Advancement to the highly parallel qPCR array targeting antibiotic resistance genes and mobile genetic elements

Robert D. Stedtfeld[a,+], Xueping Guo[b,c,d,+], Tiffany M. Stedtfeld[a], Hongjie Sheng[d], Maggie R. Williams[a], Kristin Hauschild[d], Santosh Gunturu[d], Leo Tift[d], Fang Wang[d], Adina Howe[e], Benli Chai[d], James R. Cole[d], James M. Tiedje[d], Syed A. Hashsham[a,d,*]

[a]Department of Civil and Environmental Engineering, East Lansing, Michigan 48824

[b]College of Environmental Science and Engineering, Tongji University, Shanghai, China

[c]Department of Plant, Soil and Microbial Sciences, Michigan State University, East Lansing, Michigan 48864

[d]Center for Microbial Ecology, Michigan State University, East Lansing, Michigan 48864

[e]Department of Agricultural and Biosystems Engineering, Ames, Iowa 50010

[+]Equal contributions

[*]Corresponding author:

Syed A. Hashsham, Ph. D.

Michigan State University

Center for Microbial Ecology

Department of Civil and Environmental Engineering

East Lansing, MI 48824-1319

Keywords: high-throughput qPCR primer set, antibiotic resistance gene, mobile genetic element, validation, qPCR array, SmartChip, primer design tool

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 newly developed primer design tool 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 mobile ARGs and MGEs 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 (termed updated ARG qPCR array 2.0) is recommended 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 five 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 5,184 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. However, multiple ARG types (e.g. NDM, mcr-1 genes) have since been discovered (Kumarasamy, KK Toleman *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 has 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), waste-water (Karkman *et al.* 2016), and fecal matrices (Qian *et al.* 2018; Do *et al.* 2107) have now been run on 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/PrimerDesign), 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 primers 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 compressive list of old and new primer sets with recommended assays for the updated ARG qPCR array (referred as ARG Array 2.0).

Materials and methods

Reference sequences collection and primer design

Selection of primer sets and gene target 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

78 genomes as described (Hu *et al.*, 2016), or genes listed as experimentally confirmed plasmid borne
79 genes in the Antibacterial and Biocide and Metal Resistance Genes Database (Pal *et al.*, 2014).
80 Reference sequences for selected target genes were assembled using the RDP FunGene Pipeline
81 (http://fungene.cme.msu.edu/), which can be used to automatically downloads, aligns, and trims
82 sequences for a given gene. For genes not currently listed in Fungene, reference sequences from the
83 latest version of ARG-ANNOT AA V3 March 2017 (Antibiotic Resistance Gene-ANNOTation) were
84 used to gather additional sequences of high similarity via NCBI.

85 Following sequence collection, primer sets were designed for these gete targets using the RDP
86 PrimerDesign Tool (https://github.com/rdpstaff/PrimerDesign). The tool highlights conserved regions
87 of a gene with thermodynamically stable primer pairs for qPCR and automatically suggests primers
88 with or without degeneracy as specified by the user. In addition, the tool evaluates conserved genomic
89 regions, which can be used for manual selection of forward and reverse positions with desired
90 amplicon length and coverage. For primers in ARG qPCR array 2.0, the following specificiations
91 were used with RDP PrimerDesign: sample select command with sliding scale, a theoretical melting
92 temperature of 60 ^{o}C, and zero degeneracy. A weighting system was implemented to select primers
93 that cover sequences which improve the diversity of target genes. The PrimerDesign tool was run on
94 the HPCC (High performance computer center) at Michigan State University.

95 *Validation of the updated ARG qPCR array*

96 Specificity, sensitivity, and amplification efficiency of primer sets were tested using a dilution
97 series of gDNA extracted from multiple organisms (28 strains) obtained from ATCC (Table S1). Two
98 SmartChips 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 plate 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 primer sets generated using the PrimerDesign tool 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 $1.6x10^{-4}$ ng per µl, corresponding to 0.166 ng to $1.6x10^{-6}$ 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 $1.6x10^{-4}$ ng per µl. All analysis regarding sensitivity and quantification described herein was 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) that had previously been run with the old array. This validation plate was performed to ensure that new primer sets would amplify in 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, Carsbad, CA, USA), and DNA concentrations were measured using the Qubit Fluorometer (Life Technology, OR, USA).

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 and 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.

Results

Summary of updated or removed primer sets

To assess questionable primer sets, data was 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). Samples used for this evaluation included wastewater, farm animal manures, soil and surface water in US, Eurozone, China, and Antarctica (Table S3). 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). 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 the PrimerDesign tool. 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, 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), the number of primer sets targeting trimethoprim resistance (n=17 new), fluoroquinolone resistance (n=10 new), aminoglycoside resistance (n=37 new), beta lactamase (n=15 new), and phenicol (n=12 new) increased significantly, among others.

While a smaller number of primers were used for some categories of resistance (e.g. beta lactamase, tetracycline), the total number of targeted resistance genes within all categories increased (excluding chromosomal multiple drug resistance genes).

One goal of updating the qPCR ARG array was to capture diverse genes with the fewest number of primers 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 targets) was lost. For example, the PrimerDesign tool 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, the PrimerDesign tool 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, the number of assays was reduced from 21 to 10 primer sets, with only 17% loss of coverage. Additional primers sets targeting housekeeping genes and chromosomal multiple drug resistance genes (MDR) that experimentally correspond with other MDR 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) and 10 environmental samples (Table S2). 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 255 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 S4). The high rate of specificity observed with mock community mixtures further demonstrate utility of the new primer sets and the PrimerDesign tool.

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 percent 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). 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 array.

203 The influence of amplicon length on linearity of standard curves, amplification efficiency, and
204 sensitivity was also examined (Fig 3). Primer sets that generated amplicons greater than 150 bp were
205 prone to lower amplification efficiencies (Fig 3B) and sensitivity (Fig 3C) as previously described
206 (Martin *et al.* 2013; Debode *et al.* 2017). For example, 76% of targeted primers sets generating
207 amplicons less than 100 bp amplified with 1 to 10 copies per reaction, while 65% of targeted primers
208 sets generating amplicons greater than 200 bp amplified with 1 to 10 copies per reaction. Certainly,
209 other parameters may also influence amplification efficiency and sensitivity such as GC content of
210 primer and target as previously described (Stedtfeld *et al.* 2008; Bustin and Huggett 2017).
211 *Experimental validation of primer sets with environmental samples*
212 Ten environmental samples (Table S2) were also run on a subset of the updated array including
213 clinical isolates, one dairy cow manure sample, two pig manure, two soil samples and two zoo animal
214 fecal samples (bongo and tree shrew). These samples were also run on the old 384 primer set ARG
215 array, in which the same DNA concentrations was used for comparison. Between zero and 42 genes
216 that were not previously targeted were detected using the updated array with the environmental
217 samples; representing a 0 to 26% increase in detected genes (Fig 4A).
218 A high Pearson correlation (R^2=0.62) of estimated copies between old and replaced primers sets
219 was also observed (Fig 4B). This correlation was skewed by lower abundance targets that amplified in
220 one primer set and not the other. Discrepancies between old and new primers are also expected for old
221 primer sets (that were replaced) due to questionable specificity. For example, the old *fox5* gene primer
222 set, which had been observed in 292 of 580 samples, including all Antarctic soil samples (Wang *et al.*
223 2016); amplified in six of the 10 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 (Table S4). 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 merely a suggestion of primer sets for future studies, which is expected to yield greater comparative analysis due to the elimination of primer sets that rarely or often amplify. Thus, comparison with previously tested samples may warrant use of the old array. 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 in underestimated. Similar instances are also expected with beta lactamase genes (e.g. *blaOXA, 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 PrimerDesign Tool provided a more automated mechanism for capturing the highest possible level of diversity for target genes of interest. In this study, the notebook 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 removed occurrence of multiple primer sets for the same gene.

QPCR on various high-throughput platforms

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. It should be noted; however, that the old ARG qPCR assays have been tested on three separate systems (Fluidigm, OpenArray, SmartChip). All three systems are comparable in terms of throughput; however, the SmartChip platform provided greater sensitivity (100 nL reaction well), and flexibility in terms of different assay/sample formats. Amplicons can also be harvested from both the SmartChip and Fluidigm platforms (Johnson *et al.* 2016); however, smaller reaction volume on the Fluidigm system requires pre-amplification, reducing accuracy and sensitivity. The TaqMan OpenArray platform also suffers in terms of flexibility, consumable chip costs, and reduced coverage inherent to use of TaqMan probes instead of SYBR based assays.

High-throughput ARG qPCR array and shotgun metagenomics

While shotgun metagenomics has been widely adopted to profile known functional and taxonomic genes in a given sample (Yergeau *et al.* 2010; Myrold, Zeglin and Jansson 2014), 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).

Further considerations

The updated array is expected to be 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 are recommended, allowing for quantitative comparison of genes via $2^{-\Delta\Delta Ct}$ 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 (Wright 2017). 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. Again, control samples should aid in differentiation of anthropogenic selective pressures for these genes.

Acknowledgments

Many thanks to Kornelia Smalla, Windi Muziasari, and Ed Topp for reviewing and suggesting new genes that are now targeted on the updated ARG qPCR array. This research was supported in part by a grant from the MSU Center for Health Impacts of Agriculture (CHIA), and a China Scholarship Council (CSC) fellowship to Xueping Guo.

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* 2107.

Forsberg K, Reyes A, Bin W *et al.* The shared antibiotic resistome of soil bacteria and human pathogens. *Science (80-)* 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 Env Microb* 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, Bagaria J *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 USA* 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**:683–690.

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

Muurinen J, Stedtfeld R, Karkman A *et al.* Influence of manure application on the environmental resistome under Finnish agricultural practice with restricted antibiotic use. *Env 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.

Myrold D, Zeglin L, Jansson J. The potential of metagenomic approaches for understanding soil microbial processes. *Soil Sci Soc Am J* 2014;**78**:3–10.

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. *Nucleic 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 Mat* 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* 2017a;**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 Env Manag* 2017b;**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 Env 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. *Env 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. *Env Sci Technol* 2016;**50**:12621–9.

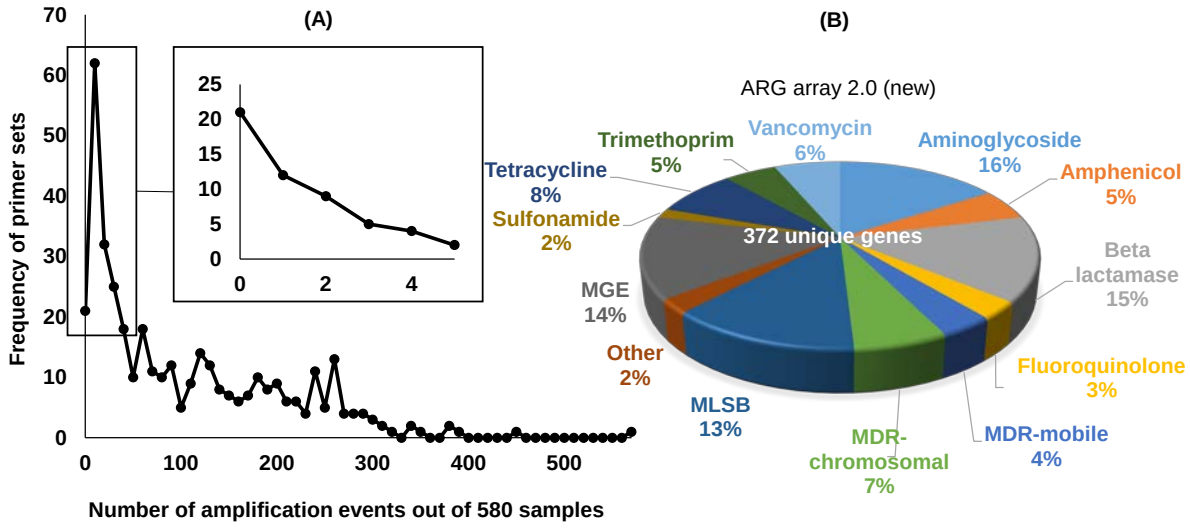
Wright G. No Title. 2017, Personal communication.

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

Yergeau E, Hogues H, Whyte L *et al.* The functional potential of high Arctic permafrost revealed by metagenomic sequencing, qPCR and microarray analyses. *ISME J* 2010;**4**:1206–14.

Zhang X, Wu B, Zhang Y *et al.* Class 1 integronase gene and tetracycline resistance genes

371 tetA and tetC in different water environments of Jiangsu Province, China. *Ecotoxicology*
372 2009;**18**:652–60.
373 Zhu Y-G, Johnson TA, Su J-Q *et al.* Diverse and abundant antibiotic resistance genes in
374 Chinese swine farms. *Proc Natl Acad Sci USA* 2013;**110**:3435–40.
375 Zhu Y-G, Zhao Y, Li B *et al.* Continental-scale pollution of estuaries with antibiotic
376 resistance genes. *Nat Microbiol* 2017;**2**:16270.
377
378

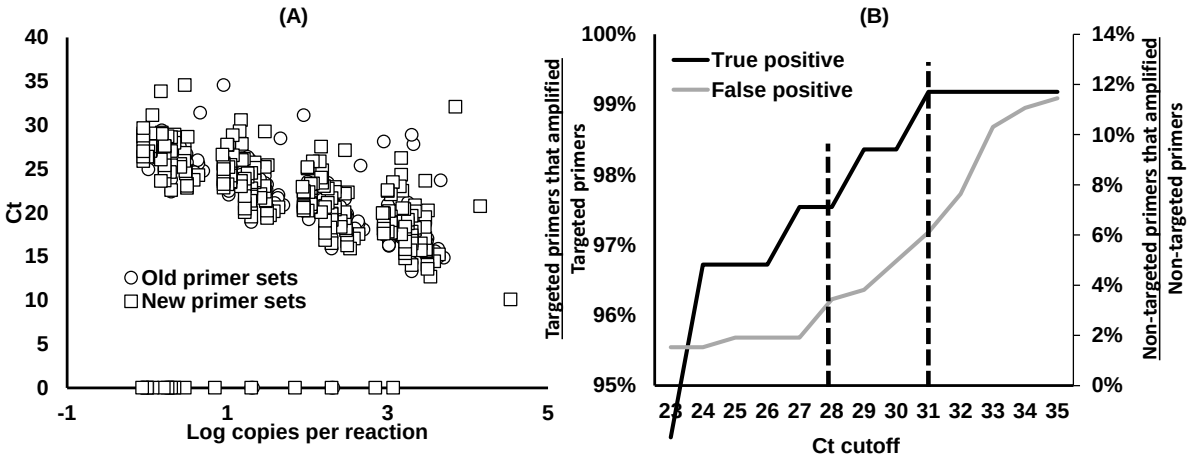


379

380 **Fig 1.** Frequency of primer sets that amplified in 580 samples tested on the old array (e.g. ~10 primer

381 sets amplified in 1 out of 580 samples), and (B) summarized distribution of target gene categories on

382 the updated ARG qPCR array.



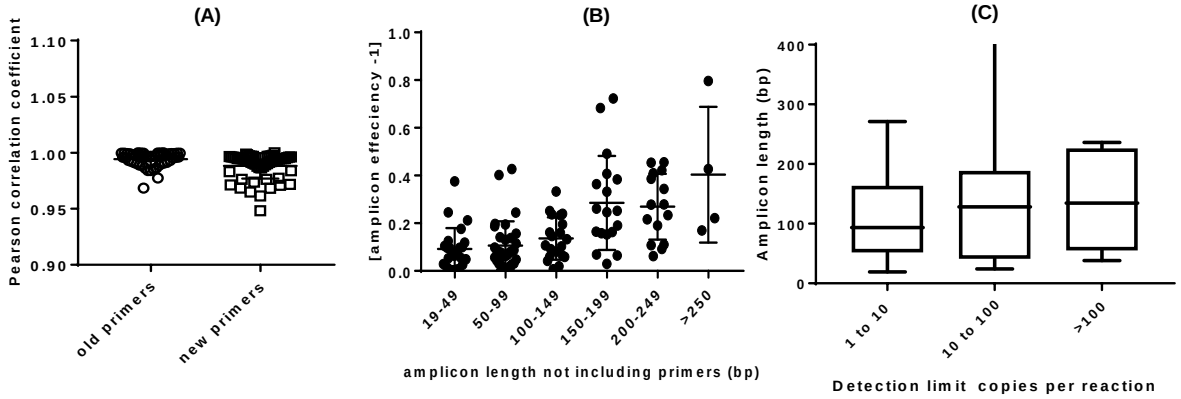
383

384 **Fig 2.** Specificity and sensitivity of primer sets targeted within a mock community of strains. (A)

385 Sensitivity of retained old and new primer sets on the updated ARG array, and (B) percent of targeted

386 (black line, primary y-axis) and non-targeted (gray line, secondary y-axis) primer sets that are deemed

387 positive amplification events based on the threshold cycle (Ct) cutoff.



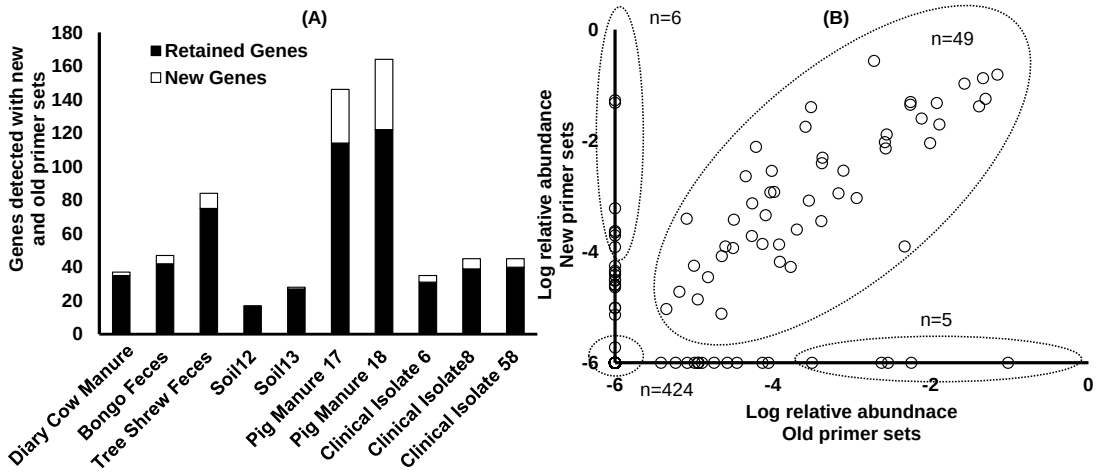
388

389 **Fig 3.** Influence of generated amplicon length on quantification and sensitivity of primer set. (A)

390 Pearson correlation coefficient based on standard curves generated with targeted primer sets and

391 dilutions of strains in a mock community sample. (B) Influence of amplicon length on amplification

392 efficiency, and (C) sensitivity.



393

394 **Fig 4.** Testing 10 environmental samples on new and old primer sets used on the ARG 2.0 array. (A)

395 Number of amplified genes with retained (old primer sets used on the new array) and new primer sets

396 targeting genes that were not previously on the old array. (B) Comparing relative abundance of new

397 primer sets (used to replace questionable primer sets) and the replaced primer sets (not used on the

398 new array) within the 10 environmental samples.