and human microbiomes. *Appl Environ Microbiol* 2016;**82**:6672–81.

Hu Y, Yang X, Qin J *et al*. Metagenome-wide analysis of antibiotic resistance genes in a large cohort of human gut microbiota. *Nat Commun* 2013;**4**:2151.

Johnson TA, Stedtfeld RD, Wang Q *et al*. Clusters of antibiotic resistance genes enriched together stay together in swine agriculture. *mBio* 2016;**7**:2214–5.

Karkman A, Johnson TA, Lyra C *et al*. High-throughput quantification of antibiotic resistance genes from an urban wastewater treatment plant. *FEMS Microbiol Ecol* 2016;**92**:1–7.

Kim H, Kang N, An K *et al*. MRPrimerW: a tool for rapid design of valid high-quality primers for multiple target qPCR experiments. *Nucleic Acids Res* 2016;**44**:W259–66.

Kumarasamy KK, Toleman M, Walsh T *et al*. Emergence of a new antibiotic resistance mechanism in India, Pakistan, and the UK: a molecular, biological, and epidemiological study. *Lancet Infect Dis* 2010;**10**:597–602.

Liu Y, Wang Y, Walsh T *et al*. Emergence of plasmid-mediated colistin resistance mechanism MCR-1 in animals and human beings in China: a microbiological and molecular biological study. *Lancet Infect Dis* 2016;**16**:161–8.

Livak K, Schmittgen T. Analysis of relative gene expression data using realtime quantitative PCR and the 2(-Delta Delta C(T)) method. *Methods* 2001;**25**:402–8.

Looft T, Johnson TA, Allen HK *et al*. In-feed antibiotic effects on the swine intestinal microbiome. *Proc Natl Acad Sci* 2012;**109**:1691–6.

Martin B, Raurich S, Garriga M *et al*. Effect of amplicon length in propidium monoazide quantitative PCR for the enumeration of viable cells of salmonella in cooked ham. *Food Anal Methods* 2013;**6**:683–90.

Martinez J, Coque T, Baquero F. What is a resistance gene? Ranking risk in resistomes. *Nat Rev Micro* 2015;**13**:116–23.

Morrison T, Hurley J, Garcia J *et al*. Nanoliter high throughput quantitative PCR. *Nucleic Acids Res* 2006;**34**:e123.

Muurinen J, Stedtfeld R, Karkman A *et al*. Influence of manure application on the environmental resistome under Finnish agricultural practice with restricted antibiotic use. *Environ Sci Technol* 2017;**51**:5989–99.

Muziasari WI, Pärnänen K, Johnson TA *et al*. Aquaculture changes the profile of antibiotic resistance and mobile genetic element associated genes in Baltic Sea sediments. *FEMS Microbiol Ecol* 2016;**92**:fiw052.

Pal C, Bengtsson-Palme J, Kristiansson E *et al*. The structure and diversity of human, animal and environmental resistomes. *Microbiome* 2016;**4**:54.

Pal C, Bengtsson-Palme J, Rensing C *et al*. BacMet: antibacterial biocide and metal resistance genes database. *Nucl Acids Res* 2014;**42**:D737–43.

Qian X, Gu J, Sun W *et al*. Diversity, abundance, and persistence of antibiotic resistance genes in various types of animal manure following industrial composting. *J Hazard Mater* 2018;**344**:716–22.

Stedtfeld R, Stedtfeld T, Fader K *et al*. TCDD influences reservoir of antibiotic resistance genes in murine gut microbiome. *FEMS Microbiol Ecol* 2017;**93**:fix058.

Stedtfeld R, Stedtfeld T, Waseem H *et al*. Isothermal assay targeting class 1 integrase gene for environmental surveillance of antibiotic resistance markers. *J Environ Manage* 2017;**198**:213–20.

Stedtfeld RD, Baushke SW, Tourlousse DM *et al*. Development and experimental validation of a predictive threshold cycle equation for quantification of virulence and marker genes by high-throughput nanoliter-volume PCR on the OpenArray platform. *Appl Environ Microbiol* 2008;**74**:3831–8.

Stedtfeld RD, Williams MR, Fakher U *et al*. Antimicrobial resistance dashboard application for mapping environmental occurrence and resistant pathogens. *FEMS Microbiol Ecol* 2016;**92**:1–9.

Thornton B, Basu C. Rapid and simple method of qPCR primer design. *Methods Mol Biol* 2015;**1275**:173–9.

Wang F-H, Qiao M, Su J-Q *et al*. High throughput profiling of antibiotic resistance genes in urban park soils with reclaimed water irrigation. *Environ Sci Technol* 2014;**48**:9079–85.

Wang F, Stedtfeld R, Kim O *et al*. Influence of soil characteristics and proximity to Antarctic research stations on abundance of antibiotic resistance genes in soils. *Environ Sci Technol* 2016;**50**:12621–9.

Xu L, Ouyang W, Qian Y *et al*. High-throughput profiling of antibiotic resistance genes in drinking water treatment plants and distribution systems. *Environ Pollut* 2016;**213**:119–26.

Zhang X, Wu B, Zhang Y *et al*. Class 1 integronase gene and tetracycline resistance genes tetA and tetC in different water environments of Jiangsu Province, China. *Ecotoxicology* 2009;**18**:652–60.

Zhu Y-G, Johnson TA, Su J-Q *et al*. Diverse and abundant antibiotic resistance genes in Chinese swine farms. *Proc Natl Acad Sci USA* 2013;**110**:3435–40.

Zhu Y-G, Zhao Y, Li B *et al*. Continental-scale pollution of estuaries with antibiotic resistance genes. *Nat Microbiol* 2017;**2**:16270.