(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
19 February 2015 (19.02.2015)

WIPO | PCT

(10) International Publication Number
WO 2015/023616 A2

(51) International Patent Classification:
C12Q 1/68 (2006.01)

(21) International Application Number:
PCT/US2014/050613

(22) International Filing Date:
12 August 2014 (12.08.2014)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
61/864,865 12 August 2013 (12.08.2013) US

(71) Applicant: **THE JOHNS HOPKINS UNIVERSITY** [US/US]; ATTN: Johns Hopkins Technology Transfer, 100 North Charles Street, 5th Floor, Baltimore, MD 21201 (US).

(72) Inventors: **YANG, Samuel**; ATTN: Johns Hopkins Technology Transfer, 100 North Charles Street, 5th Floor, Baltimore, MD 21201 (US). **WANG, Jeff**; ATTN: Johns Hopkins Technology Transfer, 100 North Charles Street, 5th Floor, Baltimore, MD 21201 (US). **FRALEY, Stephanie, I.**; ATTN: Johns Hopkins Technology Transfer, 100 North Charles Street, 5th Floor, Baltimore, MD 21201 (US).

(74) Agents: **MORALES, Joseph** et al.; Whiteford, Taylor & Preston LLP, Seven Saint Paul Street, Baltimore, MD 21202 (US).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Published:

— *without international search report and to be republished upon receipt of that report (Rule 48.2(g))*

— *with sequence listing part of description (Rule 5.2(a))*

WO 2015/023616 A2

(54) Title: METHOD FOR COMPREHENSIVE, QUANTITATIVE, AND HIGHLY SENSITIVE DISCRIMINATION OF NUCLEIC ACID SEQUENCES IN HOMOGENEOUS AND HETEROGENEOUS POPULATIONS

(57) Abstract: A method for the detection of nucleic acids in heterogeneous samples is provided. Very small quantities of nucleic acids can be identified in a sample through the use of digital High Resolution Melt Analysis (dHRMA). In one step of the method, melt curves for specific target sequences are identified. In a second step, primers for the amplification of target sequences are utilized in optimized digital PCR for detection of the target sequences. In a subsequent step, the melt curves of the sequences present in the sample are obtained. In the final step, the melt curves obtained from the sample are compared to the standard melt curves for the target sequences.

METHOD FOR COMPREHENSIVE, QUANTITATIVE, AND HIGHLY SENSITIVE DISCRIMINATION OF NUCLEIC ACID SEQUENCES IN HOMOGENEOUS AND HETEROGENEOUS POPULATIONS

CROSS REFERENCE TO RELATED APPLICATIONS

This application is based upon and claims priority from co-pending U.S. Provisional Patent Application Serial No. 61/864,865 entitled "METHOD FOR COMPREHENSIVE, QUANTITATIVE, AND HIGHLY SENSITIVE DISCRIMINATION OF NUCLEIC ACID SEQUENCES IN HOMOGENEOUS AND HETEROGENEOUS POPULATIONS," filed with the U.S. Patent and Trademark Office on August 12, 2013, by the inventors herein, the specification of which is incorporated herein by reference.

STATEMENT OF GOVERNMENT RIGHTS

The U.S. Government has a paid-up license in this invention and the right in limited circumstances to require the patent owner to license others on reasonable terms as provided by the terms of Grant No. 1159771 awarded by the National Science Foundation and by the National Institute of Health.

BACKGROUND OF THE INVENTION

FIELD OF THE INVENTION

This invention relates to the field of biotechnology and, more specifically, to methods of measuring nucleic acids in homogeneous and heterogeneous samples.

DESCRIPTION OF THE BACKROUND

A key challenge in the fields of clinical diagnostic development and basic biomolecular research involves the ability to quantitatively and accurately identify even single copies of all nucleic acid targets of interest in a heterogeneous sample. Pattison, S.H. et al., Molecular detection of CF lung pathogens: Current status and future potential, J. of Cystic Fibrosis (2013); Park, S.Y., et al. Clinical significance and outcome of polymicrobial

Staphylococcus aureus bacteremia, J Infect., 65, 119-127 (2012); Pasic, M.D., et al. Genomic Medicine: New Frontiers and New Challenges. *Clin Chem*, **5**, 158-167 (2013); Pritchard, C.C., et al. MicroRNA profiling: approaches and considerations. *Nature Reviews Genetics*, 13:358-369 (2012); Blainey, P.C. The future is now: single-cell genomics of bacteria and archaea. *FEMS Microbiol Rev* (2013); Yang, S. and Rothman, R.E. PCR-based diagnostics for infectious diseases: uses, limitations, and future applications in acute-care settings. *Lancet Infect Dis*, 4:337-348 (2004). Typically, hybridization assays such as microarrays are used for broad, semi-quantitative profiling, while multi-reaction quantitative PCR (qPCR) is used for enumeration of multiple species. The qPCR arrays currently available require that the sample input be divided among multiple wells resulting in multiple target molecules per well and only one primer set per well targeting an individual species as shown in Figure 1(b); sensitivity is inversely related to the number of reactions. Low level species may not be distributed to a well containing appropriate primers, and unknown species will not be discovered. Multiplexed qPCR involves a single reaction containing all target molecules and multiple sets of specific primers and probes targeting individual species; single copy sensitivity is possible as shown in Figure 1(c). However, the number of detectable target species is limited by the resolution of fluorescent probe spectra to about 4 species and unknown targets will not be discovered.

Microarrays and qPCR techniques lack the sensitivity to detect species in low level concentrations since the sample must be split across many reactions containing distinct sets of primers or probes. Chuaqui, R.F., *et al.* Post-analysis follow-up and validation of microarray experiments. *Nature Genetics*, 32:509-514 (2002); Whale, A.S., et al. Comparison of microfluidic digital PCR and conventional quantitative PCR for measuring copy number variation. *Nucleic Acids Res*, 40:e82 (2012); Schmittgen, T.D., et al. Real-time PCR quantification of precursor and mature microRNA. *Methods*, 44:31-38 (2008). Likewise, their

multiplexed formats rely on highly specific primer or probe annealing to discriminate single nucleotide differences, often resulting in inaccuracies. Prior knowledge of target molecule sequences within a sample is required and discovery of novel species is not possible. Broad-based amplification followed by "sloppy" molecular beacon melt can identify a variety of singular target molecules as shown in Figure 1(d), but in a heterogeneous sample, amplification competition may obscure low level species and complex melt curves involving unknown targets cannot be resolved. The same is true for broad-based qPCR for amplification of all target molecules in the sample followed by bulk high resolution met analysis (HRMA), cannot distinguish components of a heterogeneous mixture involving unknown targets.

With the advent of next generation sequencing (NGS), highly sensitive, specific, and multiplexed detection of both known and unknown target molecules in a mixed sample is possible, and NGS's utility as a diagnostic and research tool is being explored. However, errors in base calling, alignment, and assembly of sequence data occur (Nielsen, R., et al. Genotype and SNP calling from next-generation sequencing data. *Nature Reviews Genetics*, 12:443-451 (2011)); and deep-sequencing remains a costly, time consuming, multi-step process that is not accessible to most clinical and basic research laboratories (Glenn, T.C. Field guide to next-generation DNA sequencers. *Molecular Ecology Resources*, 11:759-769 (2011)).

The use of Real Time – PCR (RT-PCR) and bulk HRMA has previously been described in WO2009/134470 (US2013/0217588A1), which is incorporated herein by reference in its entirety.

SUMMARY OF THE INVENTION

A method for the detection of nucleic acid sequences in a sample is provided. In one preferred embodiment, the first step of the method is to conduct an optimized digital PCR

(dPCR) reaction of the sample to amplify the target sequences. In a second step, a HRMA is conducted to obtain the melt curves of the dPCR amplified product. In a third step, the melt curves of the sample are compared to the standard melt curves of the target sequence to identify the particular sequences found in the sample.

A method for the detection of nucleic acid sequences in a heterogeneous sample, comprises: amplifying at least one target sequence individually in a sample through a digital PCR (dPCR) procedure, obtaining the individual melt curves of the at least one individually amplified target sequence in the sample through dHMRA, wherein the dHRMA procedure allows resolution of individual melt curves for each target sequence as opposed to an average melt curve for the sample, and comparing the melt curves in the sample to a set of standard melt curves for the target sequence. The sequences are amplified through broad-based primers or universally ligated adapter sequences. In one embodiment, the sample contains at least two target sequences. In a preferred embodiment, the dPCR procedure comprises amplification by digitizing the sample to have a single molecule per reaction well. In one embodiment of the method the digitizing step comprises diluting the sample to achieve a digital dilution within a digital range of one or two out of three reactions providing a negative result.

A method for identifying infectious pathogens comprises individually amplifying at least one target sequence in a sample containing a heterogeneous mixture of pathogen DNA, wherein said amplification is conducted through dPCR, obtaining the individual melt curves of the at least one target sequence in the sample through dHMRA, wherein the dHRMA procedure allows resolution of individual melt curves for each target sequence as opposed to an average melt curve for the sample, and comparing the melt curves in the sample to a set of standard melt curves for target sequences, wherein the target sequence is a unique sequence of one infectious pathogen. The method can be used to identify one or more pathogens

selected from the group consisting of *S. aureus, S. epidermidis, S. saprophyticus, S. lugdunesis, S. agalactiae, E. faecalis, P. acnes, P. aeruginosa, K. pneumoniae, S. choleraesuis, S. enteritidis,* and *S. dublin* among others. The sequences are amplified through broad-based primers or universally ligated adapter sequences.

A method for miRNA profiling comprises individually amplifying at least one target sequence in a sample containing a heterogeneous mixture of miRNAs, wherein said amplification is conducted through optimized dPCR reactions that enhance specificity of amplification, obtaining the individual melt curves of the at least one target sequence in the sample through dHMRA, wherein the dHRMA procedure allows resolution of individual melt curves for each target sequence as opposed to an average melt curve for the sample, and comparing the melt curves in the sample to a set of standard melt curves for target sequences, wherein the target sequence is a unique sequence of one miRNA.

BRIEF DESCRIPTION OF THE DRAWINGS

The above and other features, aspects, and advantages of the present invention are considered in more detail, in relation to the following description of embodiments thereof shown in the accompanying drawings, in which:

Figure 1 (a) shows a graphical representation of digital High Resolution Melt Analysis (dHRMA). Figure 1 (b) shows a graphical representation of qPCR assays. Figure 1 (c) shows a graphical representation of multiplex qPCR assays. Figure 1 (d) shows a graphical representation of qPCR assays followed by bulk HRMA.

Figure 2 (a) is a graphical representation of the raw HRM data showing the loss of fluorescence in wells containing standard dilutions of four tagged Let-7 miRNA sequences (unmixed). Figure 2 (b) is a derivative plot of raw fluorescence data from two separate experiments containing 4 dilutions each that has been temperature shifted by alignment of the temperature calibrator curve peaks, vertical dotted lines. Figure 2 (c) is graphical

representation of multi-experiment data from Figure 2(b) after normalization and temperature calibration showing highly reproducible, unique melting curves for each tagged miRNA sequence and no template controls are clearly distinguishable (gray curves). Figure 2 (d) shows normalized, temperature calibrated, standard curves from two independent dilution series experiments using each Let-7 family and related miRNA sequence give database references for future dHRMA target identification.

Figure 3 (a) is a graphical representation of the raw HRM data showing the loss of fluorescence in each well of a 96-well plate across which a dilute mixture of Let-7a, b, c, and miR-29 was dispersed. Figure 3 (b) shows normalized, temperature calibrated, and database matched melt curves show universally amplified, single copy detection of each of the four input miRNA (colored melt curves), detection of wells containing multiple copies of targets (light gray, non-matching curves), and amplification negative wells (dark gray melt curves). The number of copies detected is shown in parentheses next to each legend label. Figure 3 (c) is a graphical representation showing confirmation of digital detection accomplished by enumerating and plotting each category of melt curves from the graph in Figure 3 (b) and comparing to the expected Poisson distributions. The experimental occupancy matched a Poisson distribution for $\lambda = 0.65$ copies/reaction. Figure 3 (d) shows standard database curves used to identify target melt curves.

Figure 4 (a) shows normalized and temperature calibrated database melt curves for each of the clinically relevant bacteria listed in Table 2 are resolvable, demonstrating the high sensitivity of HRMA. Figure 4 (b) shows difference curves of each bacteria strain using *S. aureus* as a reference. In previously work which didn't include temperature calibrators or PCR optimal buffers, V6 amplicons of *S. aureus*, *S. epidermidis*, and *S. saprophyticus* were not resolvable by HRMA (14).

Figure 5 (a and b) show derivative plots of HRM fluorescence data before temperature shifting; melting curves for the same species do not accurately overlap. Inset in *a* is an enlarged view of the low temperature calibrator melt curve showing slight differences in melting peak. Figure 5 (c and d) show derivative plots of HRM fluorescence data after temperature shifting showing improved matching and resolution. Four 10-fold dilutions of each of the four species are plotted.

Figure 6 (a) shows raw dHRMA data showing the loss of fluorescence in all but one well of a 96-well plate across which two dilute mixtures of *S. aureus*, *E. faecalis*, and *P. acnes* were dispersed in two separate experiments. Figure 6 (b) shows normalized, temperature calibrated, and database matched digital melt curves after 70 cycles of PCR show broad-based amplified, single copy detection of each of the four input bacteria (colored melt curves), detection of wells containing multiple copies of input or contaminating gDNA templates from Taq polymerase (light gray, non-matching curves), and amplification negative wells (dark gray melt curves). The number of copies detected is shown in parentheses next to each legend label. The corresponding "bulk" well (red curve) where an amount of each target gDNA equivalent to that diluted across the rest of the plate, and also including contaminating Taq gDNA, was assayed by conventional bulk HRMA. Figure 6 (c) is a graphical representation of Poisson distribution matching to the results of broad-based digital detection of polymicrobial input by dHRMA. Enumeration of negatives allows us to calculate that an experimental occupancy of approximately 0.943 was achieved. Figure 6 (d) shows standard database curves used to identify target melt curves.

Figure 7 shows uMELT model predictions of 100 distinct bacterial amplicon melt curves demonstrating the potential for resolving numerous organisms by dHRMA.

DETAILED DESCRIPTION OF THE INVENTION

The invention summarized above may be better understood by referring to the following description. This description of an embodiment, set out below to enable one to practice an implementation of the invention, is not intended to limit the preferred embodiment, but to serve as a particular example thereof. Those skilled in the art should appreciate that they may readily use the conception and specific embodiments disclosed as a basis for modifying or designing other methods and systems for carrying out the same purposes of the present invention. Those skilled in the art should also realize that such equivalent assemblies do not depart from the spirit and scope of the invention in its broadest form.

Sensitive, quantitative, and comprehensive profiling of genetically heterogeneous populations is important for both clinical and basic research applications. Digital high resolution melt analysis (dHRMA), for the rapid, sensitive, quantitative, and broad-based characterization of diverse samples is presented. Nucleic acid molecules of interest are partitioned, as in digital PCR, and amplified using broad-based primers or universally ligated adapter sequences. Each individual, homogeneous reaction undergoes HRMA with generic DNA intercalating dye. Extensive optimization of the assay enables a database of known target melt curves to be used for identification. Absolute quantitation of numerous and even unknown genotypes with single molecule and single nucleotide sensitivity is possible in a single assay. Unlike the methods described in WO2009/134470 (US2013/0217588A1), the present invention further provides for the use of dPCR followed by HRMA to isolate and identify target sequences from heterogeneous mixtures. The combination of dPCR and HRMA provides a digitized HRMA method, or dHRMA, in which the melt curves of each individual target sequence can be determined.

One method of the present invention consists of the following steps. In a first step, the sample is "digitized" via dilutional or microfluidic methods. Digitization partitions target DNA/RNA species into individual single molecule samples. In a second step, the samples are amplified through universal amplification as described in more detail below. In a third step, HRMA of the digitally amplified product is conducted providing individual melt curves for each partitioned molecule. Finally, the melt curves for each individually partitioned molecule are compared to a database of melt curves for identification of the target sequence.

Digitization enables HRMA to accomplish both identification and quantification of many more target molecules than traditional HRMA methods (Figure 1). As shown in Figure 1(a) dHRMA uses a diluted sample input such that either zero or one copy of the target molecule is distributed per reaction well. It is contemplated that other methods may be used to achieve digitation. For example, microfluidic technologies may be used that partition samples into minute reaction volumes, each holding a single target molecule of interest. Broad-based primers amplify all targets individually giving a singular melt curve for each target/reaction and allowing quantification and identification of all species, including unknown or unexpected species, in the mixture. In one preferred embodiment, a "digital level" of amplification is reached by serial dilutions, where each dilution is run through PCR and HRM in triplicate. The digital dilution is identified as the particular dilution that after PCR and HRM gives some positive and some negative signals out of the three replicate reactions. Each tenfold dilution is run in triplicate, and the first dilution where one or two out of three reactions is negative is considered in the digital range. For quantitative digital experiments where a sample is tested across many more reactions (not just three), the level of digitization is determined by quantifying the number of positive and negative reactions and fitting those numbers to a Poisson distribution. Because dHRMA is an extension of dPCR, the same principles apply: the distribution of digital melt curves is governed by stochasticity

and the quantification of species by summation of dHRMA melt curve types can achieve greater precision than qPCR quantification. The number of each type of melt curve relates to the original concentration of individual species in the heterogeneous sample by Poisson statistics.

In addition to digitization, the integration of three key techniques enable the multiplexing and accuracy achieved by the dHRMA assay: (1) the incorporation of broad-based primers or universally ligated priming sites for unbiased amplification of all molecules of interest; (2) shifting the burden of discriminating all amplified sequences from primers and probes to digital melt curves for specificity that relies on the inherent physical properties of the sequence flanked by conserved primer sequences; (3) highly optimized reaction conditions that incorporate tools for normalization to enable database matching. This involves buffer conditioned with an optimized level of positive ions that prevent non-specific binding, temperature calibrating sequences, and computational melt curve normalization that involves temperature shifting based on the calibrator melt peaks, amplification efficiency normalization based on total fluorescence intensity, and background removal as understood by a person of ordinary skill in the art. Under these HRMA conditions, single nucleotide resolution of numerous individual targets is possible. The resulting digital, temperature calibrated, stabilized melt curves are reliably sequence-specific and accurately identifiable by matching to a previously generated database of temperature calibrated melt curves. Traditional bulk HRMA of a heterogeneous sample cannot accomplish the same feat, since each nucleic acid sequence in the mixture will contribute to a single complex melt curve, which is impossible to decouple into individually contributing species when some of the targets in the sample are unknown, e.g. contamination.

The accuracy and reproducibility of dHRMA melt curves relies first on optimal dPCR reaction conditions such that single copies of template are reliably amplified, primer-dimers

and non-specific amplification products are averted, and the reaction is cycled to completion. This precludes false negative or erroneous melt curves in the downstream analysis. Thus primer concentration is minimized and cycling extended during optimization. Primer specificity is also important since dHRMA is highly sensitive to even single nucleotide differences in amplicon sequence, but relying on annealing temperature for control of specificity is risky due to the inherent technical challenges of ensuring uniform heating across all reactions and the need to adapt for each primer set involved. Instead, buffer conditions may be optimized by including ammonium and potassium ions which universally stabilize specific annealing and destabilize non-specific hydrogen bonding respectively. In one preferred embodiment, the final concentrations are the following: Potassium or ammonium salt at 50mM, betaine at 500mM, 10% poly(ethylene glycol) by weight, universal primers at 400nM, 10nM fluorescein, 3.5mM MgCl2, 50nM temperature calibrators, 1X Evagreen dye, 0.2mM dNTP, 0.05U/ul Taq polymerase. This promotes specificity across a wide range of annealing temperatures for any primers. Using universal and broad-based primers ensures that primer specificity is equivalent across all targets so that discrimination of species relies only on the sequence between the conserved priming sites.

Polymicrobial dHRMA reactions required further optimization due to challenges associated with bacterial nucleic acid contamination. PCR reagents often contain background levels of bacterial genomic DNA (gDNA), particularly Taq polymerases generated with recombinant DNA in bacterial cultures. Indeed, approximately 1-1.5 copies of contaminating gDNA per well on average was observed in initial dHRMA experiments. This level of background in the PCR reagents obscured digitization of the target sequences and resulted in complex multi-species melting curves within the majority of the wells. Contamination is reduced using a filtration protocol, as recognized by a person of ordinary skill in the art. In one exemplary embodiment, all PCR reagents except Taq polymerase and target gDNA are

first mixed and then filtered to remove contaminating microbial DNA. A highly purified Taq polymerase with very low background DNA contamination is preferably used. All water used for polymicrobial experiments is treated with DNase and then heat inactivated prior to use. In addition, reduction in reaction volume assists in overcoming contamination problems and facilitates digitization. Reaction volumes are reduced in half, which, in turn, reduces the average level of contamination by half but maintains the same reagent concentrations. These additional steps assist in a more efficient bacterial dHRMA.

"Universal amplification," as described here, means the co-amplification of a heterogeneous nucleic acid sequence population of interest (e.g., polyallelic genetic loci, small or large RNA species, etc.). Universal amplification is accomplished through the use of universal tag sequences. The universal tag sequences, in one embodiment, are primers designed against highly conserved sequences flanking hypervariable regions of the target DNA/RNA. The tag sequences are screened for homology with other non-target sequences and tested for their propensity to form hairpin structures. Structures with low propensity to form such structures are preferred as they would provide a more efficient replication reaction. In one such alternative embodiment, polymicrobial pathogen primers are designed based on common sequences among various pathogens.

In an alternative embodiment, universal tag sequences consist of primers that include ligation-mediated adaptors or multiplexing primers that include built-in label sequences for universal priming. In such alternative embodiment, multiplex ligation dependent probe amplification (also known in the art as MPLA) may be utilized. Ligation-mediated adaptors are small oligonucleotides designed bind to target nucleic acid fragments of varying sequences but which only undergo replication if the complementary adaptors are present. PCR primers, which anneal to the built-in label sequences on the adaptors, are then used to co-amplify all the target fragments. Alternatively, multiplexing primers (e.g. molecular

inversion probes), designed to capture multiple target sequences of interest, have built-in label sequences to allow universal amplification of all captured sequences.

The universal tag sequences are utilized to optimize the reactions for dHRMA. Each universal tag sequence is used in known samples to develop standard melt curves and a database of standard melt curves is assembled. The universal tag sequence standard melt curves or target sequence curves are used to assess the presence of the sequence of interest in a heterogeneous sample in the final step of the method described above.

Once the target sequence curves have been defined, a genomic DNA sample is subjected to the optimized dHRMA reactions for the particular target sequences. The gDNA sample, or any target sample suspected to have the target sequences, is subjected to dPCR in accordance with the optimal conditions for the dHRMA reactions. As described above the sample is first digitized and then subjected to PCR utilizing the target sequence primers. The dPCR reaction amplifies the sequences found in the heterogeneous sample. In a subsequent step, the dPCR reactions are subjected to HRMA generating individual melt curves per sample. The melt curves for the sample are identified utilizing standard scanning techniques. In order to detect the amplified sequences, DNA intercalating agents are added to the solution in some preferred embodiments. It is contemplated that other methods of identifying and measuring the presence of the amplified products in the reactions may be utilized, such as fluorescent or radiographic tags. In a final step, the generated melt curves are compared to the standard melt curves in order to identify the amplified sequences.

The dHRMA method described allows highly sensitive, specific, broad-based detection and discovery of potentially thousands of nucleic acid genotypes in a rapid single assay format that is less expensive and generally more accessible than NGS. This new technique unites and reinvents aspects of limiting dilution/ digital PCR (dPCR) (Vogelstein, B. and Kinzler, K.W. Digital PCR. *Proc. Natl. Acad. Sci.*, 96:9236-9241(1999); Sykes, P.J.,

et al. Quantitation of Targets for PCR by Use of Limiting Dilution. *Biotechniques*, 13:444-449 (1992)), broad-based/universal amplification (Yang, S., et al. Rapid identification of biothreat and other clinically relevant bacterial species by use of universal PCR coupled with high-resolution melting analysis. *J Clin Microbiol*, 47:2252-2255 (2009)), high resolution melt analysis (HRMA) (Dwight, Z., et al. uMELT: prediction of high-resolution melting curves and dynamic melting profiles of PCR products in a rich web application. *Bioinformatics*, 27:1019-1020 (2011); Wittwer, C.T., et al. High-resolution genotyping by amplicon melting analysis using LCGreen. *Clinical Chemistry*, 49:853-860 (2003); Wittwer, C.T., et al. High-resolution DNA melting analysis for simple and efficient molecular diagnostics. *Pharmacogenomics*, 8:597-608 (2007); Wittwer, C.T. High-Resolution DNA Melting Analysis: Advancements and Limitations. *Human Mutation*, 30:857-859 (2009)), and microfluidic theory.

Unlike current microbiological methods, qPCR, and bulk HRMA assays, dHRMA for pathogen detection overcomes sensitivity and specificity issues resulting from contamination by environmental microbes, contamination within PCR reagents, and multi-species polymicrobial infections, which can be problematic to diagnostic assays (Corless, C.E., et al. Contamination and Sensitivity Issues with a Real-Time Universal 16S rRNA PCR. *Journal of Clinical Microbiology*, 38:1747-1752 (2000); Garcia, P. Coagulase-negative staphylococci: clinical, microbiological and molecular features to predict true bacteraemia. *Journal of Medical Microbiology*, 53:67-72 (2004); Weinstein, M.P., et al. The clinical significance of positive blood cultures in the 1990s: A prospective comprehensive evaluation of the microbiology, epidemiology, and outcome of bacteremia and fungemia in adults. *Clinical Infectious Diseases*, 24:584-602 (1997); Spangler, R., et al. Optimizing Taq Polymerase Concentration for Improved Signal-to-Noise in the Broad Range Detection of Low Abundance Bacteria, *PLOS ONE*, 4:e7010 (2009)). The potential of dHRMA to identify and

quantify at the single molecule level in a mixed sample may also, as semi-quantitative studies suggest, further improve clinical diagnosis by helping to distinguishing whether a microbe is a pathogen or contaminant in instances where it could be either.

As shown below in the examples, the feasibility and utility of the approach is demonstrated through application to the concepts of detecting polymicrobial blood infection and profiling microRNA (miRNA) involved in infections. Several groups have proposed broad-based bulk HRMA methods for identification of bacterial species involved in sepsis (Jeng, K., *et al.* Comparative analysis of two broad-range PCR assays for pathogen detection in positive-blood-culture bottles: PCR-high-resolution melting analysis versus PCR-mass spectrometry. *J Clin Microbiol*, 50:3287-3292 (2012); Won, H., et al. Rapid identification of bacterial pathogens in positive blood culture bottles by use of a broad-based PCR assay coupled with high-resolution melt analysis. *Journal of Clinical Microbiology*, 48:3410-3413 (2010); Chakravorty, S., et al. Rapid universal identification of bacterial pathogens from clinical cultures by using a novel sloppy molecular beacon melting temperature signature technique. *Journal of Clinical Microbiology*, 48:258-267(2010); El-Hajj, H.H., et al. Use of sloppy molecular beacon probes for identification of mycobacterial species. *Journal of Clinical Microbiology*, **47**, 1190-1198 (2009)). These methods typically use universal primers to amplify hypervariable regions of the 16S rRNA gene. Subsequently, HRMA is performed using DNA saturating dyes or multiple color probes to generate a melt signature that uniquely identifies each specific bacterial species. Even though a generally reproducible melting signature was obtained when each of the > 50 bacterial species was measured independently, the species could not be individually identified when multiple species were present in the sample simultaneously or when contaminating bacteria were present, as often occurs clinically.

Additionally, dHRMA can be used for profiling circulating miRNA in clinical samples and monitoring host immune response. miRNA are short (19-22nts) non-coding RNA that interact with messenger RNA to regulate gene expression. Circulating miRNA, released by cells into the blood are ideal candidates for quantitative profiling by the highly sensitive dHRMA assay. Quantitation, despite their low level presence, is possible via digitization, and their short lengths give greater melt curve shifts among different sequences for accurate resolution of single nucleotide differences.

The absolute quantification and identification of numerous target genotypes, including discovery of unexpected or unknown species, in a heterogeneous sample using a generic reporter dye, such as EvaGreen (Sigma), is possible by dHRMA. This technique uses limiting dilution digitization to partition target nucleic acids across many reactions allowing discretized HRMA, where each sample molecule is represented by a single and specific melting profile. Broad-based primers or universal tagging allows unbiased amplification of all nucleic acids of interest. Sample-to-answer is achieved with a single assay in a few hours, circumventing the costly, lengthy, and multi-step process of NGS. Single nucleotide resolution, single molecule sensitivity, and broad-based multiplexing offer improvements to traditional microarray and qPCR profiling.

In the exemplary embodiment described in this application, the dHRMA approach is used for sensitive, quantitative, and accurate detection in mock heterogeneous samples containing multiple pathogens and contaminants commonly involved in polymicrobial sepsis (2) or members of the Lethal-7 (Let-7) family of closely related host miRNA known to be key infection-related biomarkers (Schulte, L.N., et al. Analysis of the host microRNA response to Salmonella uncovers the control of major cytokines by the let-7 family. *EMBO J*, 30:1977-1989 (2011)). This technology can identify many more microbes or miRNAs. Strategies to further expand the number of targets identifiable by dHRMA include

incorporating multiple broad-based primer sets and unlabeled probes or generating longer amplicons (Figure 7).

The dHRMA curve matching method allows distinction of single sequences in polymicrobial and miRNA experiments (Figures 3B and 6B). In some embodiments, absolute quantities of each sequence are resolved by expanding the number of reactions across which the sample is diluted. More reactions assist in ensuring that an entire sample can be processed, that individual species are identified separately, and that the likelihood of two or more sequences occupying the same well is extremely low (Table 3). The precision of digital quantification is also improved by increasing the number of reactions Dube, S., et al. Mathematical Analysis of Copy Number Variation in a DNA Sample Using Digital PCR on a Nanofluidic Device. *PLOS ONE*, 3 (2008). Increasing the number and reducing the volumes of digital reactions results in improved resolution in the presence of contaminants, higher content, higher throughput, and reduced reagent costs. High-throughput droplet microfluidic technologies (Heyries, K.A., et al. Megapixel digital PCR. *Nat Methods*, 8:649-651 (2011); Rane, T.D., et al. Counting single molecules in sub-nanolitre droplets. *Lab on a Chip*, 10:161-164 (2010); Zhang, Y., et al. Detecting Genetic Variations in a Droplet. *Proc. 15th International Conference on Miniaturized Chemical and Biochemical Analysis Systems*, 1179-1181 (2011)) that incorporate simultaneous highly controlled heating and sensitive fluorescence detection for millions of reactions are used. In a real heterogeneous sample where unknown sequences are expected and starting concentrations of targets may be unknown, millions of broad-based dHRMA reactions will ensure enough dynamic range for successful single molecule detections (Table 3). Also, characterization of the contaminants in Taq and other reagents by dHRMA database building in accordance with the method described herein will allow non-relevant targets to be rigorously identified and excluded from profiling analysis.

Clinical microbiology currently relies on lengthy culture-based assays to diagnose infections leading to sepsis, which has a high mortality rate that continuously increases with every hour of inappropriate treatment. Generally, immediate conservative treatment with broad-spectrum intravenous antibiotic therapy is initiated without any diagnostic information leading to inaccurate and overtreatment as well as misuse of multiple antibiotics giving rise to the emergence of drug resistant pathogens. The ability of dHRMA to quantify even low level targets within hours may prove especially useful for diagnosing sepsis, especially polymicrobial sepsis which is associated with an even higher mortality rate. The participation of other Gram positive and Gram negative bacteria, anaerobic bacteria, fungi, or novel and emerging pathogens in infection can demand tailored treatments, making fast and accurate broad-based profiling to identify and quantify microbial loads highly desirable. Likewise, the ability to quantify the heterogeneity of species present in a sample may aid clinicians in identifying true sepsis versus contamination in an otherwise ambiguous sample where some bacteria can function in either category, e.g. coagulase negative staphylococci. Biomarker research may also significantly benefit from this technology. For miRNA profiling, the need for highly specific, quantitative, and all-encompassing profiling that is quickly and easily obtained has not been met by any singular technology. The ability of a single platform to accomplish all of these tasks increases the interpretability and reproducibility of findings in the field of nucleic acid identification.

In one embodiment of the present invention, a method for building a dHRMA curve standard library is provided. A standard library of melt curves is developed through the digitization process described herein. The melt curves of known sequences are identified and stored for use as a reference library. In one step of the method for identifying unknown samples, a mixture of known target molecules is prepared for dHRMA by diluting the sample to achieve a single molecule per reaction configuration in a multiwell plate. The known

targets are then amplified utilizing dPCR. The amplified targets are then subject to HRMA and the melt curves are identified. The melt curve standard library identified can then be utilized to evaluate heterogeneous samples containing unknown mixtures of DNA.

Examples

A database of twelve clinically relevant bacteria was created. Out of a spiked mixture, *S. aureus*, *E. faecalis*, and *P. acnes* were identified at the single cell level by dHRMA using broad-based 16s rRNA primers. A database of the Lethal-7 (Let-7) family and two related microRNA was also generated. Out of a spiked mixture, single copies of Let-7a, Let-7b, Let-7c, and miR-29 were identified by dHRMA using universal primers.

MATERIAL AND METHODS

miRNA Universal Tag Design

Multiple short tag sequences were generated following basic primer design rules. Each tag pair was entered into NCBI-BLAST (Altschul, S.F., et al. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res*, 25:3389-3402 (1997)) to screen for homology with other human or bacterial sequences, then tested using OligoCalc (Kibbe, W.A. OligoCalc: an online oligonucleotide properties calculator. *Nucleic Acids Research*, 35:W43-W46 (2007)) to determine the propensity for each oligonucleotide to form hairpins and self-anneal, as well as the degree of 3' end complementarity. The tag primer sequences used for these experiments are: Tag-F (5'-CCATAGACGTAGCAACGATCG-3')(SEQ ID No. 1) and Tag-R (5'-GATGCAAGGACT ATCCACTCAC-3')(SEQ ID No. 2). For this study, tagged cDNA corresponding to the ten miRNA sequences in Table 1 were synthesized by Integrated DNA technologies (IDT, Coralville, IA).

Polymicrobial Primer Design

An alignment was performed using BioEdit (Ibis Biosciences) of 16S rRNA gene sequences for the pathogens listed in Table 2. The V6 region was chosen to discriminate these pathogens based on uMELT model predictions of amplicon melt curves. The V6 primers used were designed previously by Yang et.al. Their sequences are: V6-F (5'-GGAGCATGTGGTTTAATTCGA-3') (SEQ ID No. 3) and V6-R (5'-AGCTGACGACANC CATGCA-3') (SEQ ID No. 4). Primers were synthesized by Integrated DNA technologies (IDT).

Nucleic Acid Extraction

DNA was extracted from clinically isolated or American Type Culture collection (ATCC) acquired bacterial strains:(*S. lugdunensis, S. enteritidis, S. aureus, S. choleraesuis, S. saprophyticus, S. epidermidis, S. Dublin, K. pneumoniae, E. faecalis, P. acnes, P. aeruginosa, S. agalactiae)* using a Roche MagNA Pure LC (Roche Diagnostics, Indianapolis, IN) with the DNA Isolation Kit I (Roche Diagnostics) using a 200 µl input volume and a 100 µl final elution volume per manufacturer's instructions.

dHRMA Reaction Design

General dHRMA Optimization

Primer concentration, annealing temperature concentration as well as ammonium and potassium concentration (Blanchard, M.M., et al. PCR buffer optimization with uniform temperature regimen to facilitate automation, *PCR Methods Appl*, 2:234-240 (1993)) were all optimized by evaluating various conditions until reactions could be performed to reliably amplify single copies of template and give reproducible dHRMA curves. The incorporation of temperature calibrator sequences to the reactions was performed and optimized in order to overcome well-to-well and experiment-to-experiment variations in heating during melt curve generation. Temperature calibrator duplexes, blocked to amplification by 3' end modification, were described previously (Gundry, C.N., et al. Base-pair neutral homozygotes

can be discriminated by calibrated high-resolution melting of small amplicons. *Nucleic Acids Res*, 36:3401-3408(2008))(34). Their sequences are: Low (5'-TTAAATTATAAAATATTTA TAATATTAATTATATATATATAAATATAATA/3SpC3/-3') (SEQ ID No. 5) High (5'-GCGCGGCCGGCACTGACCCGAGACTCTGAGCGGCTGCTGGAGGTGCGGAAGCGG AGGGGCGGG/3SpC3/-3') (SEQ ID No. 6). White 96-well plates with black semi-skirting were used to maximize detection of fluorescence signal and minimize well-to-well fluorescence cross-talk (Eppendorf).

miRNA Experiments

Optimized dHRMA miRNA reactions were performed as follows: 10ul total reaction volume consisting of 1X PCR buffer (Qiagen), 10nM fluorecene (Bio-rad), 3.5mM total MgCl2 (Qiagen), 400nM of each tag primer (IDT), 50nM temperature calibrator sequences (IDT), 1X EvaGreen (Sigma), 200uM dNTP (Invitrogen), 0.05U/ul HotStart Taq polymerase (Qiagen), 2ul gDNA dilution, and ultrapure water (Quality Biological), with a 15ul overlay of high quality DNase-free mineral oil (Sigma). Thermocycling proceeded as follows: hold-(95°C 5min), cycle 60 times-(95°C 30sec, 59°C 30sec), and cycle 1 time-(95°C 30sec, 25°C hold). Single copy amplification by tag primers clustered around a PCR cycle threshold (Ct) of 45.

Polymicrobial Experiments

Further optimization experiments and reduction in reaction volume combined with a previous filtration protocol (Yang, S., *et al.* Rapid PCR-based diagnosis of septic arthritis by early Gram-type classification and pathogen identification. *J Clin Microbiol*, 46:1386-1390 (2008)) resulted in the final reaction conditions of: 5ul total reaction volume consisting of 1X PCR buffer (Qiagen), 10nM fluorecene (Bio-rad), 3.5mM total MgCl2 (Qiagen), 150nM of each V6 primer (IDT), 50nM temperature calibrator sequences (IDT), 1X EvaGreen (Sigma), 200uM dNTP (Invitrogen), 0.05U/ul AmpliTaq Gold Low DNA Taq polymerase

(Invitrogen), 2ul genomic DNA (gDNA) dilution, and ultrapure water (Quality Biological), with a 20ul overlay of high quality DNase-free mineral oil (Sigma).

For each bacterial gDNA template (Table1), standard 10-fold dilutions of template gDNA in ultrapure, DNase treated water ranging from approximately 10^6 to 0 copies per PCR reaction were run using a Biorad MyIQ qPCR machine. The thermocycling conditions that achieved reliable single copy amplification near a Ct of 45 were chosen. To ensure amplification completion, 70 cycles were run. The optimized thermocycling conditions were: hold-(95°C 5min), cycle 70 times-(95°C 30sec, 65°C 30sec, 72°C 30sec), and cycle 1 time-(95°C 30sec, 25°C hold).

High Resolution Melt Analysis

Directly after dPCR, dHRMA was performed on the 96-well digital plate with LightScanner equipment (BioFire Diagnostics, Salt Lake City, UT) using a temperature range from 55 to 95°C. Analysis was accomplished using the LightScanner software with Call-IT 2.5 small amplicon genotyping algorithm, which incorporates temperature shifting and normalization using the low and high temperature calibrators. The multi-plate analysis tool was used to match standard database curves to experimental curves and identify the templates.

Quantitation of Species

For each type of positively identified melt curve grouped by the LightScanner software, the sum of matching curves was calculated along with the number of negative wells. Using these values, the concentration of each type of target molecule is calculated by Poisson statistics. The Poisson probability (P) of occupancy (λ) in any well, i.e. the fraction of wells having 0, 1, etc. copies of target, is given by $P= (\lambda^n/n!) * e^{(-\lambda)}$, where *n* is the total number of wells. The overall average occupancy is given by the sum of the occupancy of each species $\lambda_{overall} = \lambda_1 + \lambda_2 + \ldots \lambda_n$.

RESULTS

dHRMA for miRNA Profiling

Here we demonstrate that the most difficult miRNA to correctly identify by current microarray and qPCR methods (Lee, I., et al. Discriminating single-base difference miRNA expressions using microarray Probe Design Guru (ProDeG). *Nucleic Acids Res*, 36:e27 (2008)), members of the Let-7 family which differ by only one to four nucleotides in sequence (Table 1, differences underlined), can be identified by universal dHRMA using DNA intercalating dye. These miRNA are of particular interest, since their down-regulation is directly implicated in the activation of host inflammatory cytokines during microbial infection (28). Likewise, miRNA having drastically different sequences can also be identified (e.g. miR-29 and miR-98, Table 1). To accomplish unbiased, universal amplification, Let-7a, b, c, d, e, f, g, i, miR-98 and miR-29 sequences were tagged for universal priming. In database generation experiments each tagged cDNA sequence was serially diluted down to the digital level, universally amplified with tag primers, and HRMA was performed on each homogeneous reaction product. Figure 2A shows the raw fluorescence melt data for Let-7a, Let-7b, Let-7c, and miR-29. A derivative plot of the fluorescence data was generated and alignment and normalization according to the low and high temperature calibrator sequences was performed (Figure 2B). Figure 2 C shows the calibrated and normalized melt data as a percentage of the highest fluorescence, i.e. when amplicons are fully annealed in a helical structure. Wells negative for amplification are easily distinguishable (gray lines, Figure 2C). Optimization of reaction conditions resulted in highly reproducible melt curves for each sequence in Table 1, and these were readily distinguishable using the LightScanner's small amplicon genotyping algorithm (Figure 2D). The melting curves in Figure 2 were collected over multiple dilutions and multiple days of experimentation, demonstrating the reliability of the optimized assay. Based on known amounts of Let-7a, b, c, and miR-29, a spiked mixture

was prepared then digitized, universally amplified, and dHRMA analysis was performed. The calculated input gave a theoretical occupancy, $\lambda_{overall} = 0.4$ copies of miRNA. Figure 3 shows the results of universal dHRMA for the mixture of miRNA. Normalized melt curves were reliably matched to a previously generated database of temperature calibrated melt curves for each miRNA (Figure 3B). From the mixture, 14 Let-7a miRNA were detected, along with 12 Let-7b, 2 Let-7c, and 8 miR-29 sequences. Of 96 reactions, 11 were unidentified, meaning the melt curve resulted from a combination of two or more of the input sequences and 49 reactions were negative. Digitization was confirmed by fitting these values to a Poisson distribution. Figure 3C shows the Poisson distribution for different values of λ, including that expected for $\lambda = 0.4$. Experimental quantification gave a Poisson distribution with $\lambda_{overall} = 0.65$ instead of the calculated 0.4 (Figure 3C). The overall concentration of the input mixture is then (0.65 copies/reaction) / (2ul input mixture/reaction) = 0.325 copies/ul of miRNA. With an increased number of digital reactions, absolute quantitation of each miRNA species can be achieved.

dHRMA for Diagnosis of Polymicrobial Sepsis

Here we demonstrate that dHRMA with broad-based bacterial 16s rDNA primers targeting the hypervariable V6 region and DNA intercalating dye can resolve each bacteria within a mixture of pathogens involved in polymicrobial sepsis and common clinical contaminants (Table 2). Melt curves were generated from standard dilutions reaching the digital level to ensure accurate curves for database creation. Calibrated and normalized helicity and difference curves of twelve clinically relevant bacteria were experimentally generated for the database (Figure 4). Highly pure and concentrated gDNA was used to generate standard melt curves, as any cross-contamination by other bacterial DNA would be amplified by the universal primers and contribute to changes in the melt curves. Difference curves of *S. aureus, S. epidermidis, and S. saprophyticus* V6 amplicons were previously

unresolvable. Surprisingly, under the currently optimized dHRMA conditions such curves are highly reproducible and discernible using only intercalating dye as the reporter (Figure 5).

Next, a spiked mixture of *S. aureus*, *E. faecalis*, and *P. acnes* was prepared, digitized, amplified by V6 broad-based primers, and assessed by dHRMA. Figure 6 contrasts a bulk HRMA experiment (red melt curve) and the dHRMA results (all other melt curves) of the polymicrobial mixture. Input concentrations were adjusted such that the same amount of the polymicrobial gDNA was added to the bulk well as was distributed across the digital wells. Four curves were positively identified by matching to the database curves (Figure 6D). These likely amplified from singular target templates. Knowing the number of negative, 37, and total positive, 58, reactions, a Poisson calculation gives $\lambda_{overall} = 0.94$ for the reaction mixture (Figure 6C). By this calculation, approximately 31 of the unidentified melt curves represent single copies of unknown gDNA templates, though not necessarily all distinct from one another. These may have originated either from the PCR reagents themselves or potentially from very low level contaminants amplified by culture and carried over into the spiked gDNA input. Estimates of Taq contamination as stated by the manufacturer are ~10 copies/ul Taq, which gives a $\lambda_{PCR\ contaminants} \sim 0.5$ for current polymicrobial dHRMA assay. The original input mixture of *S. aureus*, *E. faecalis*, and *P. acnes* gDNA then has $\lambda_{input} \sim 0.44$. With an increased number of digital reactions, absolute quantitation of each bacterial species can be achieved.

Table 1 Shows the differences in nucleotide sequences for the Lethal-7 family and related miRNA. The differences are underlined.

Name	Sequence	Difference	SEQ ID No.
let-7a	TGAGGTAGTAGGTTGTATAGTT	reference	7
let-7b	TGAGGTAGTAGGTTGT***G***T***G***GTT	2 nt	8
let-7c	TGAGGTAGTAGGTTGTAT***G***GTT	1 nt	9
let-7d	***A***GAGGTAGTAGGTTG***C***ATAGT	2 nt	10
let-7e	TGAGGTAG***G***AGGTTGTATAGT	1 nt	11

let-7f	TGAGGTAGTAG***A***TTGTATAGTT	1 nt	12
let-7g	TGAGGTAGTAG***T***TTGTA***C***AGT	2 nt	13
let-7i	TGAGGTAGTAG***T***TTGT***GCT***GT	4 nt	14
miR-98	TGAGGTAGTA***A***GTTGTAT***T***GTT	2 nt	15
miR-29	T***AGCACCA***T***CT***G***AAA***T***CGGT***T***A***	17 nt	16

Table 2 provides a list of clinically relevant bacteria involved in sepsis.

Gram	Genus	Species	Common role in blood culture
+	Staphylococcus	lugdunesis	Contaminant/Emerging Pathogen
+	Staphylococcus	aureus	Pathogen
+	Staphylococcus	saprophyticus	Contaminant/Emerging Pathogen
+	Staphylococcus	epidermidis	Contaminant/Emerging Pathogen
+	Streptococcus	agalactiae	Pathogen
+	Enterococcus	faecalis	Pathogen
+	Propionibacterium	acnes	Contaminant
-	Pseudomonas	aeruginosa	Pathogen
-	Klebsiella	pneumoniae	Pathogen
-	Salmonella	choleraesuis	Pathogen
-	Salmonella	enteritidis	Pathogen
-	Salmonella	dublin	Pathogen

Table 3 shows the dynamic range of dHRMA detection as a function of the number of reactions.

λ	% negative wells	% 1 copy	% 2 copy
0.001	99.90%	0.10%	0.005%
0.01	99.00%	0.99%	0.005%
0.1	90.48%	9.05%	0.452%
0.15	85.07%	12.51%	0.982%
1	36.79%	36.79%	18.39%

Total Reactions	λ	Number of Single Bacteria Resolved	Negative Wells
1,000,000	0.001	1.0E+03	1.0E+06
10,000,000	0.001	1.0E+04	1.0E+07
100,000,000	0.001	1.0E+05	1.0E+08

Total Reactions	λ	Number of Single Bacteria Resolved	Negative Wells
1,000,000	0.001	1.0E+03	1.0E+06
1,000,000	0.01	9.9E+03	9.9E+05
1,000,000	0.1	9.05E+04	9.0E+05
1,000,000	1	3.68E+05	3.7E+05

The invention has been described with references to a preferred embodiment. While specific values, relationships, materials and steps have been set forth for purposes of

describing concepts of the invention, it will be appreciated by persons skilled in the art that numerous variations and/or modifications may be made to the invention as shown in the specific embodiments without departing from the spirit or scope of the basic concepts and operating principles of the invention as broadly described. It should be recognized that, in the light of the above teachings, those skilled in the art can modify those specifics without departing from the invention taught herein. Having now fully set forth the preferred embodiments and certain modifications of the concept underlying the present invention, various other embodiments as well as certain variations and modifications of the embodiments herein shown and described will obviously occur to those skilled in the art upon becoming familiar with such underlying concept. It is intended to include all such modifications, alternatives and other embodiments insofar as they come within the scope of the appended claims or equivalents thereof. It should be understood, therefore, that the invention may be practiced otherwise than as specifically set forth herein. Consequently, the present embodiments are to be considered in all respects as illustrative and not restrictive.

INDUSTRIAL APPLICABILITY

The present invention is applicable to methods for identification of biological molecules. The invention discloses a method for the detection of nucleic acids in heterogeneous samples. The method and devices described herein can be made and practiced in industry in the field of biotechnology.

CLAIMS

What is claimed is:

Claim 1. A method for the detection of nucleic acid sequences in a heterogeneous sample, comprising:

amplifying at least one target sequence individually in a sample through a digital PCR (dPCR) procedure,

obtaining the individual melt curves of the at least one individually amplified target sequence in the sample through dHMRA, wherein the dHRMA procedure allows resolution of individual melt curves for each target sequence as opposed to an average melt curve for the sample, and

comparing the melt curves in the sample to a set of standard melt curves for the target sequence.

Claim 2. The method of claim 1, wherein the sequences are amplified through broad-based primers or universally ligated adapter sequences.

Claim 3. The method of claim 1, wherein the sample contains at least two target sequences.

Claim 4. The method of claim 1, wherein the dPCR procedure comprises amplification by digitizing the sample to have a single molecule per reaction well.

Claim 5. The method of claim 4, wherein the said digitizing of the sample comprises diluting the sample to achieve a digital dilution within a digital range.

Claim 6. The method of claim 5, wherein the digital range consists of one or two out of three reactions providing a negative result.

Claim 7. A method for identifying infectious pathogens, comprising:

individually amplifying at least one target sequence in a sample containing a heterogeneous mixture of pathogen DNA, wherein said amplification is conducted through dPCR,

obtaining the individual melt curves of the at least one target sequence in the sample through dHMRA, wherein the dHRMA procedure allows resolution of individual melt curves for each target sequence as opposed to an average melt curve for the sample, and

comparing the melt curves in the sample to a set of standard melt curves for target sequences, wherein the target sequence is a unique sequence of one infectious pathogen.

Claim 8. The method of Claim 7, wherein the pathogen is selected from the group consisting of *S. aureus, S. epidermidis, S. saprophyticus, S. lugdunesis, S. agalactiae, E. faecalis, P. acnes, P. aeruginosa, K. pneumoniae, S. choleraesuis, S. enteritidis,* and *S. dublin* among others.

Claim 9. The method of claim 7, wherein the sequences are amplified through broad-based primers or universally ligated adapter sequences.

Claim 10. The method of claim 7, wherein the sample contains at least two target sequences.

Claim 11. The method of claim 7, wherein the dPCR procedure comprises amplification by digitizing the sample to have a single molecule per reaction well.

Claim 12. The method of claim 11, wherein the said digitizing of the sample comprises diluting the sample to achieve a digital dilution within a digital range.

Claim 13. The method of claim 12, wherein the digital range consists of one or two out of three reactions providing a negative result.

Claim 14. A method for miRNA profiling, comprising:

individually amplifying at least one target sequence in a sample containing a heterogeneous mixture of miRNAs, wherein said amplification is conducted through optimized dPCR reactions that enhance specificity of amplification,

obtaining the individual melt curves of the at least one target sequence in the sample through dHMRA, wherein the dHRMA procedure allows resolution of individual melt curves for each target sequence as opposed to an average melt curve for the sample, and

comparing the melt curves in the sample to a set of standard melt curves for target sequences, wherein the target sequence is a unique sequence of one miRNA.

Claim 15. The method of claim 14, wherein the sequences are amplified through broad-based primers or universally ligated adapter sequences.

Claim 16. The method of claim 14, wherein the sample contains at least two target sequences.

Claim 17. The method of claim 14, wherein the dPCR procedure comprises amplification by digitizing the sample to have a single molecule per reaction well.

Claim 18. The method of claim 17, wherein the said digitizing of the sample comprises diluting the sample to achieve a digital dilution within a digital range.

Claim 19. The method of claim 18, wherein the digital range consists of one or two out of three reactions providing a negative result.

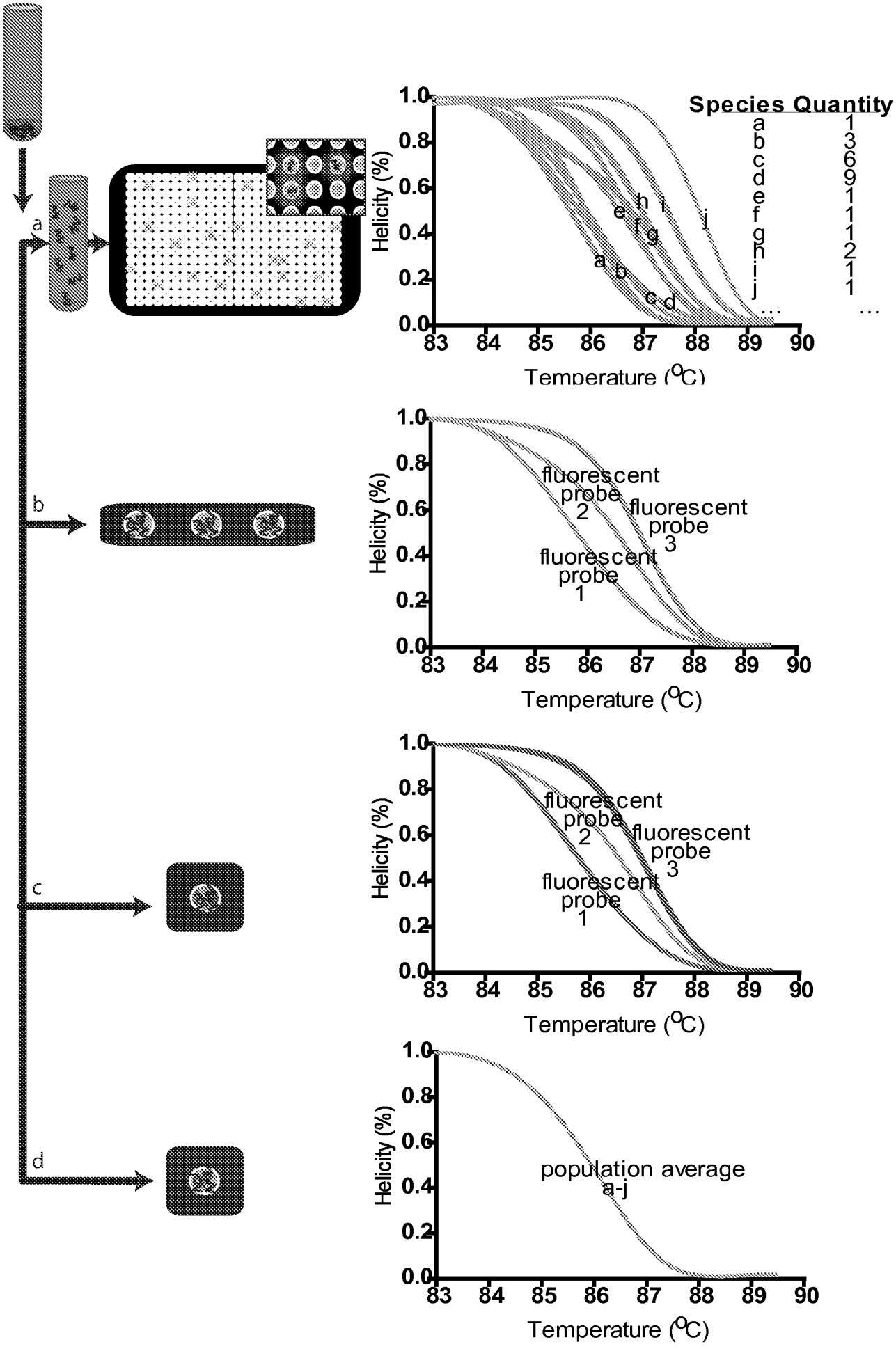

a
b
c
d
Species Quantity
a 1
b 3
c 6
d 9
e 1
f 1
g 1
h 2
i 1
j 1
...
...
Helicity (%)
Temperature (°C)
fluorescent probe 1
fluorescent probe 2
fluorescent probe 3
population average a-j

FIG. 1

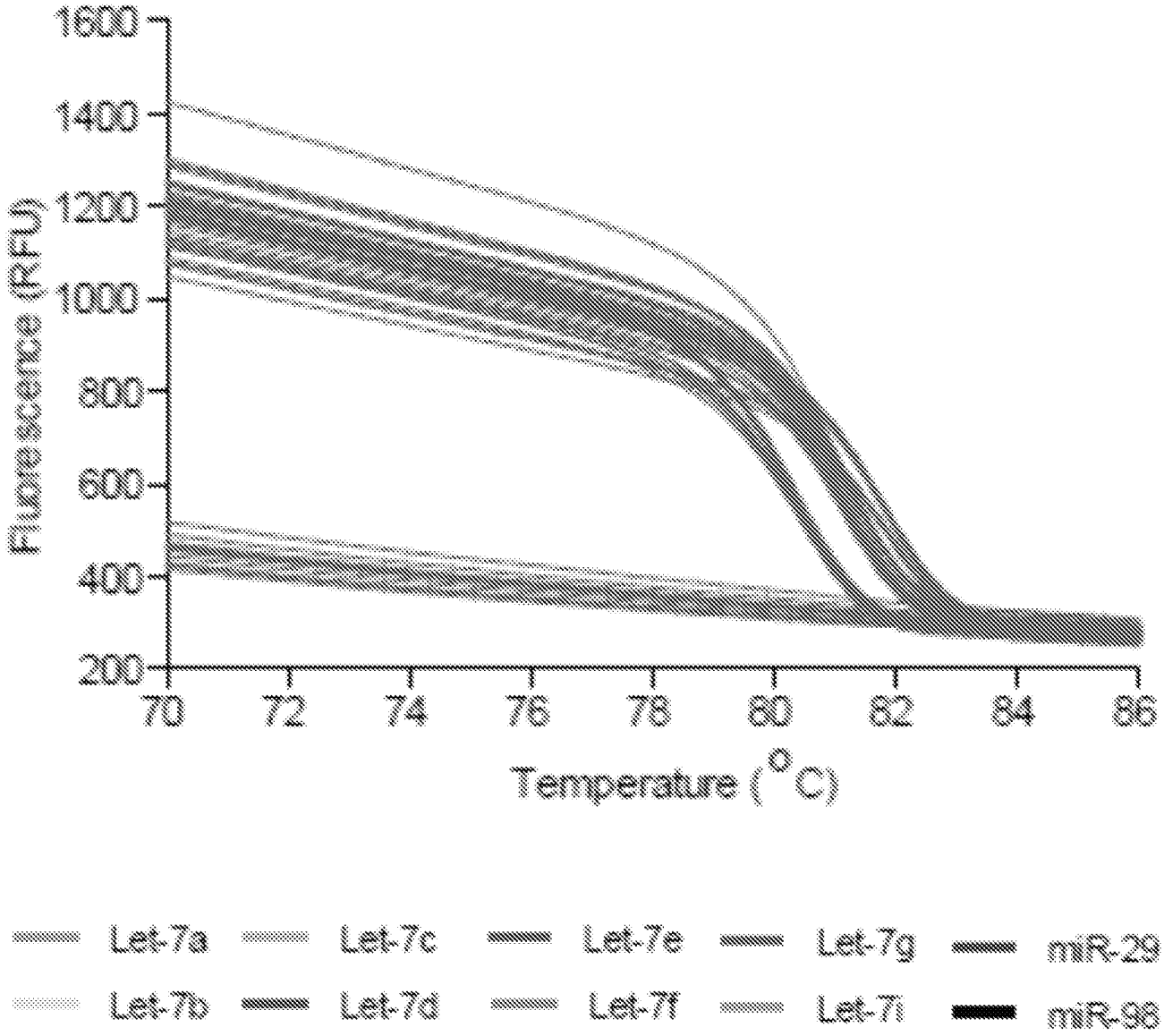

1600
1400
1200
1000
800
600
400
200
Fluorescence (RFU)
70
72
74
76
78
80
82
84
86
Temperature (°C)
Let-7a
Let-7b
Let-7c
Let-7d
Let-7e
Let-7f
Let-7g
Let-7i
miR-29
miR-98

FIG. 2A

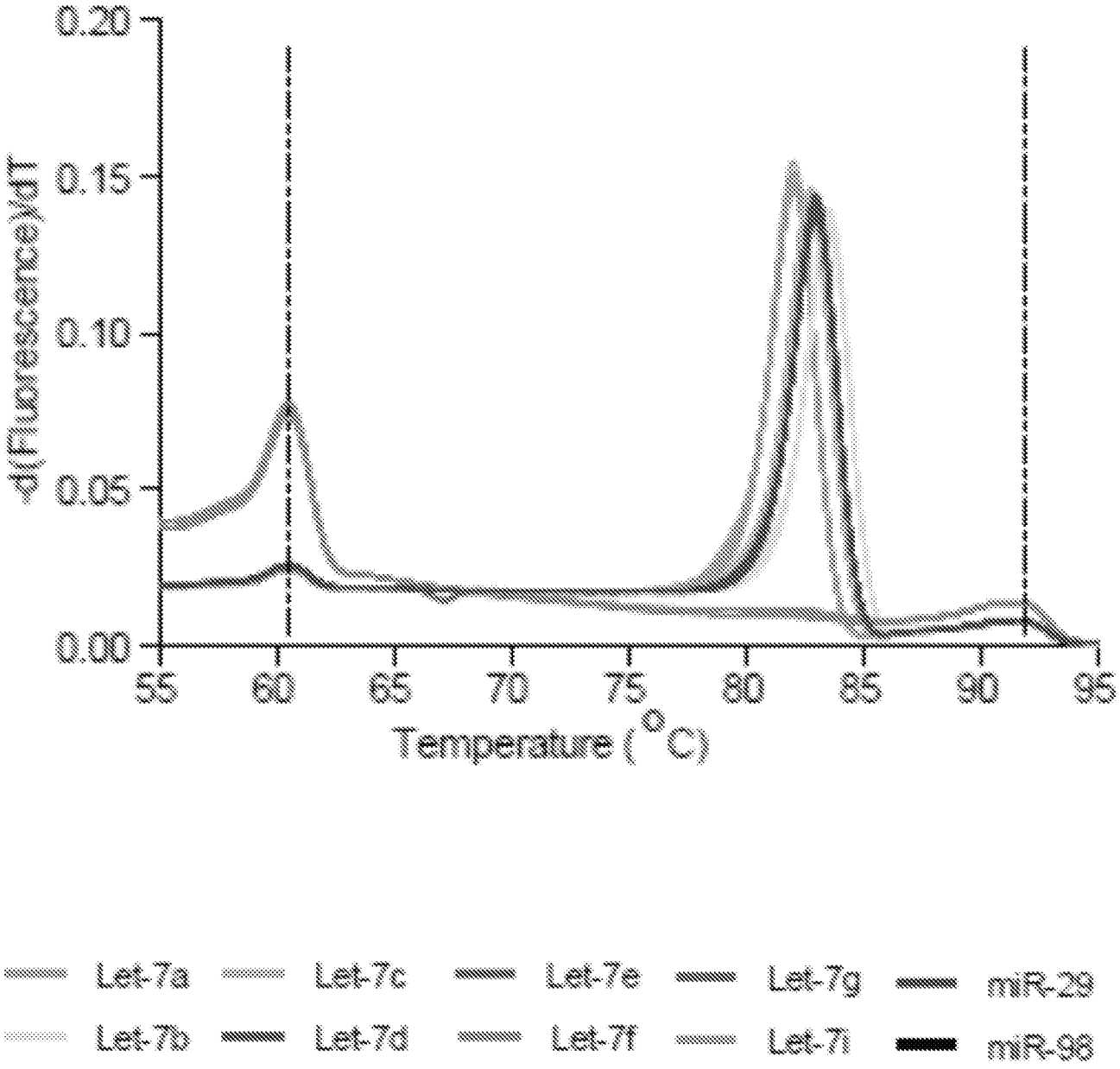

0.20
0.15
0.10
0.05
0.00
-d(Fluorescence)/dT
55
60
65
70
75
80
85
90
95
Temperature (°C)
Let-7a
Let-7b
Let-7c
Let-7d
Let-7e
Let-7f
Let-7g
Let-7i
miR-29
miR-98

FIG. 2B

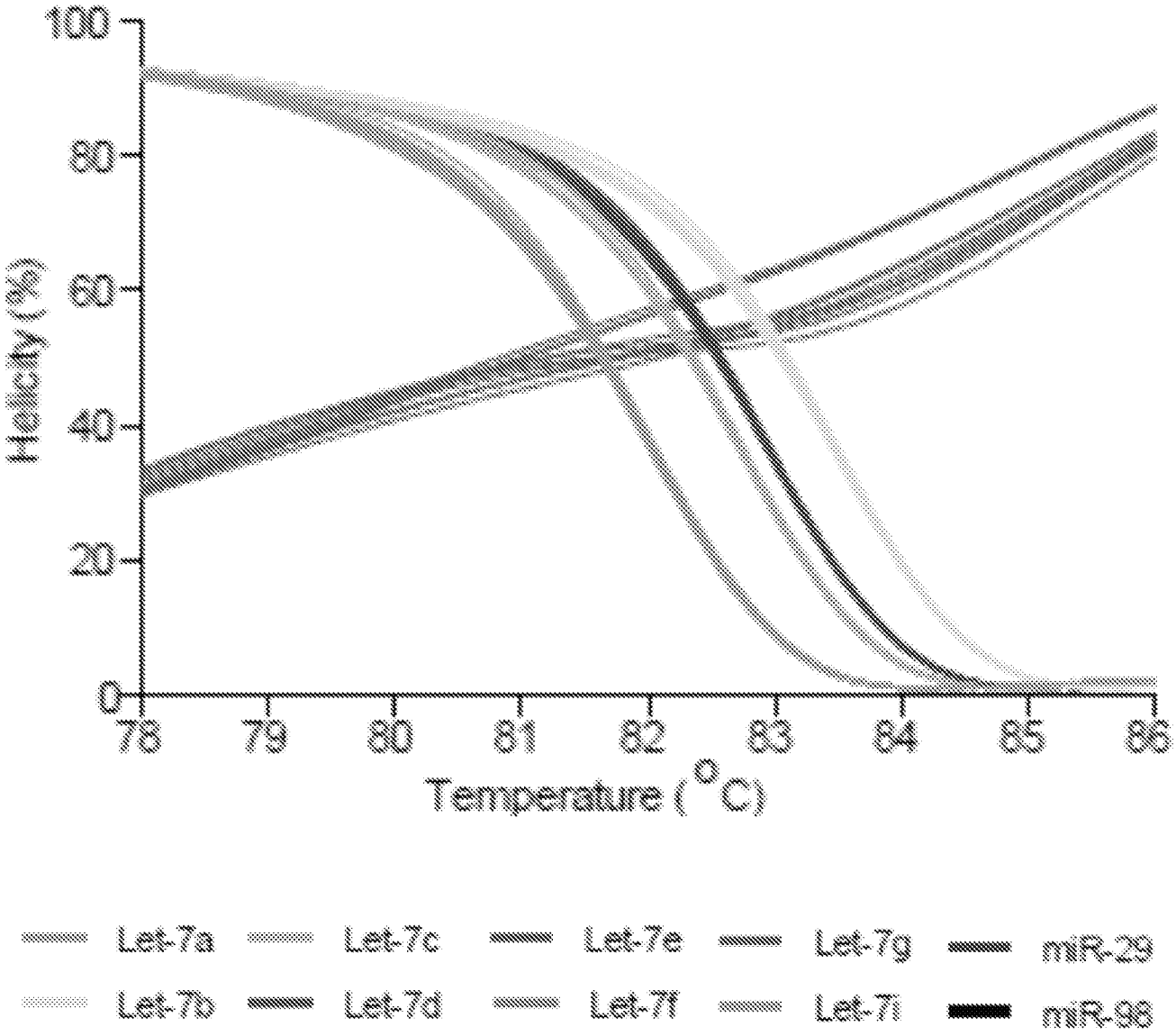

100
80
60
40
20
0
Helicity (%)
78
79
80
81
82
83
84
85
86
Temperature (°C)
Let-7a
Let-7c
Let-7e
Let-7g
miR-29
Let-7b
Let-7d
Let-7f
Let-7i
miR-98

FIG. 2C

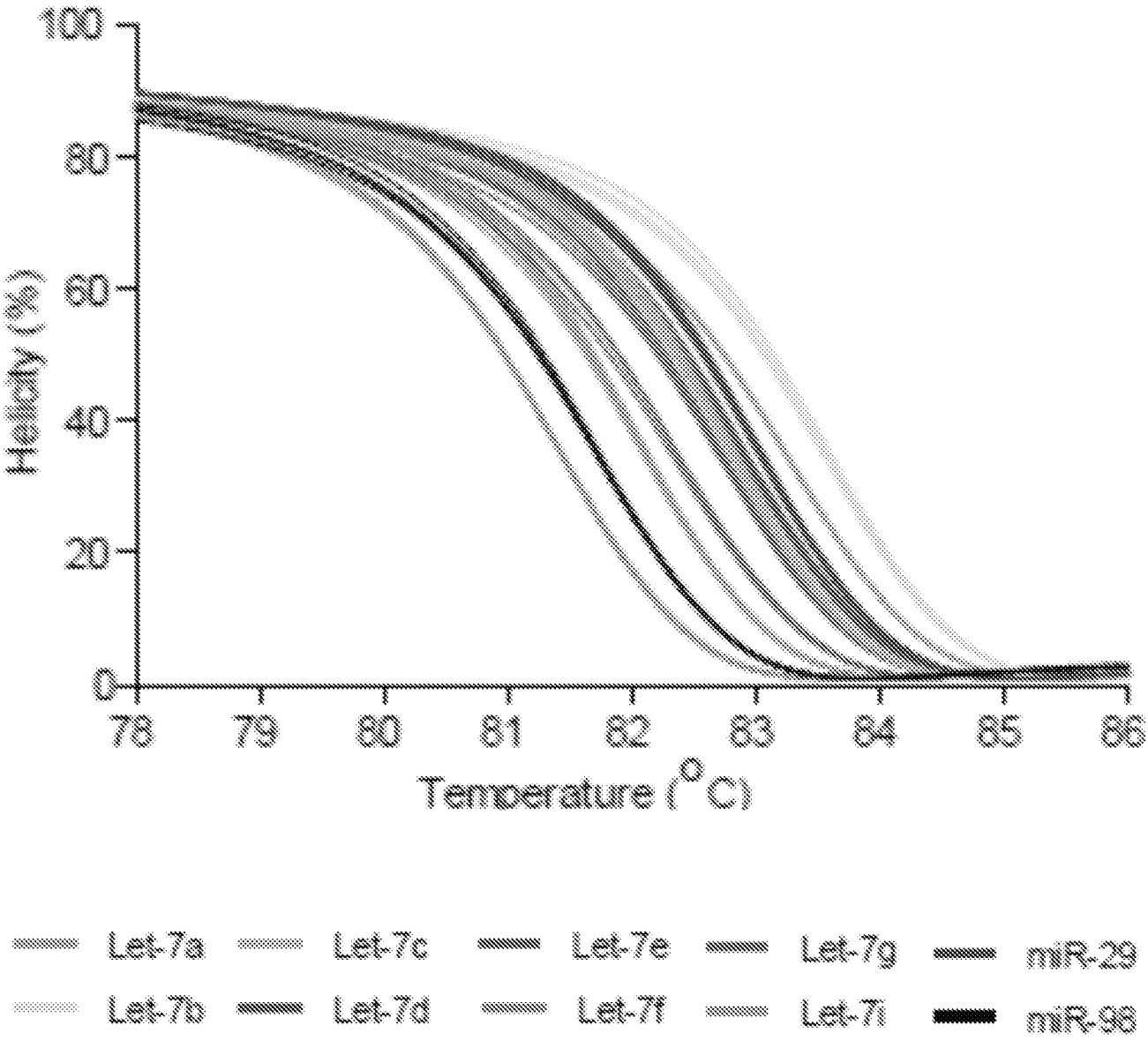

100
80
60
40
20
0
Helicity (%)
78
79
80
81
82
83
84
85
86
Temperature (°C)
Let-7a
Let-7b
Let-7c
Let-7d
Let-7e
Let-7f
Let-7g
Let-7i
miR-29
miR-98

FIG. 2D

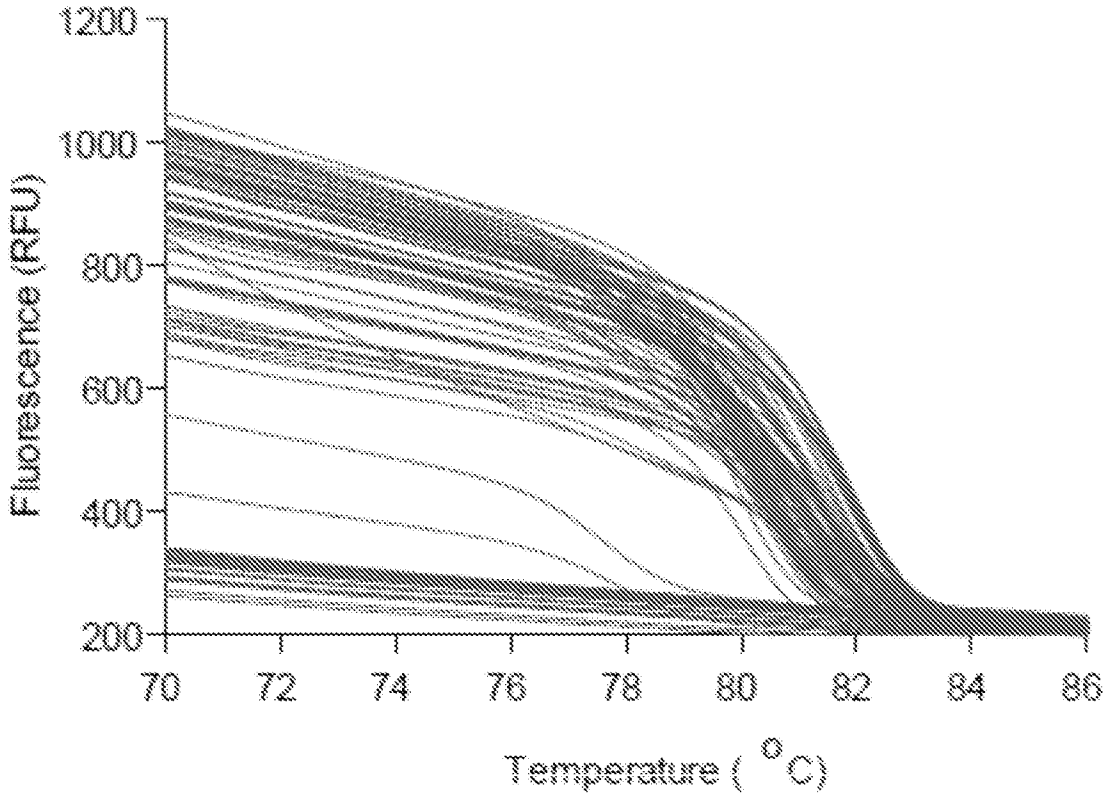

1200
1000
800
600
400
200
Fluorescence (RFU)
70
72
74
76
78
80
82
84
86
Temperature (°C)

FIG. 3A

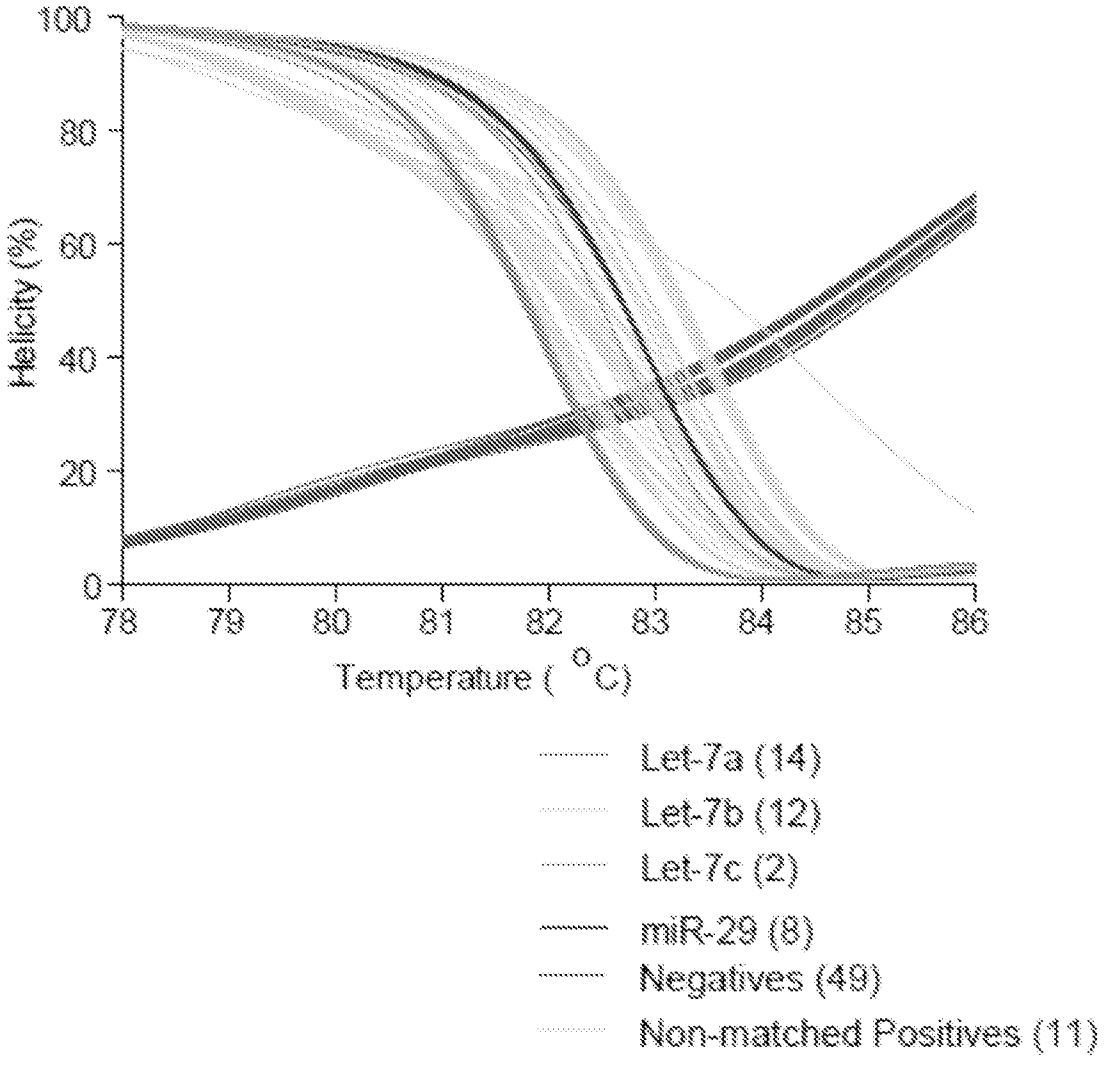

100
80
60
40
20
0
Helicity (%)
78
79
80
81
82
83
84
85
86
Temperature (°C)
Let-7a (14)
Let-7b (12)
Let-7c (2)
miR-29 (8)
Negatives (49)
Non-matched Positives (11)

FIG. 3B

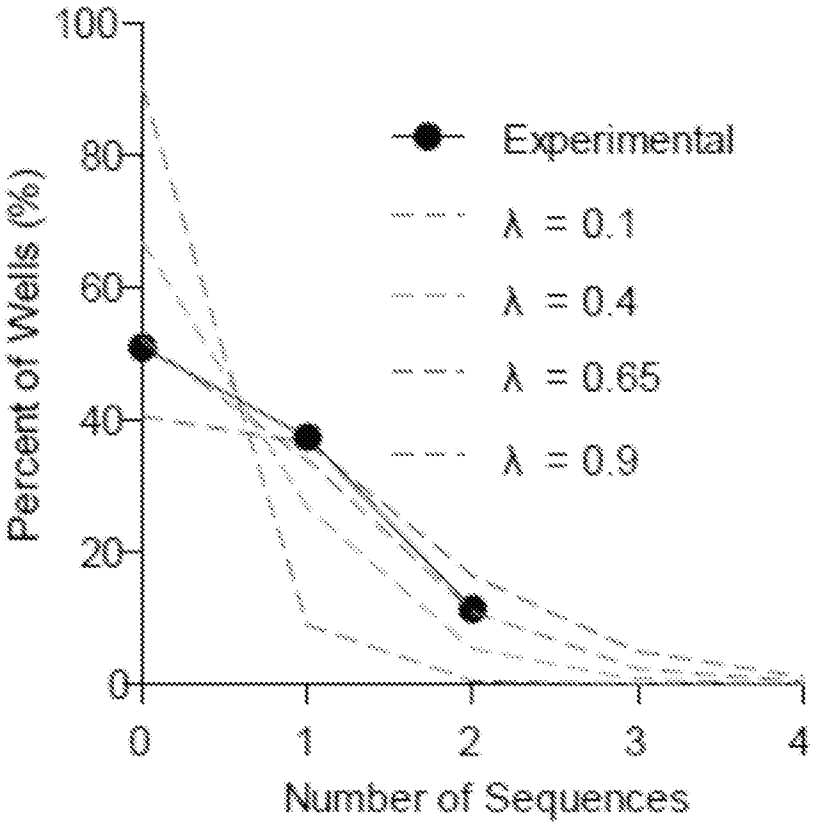

100
80
60
40
20
0
Percent of Wells (%)
Experimental
λ = 0.1
λ = 0.4
λ = 0.65
λ = 0.9
0
1
2
3
4
Number of Sequences

FIG. 3C

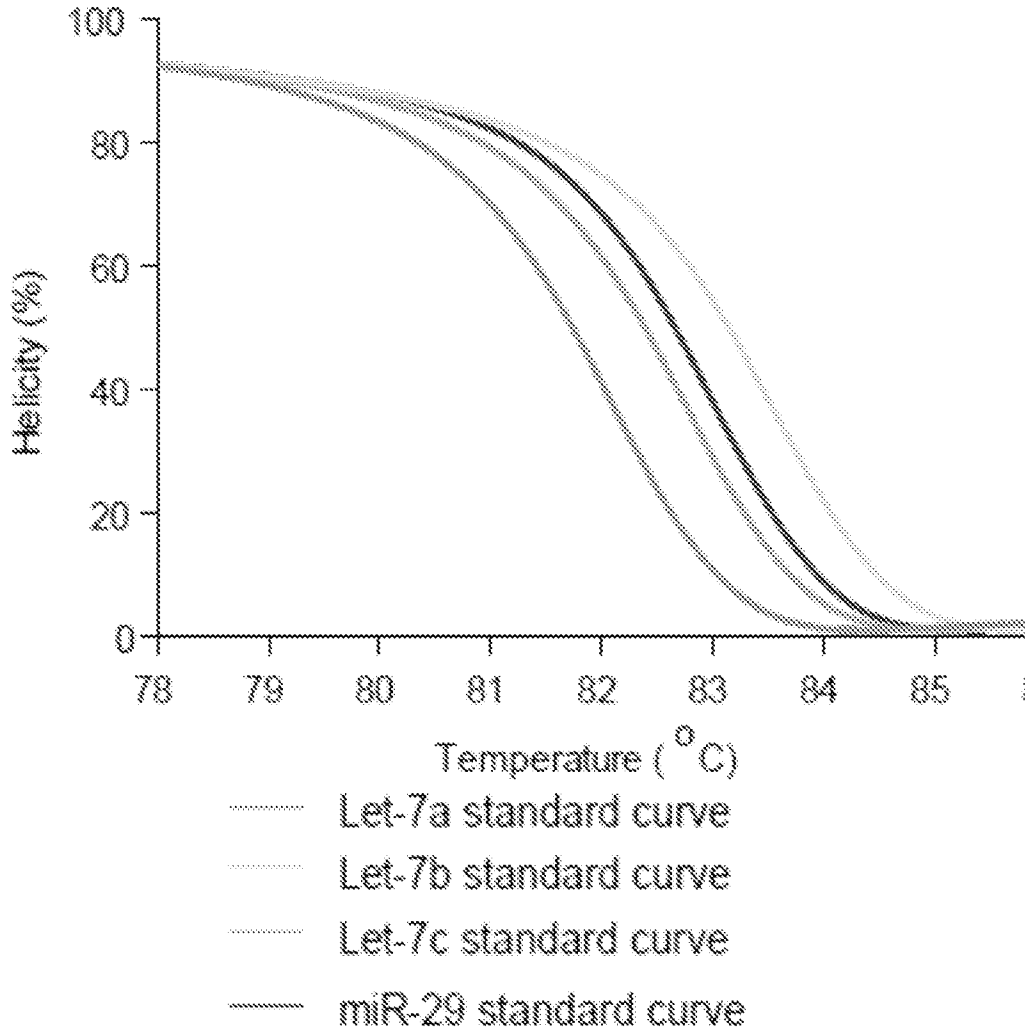

100
80
60
40
20
0
Helicity (%)
78
79
80
81
82
83
84
85
Temperature (°C)
Let-7a standard curve
Let-7b standard curve
Let-7c standard curve
miR-29 standard curve

FIG. 3D

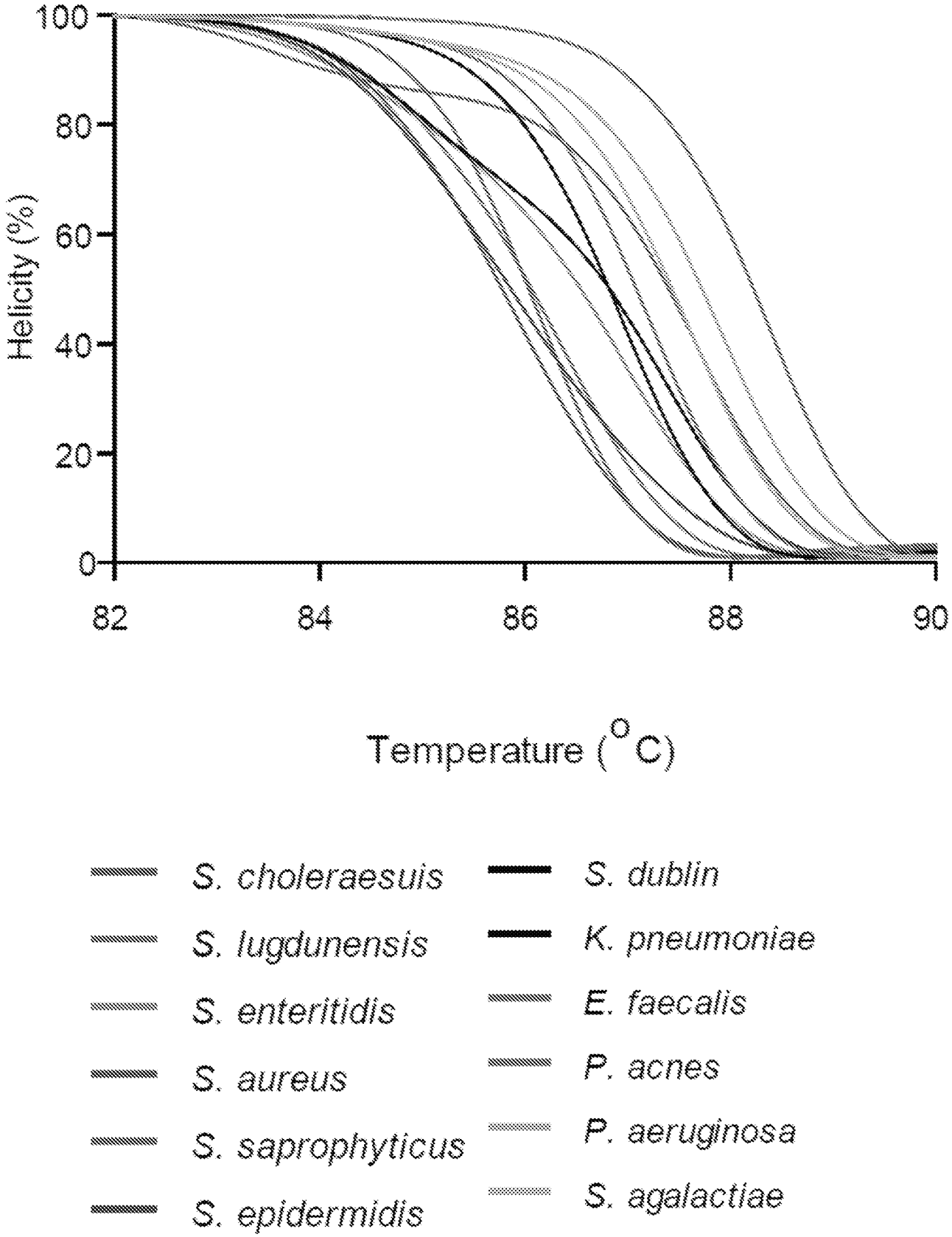
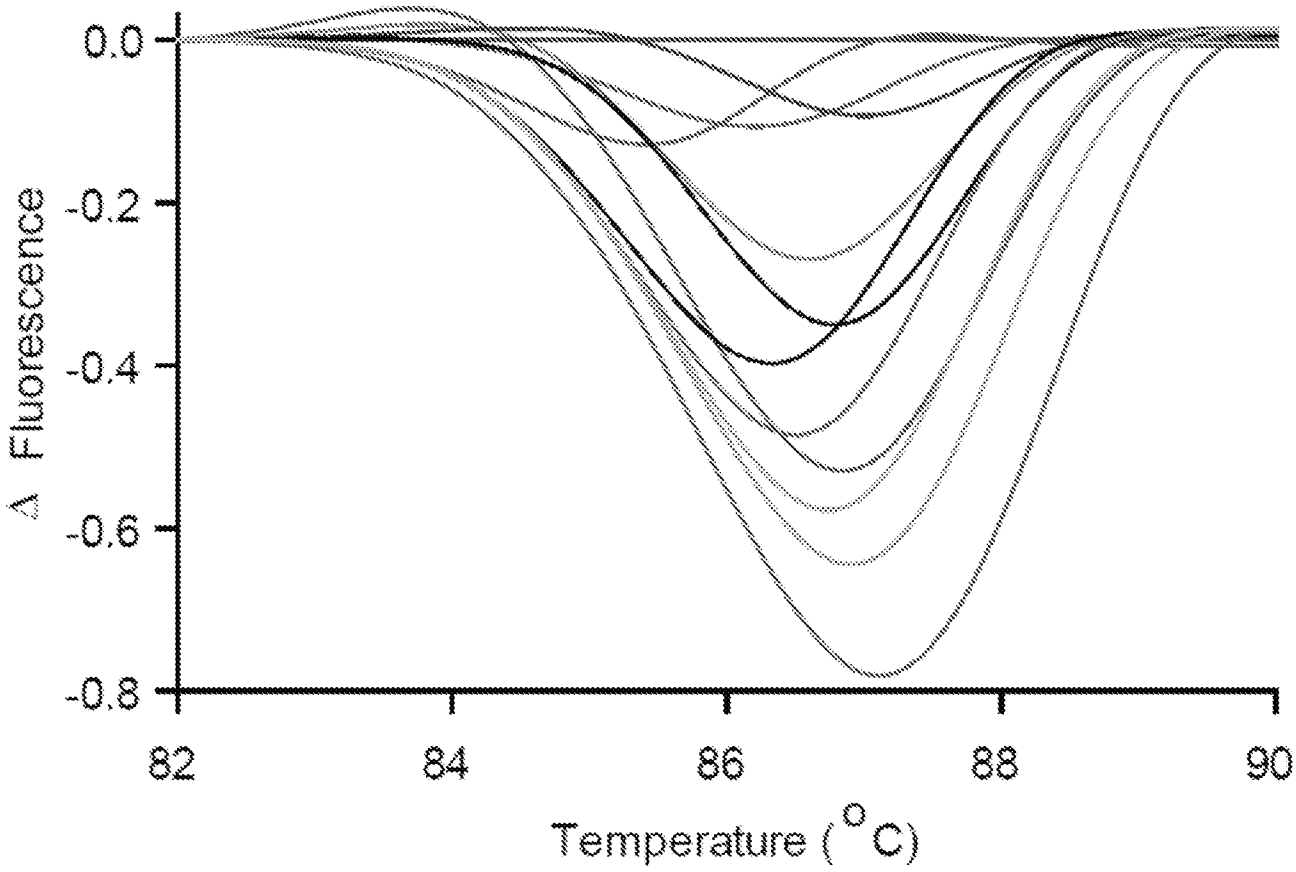

100
80
60
40
20
0
Helicity (%)
82
84
86
88
90
Temperature (°C)
S. choleraesuis
S. lugdunensis
S. enteritidis
S. aureus
S. saprophyticus
S. epidermidis
S. dublin
K. pneumoniae
E. faecalis
P. acnes
P. aeruginosa
S. agalactiae

FIG. 4A



S. choleraesuis
S. lugdunensis
S. enteritidis
S. aureus
S. saprophyticus
S. epidermidis
S. dublin
K. pneumoniae
E. faecalis
P. acnes
P. aeruginosa
S. agalactiae

FIG. 4B

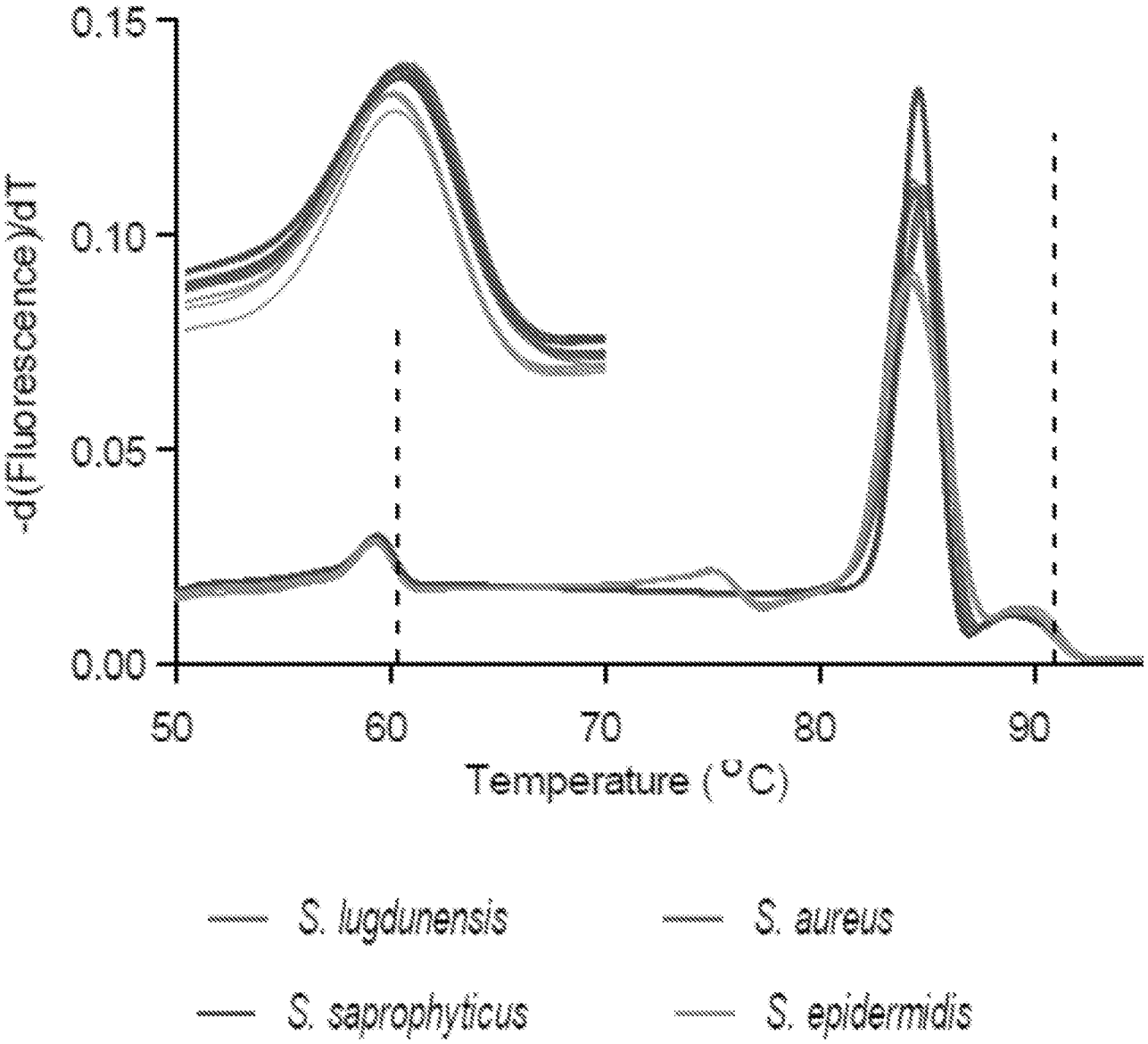

0.15
0.10
0.05
0.00
-d(Fluorescence)/dT
50
60
70
80
90
Temperature (°C)
S. lugdunensis
S. aureus
S. saprophyticus
S. epidermidis

FIG. 5A

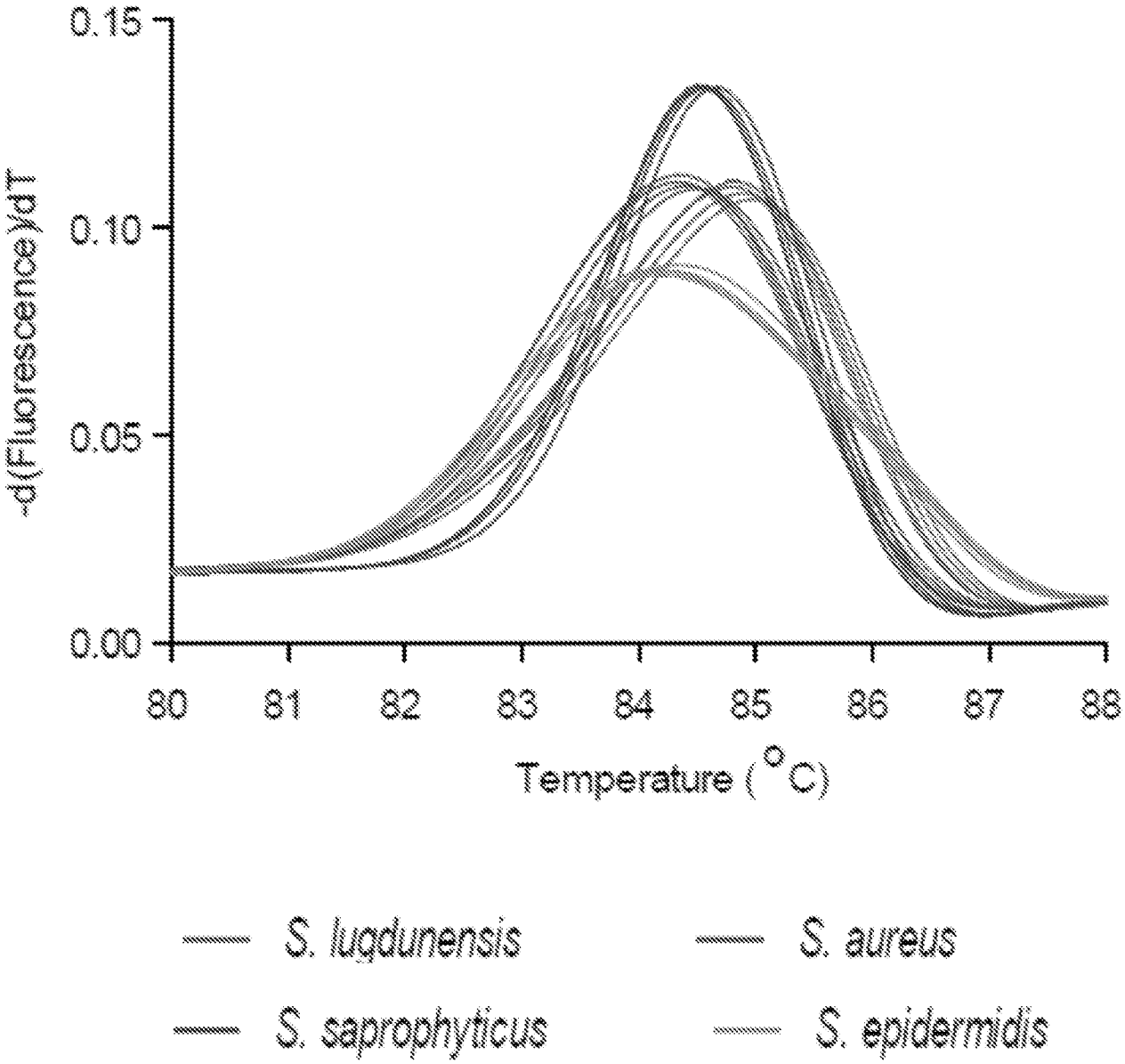

0.15
0.10
0.05
0.00
-d(Fluorescence)/dT
80
81
82
83
84
85
86
87
88
Temperature (°C)
S. lugdunensis
S. aureus
S. saprophyticus
S. epidermidis

FIG. 5B

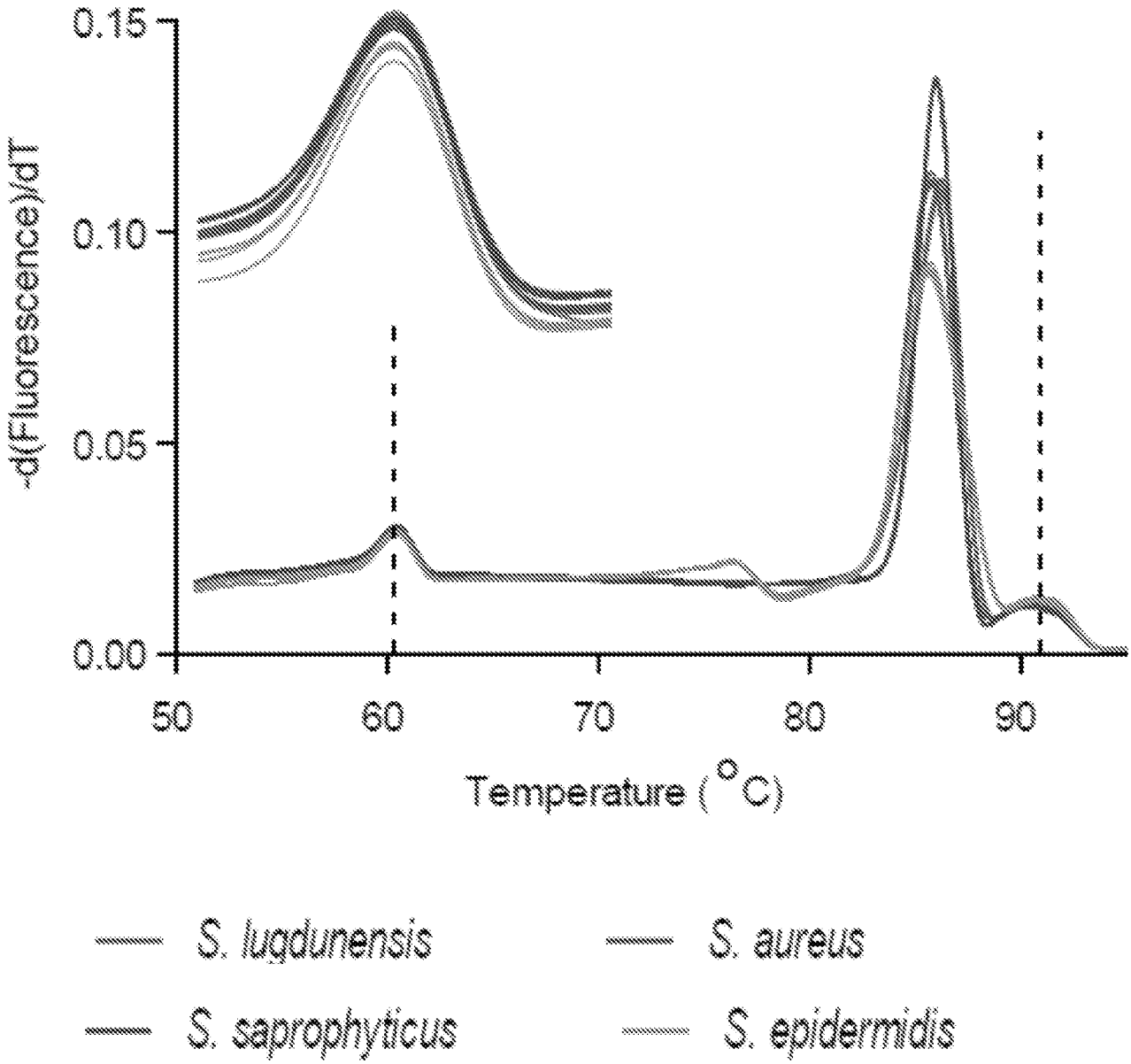

0.15
0.10
0.05
0.00
-d(Fluorescence)/dT
50
60
70
80
90
Temperature (°C)
S. lugdunensis
S. aureus
S. saprophyticus
S. epidermidis

FIG. 5C

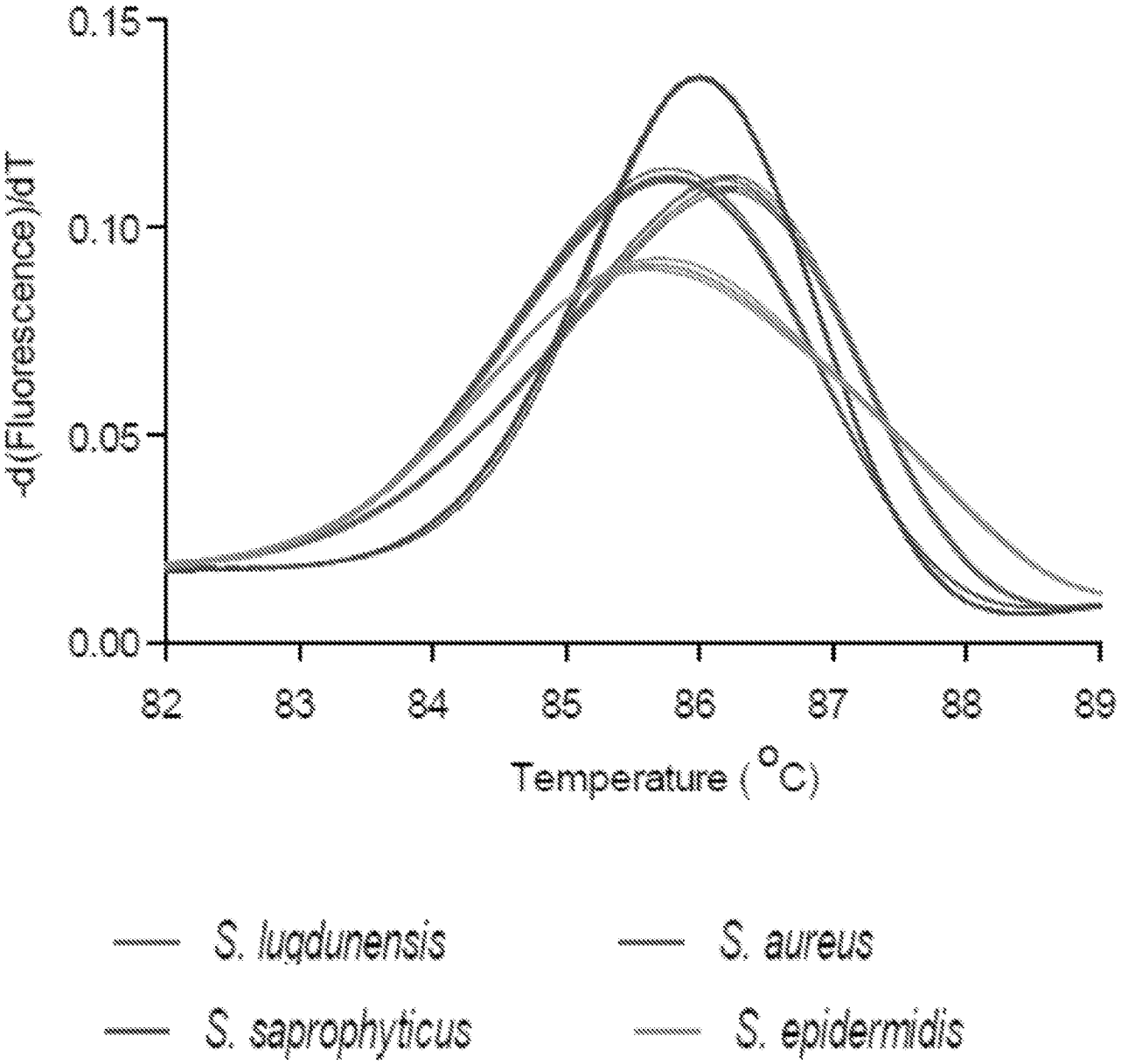

0.15
0.10
0.05
0.00
-d(Fluorescence)/dT
82
83
84
85
86
87
88
89
Temperature (°C)
S. lugdunensis
S. aureus
S. saprophyticus
S. epidermidis

FIG. 5D

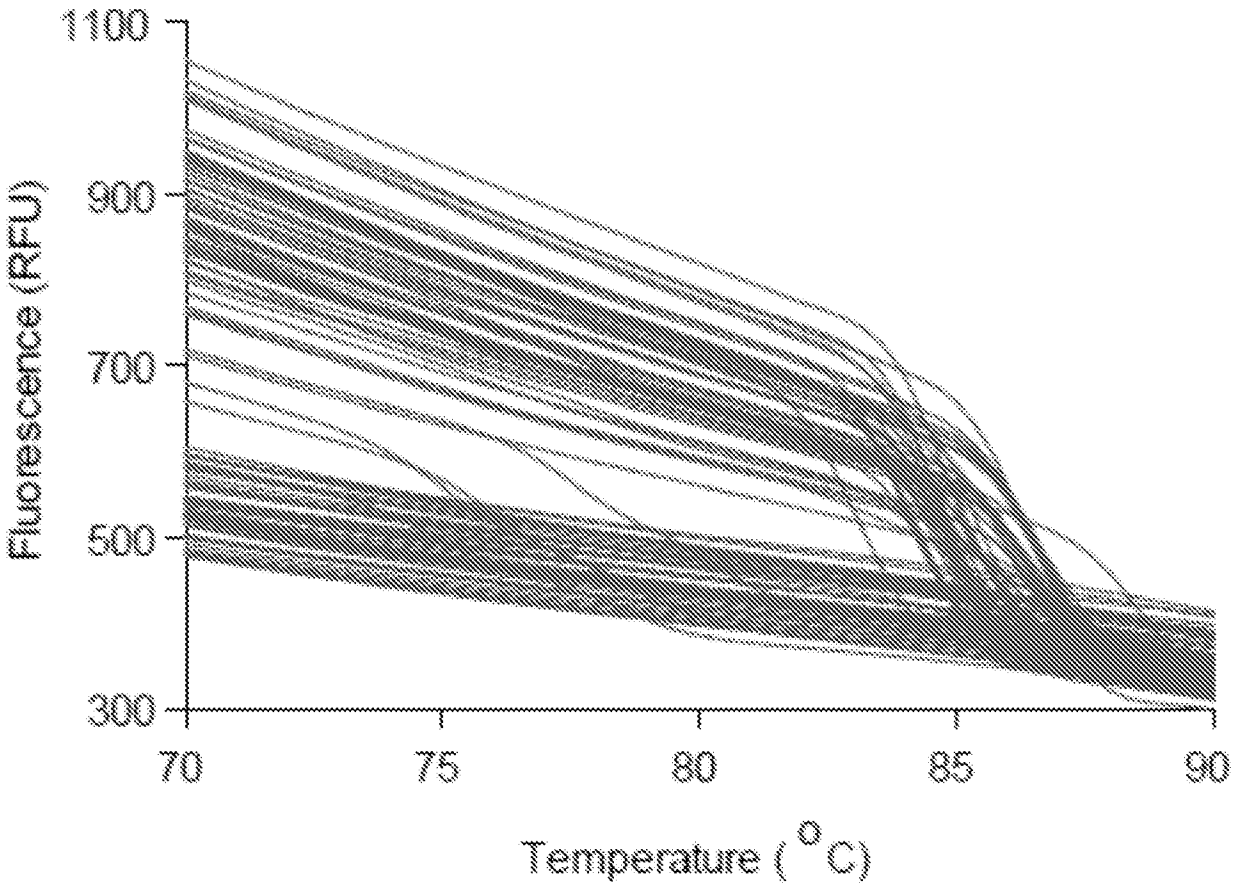

1100
900
700
500
300
Fluorescence (RFU)
70
75
80
85
90
Temperature (°C)

FIG. 6A

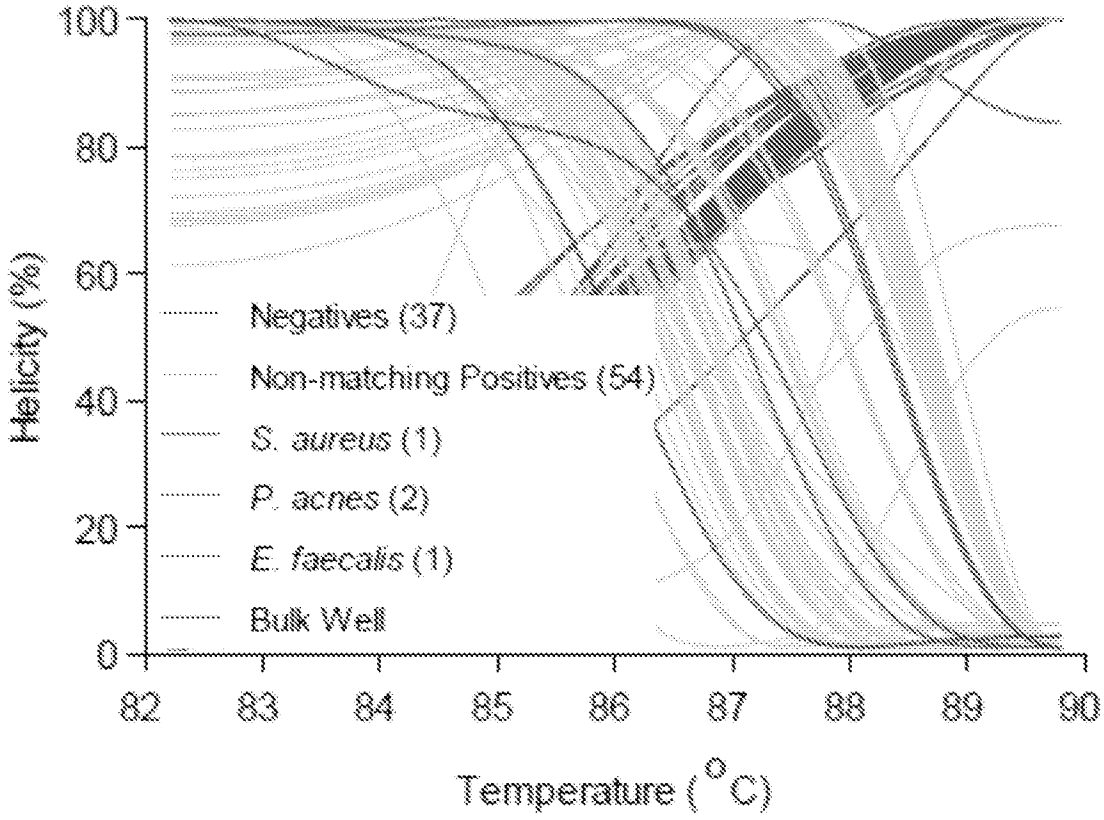

100
80
60
40
20
0
Helicity (%)
Negatives (37)
Non-matching Positives (54)
S. aureus (1)
P. acnes (2)
E. faecalis (1)
Bulk Well
82
83
84
85
86
87
88
89
90
Temperature (°C)

FIG. 6B

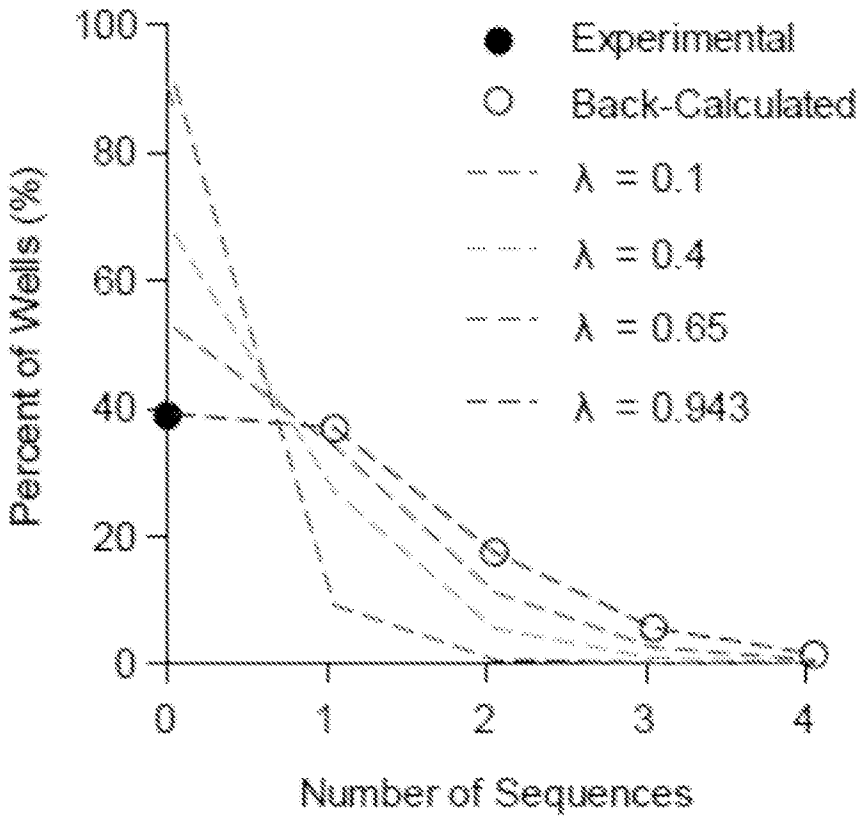

100
80
60
40
20
0
Percent of Wells (%)
0
1
2
3
4
Number of Sequences
Experimental
Back-Calculated
λ = 0.1
λ = 0.4
λ = 0.65
λ = 0.943

FIG. 6C

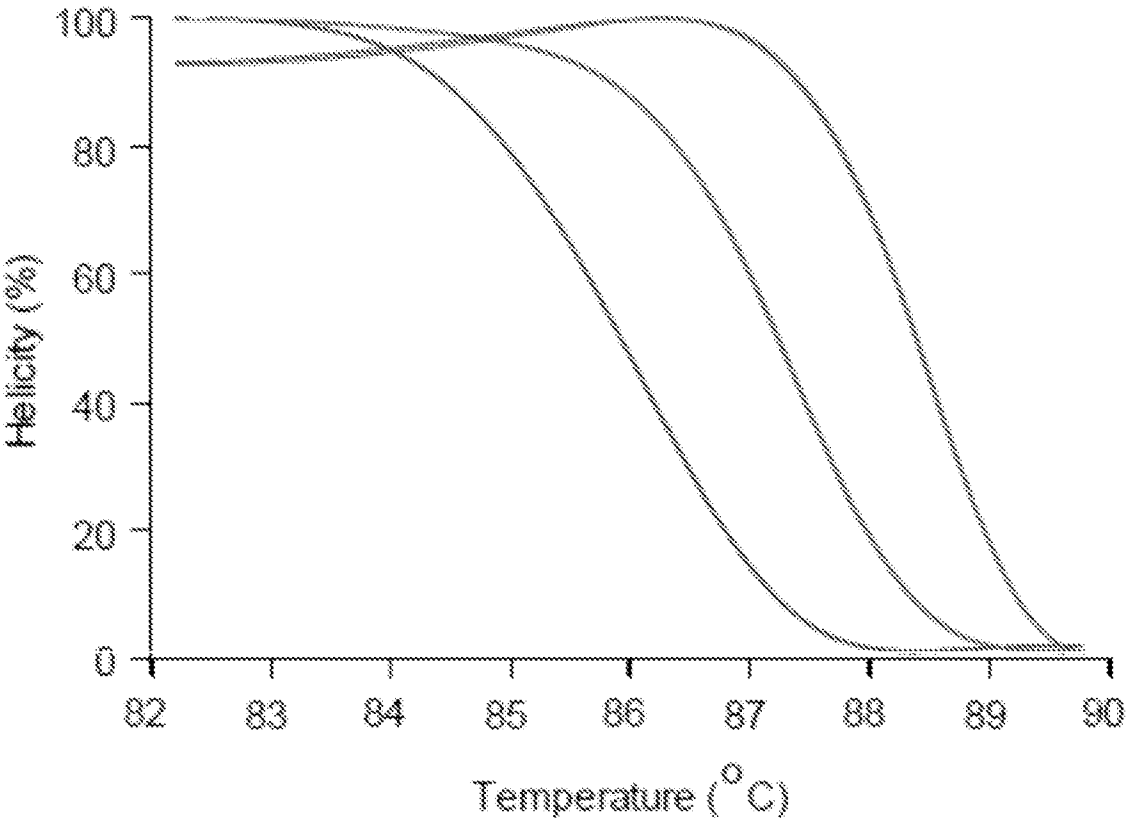

100
80
60
40
20
0
Helicity (%)
82
83
84
85
86
87
88
89
90
Temperature (°C)
S. aureus standard curve
E. faecalis standard curve
P. acnes standard curve

FIG. 6D

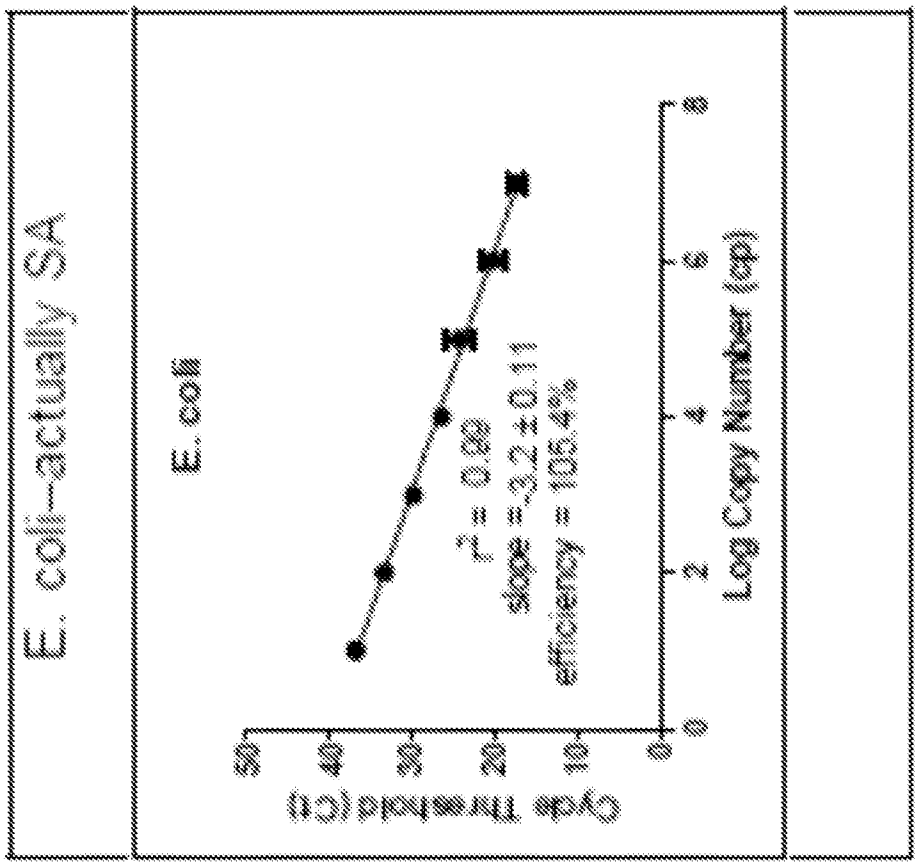

E. coli–actually SA
E. coli
Cycle Threshold (Ct)
Log Copy Number (cp)
r² = 0.99
slope = -3.2 ± 0.11
efficiency = 105.4%

FIG. 7B

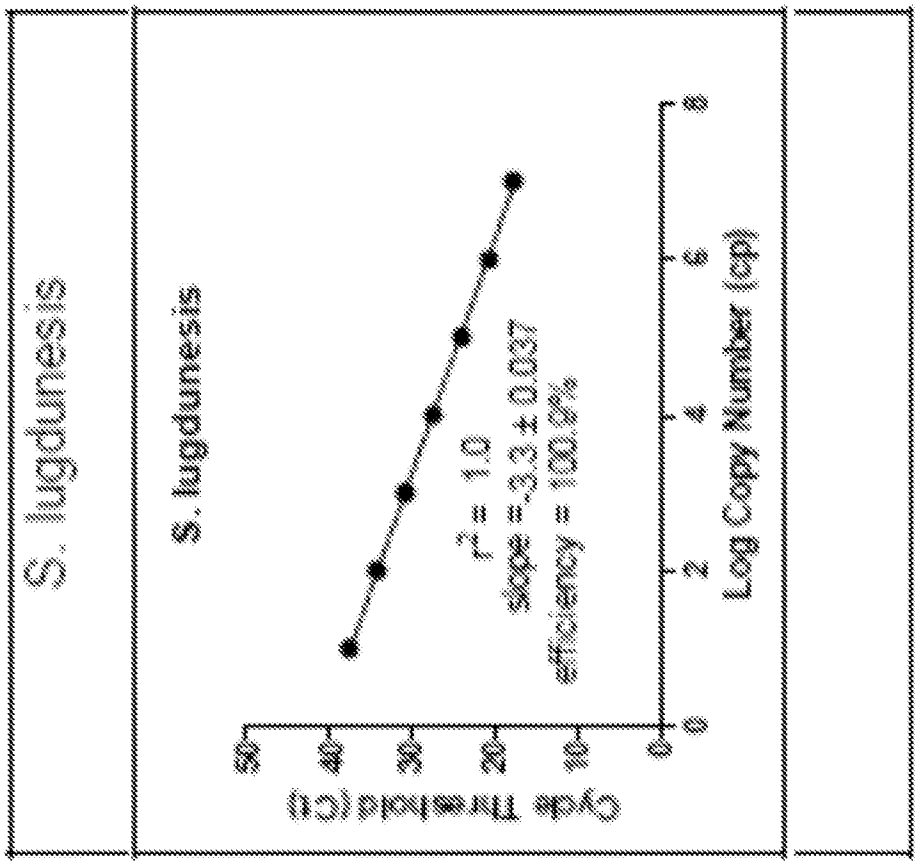

S. lugdunesis
S. lugdunesis
Cycle Threshold (Ct)
Log Copy Number (cp)
r² = 1.0
slope = -3.3 ± 0.037
efficiency = 100.6%

FIG. 7A

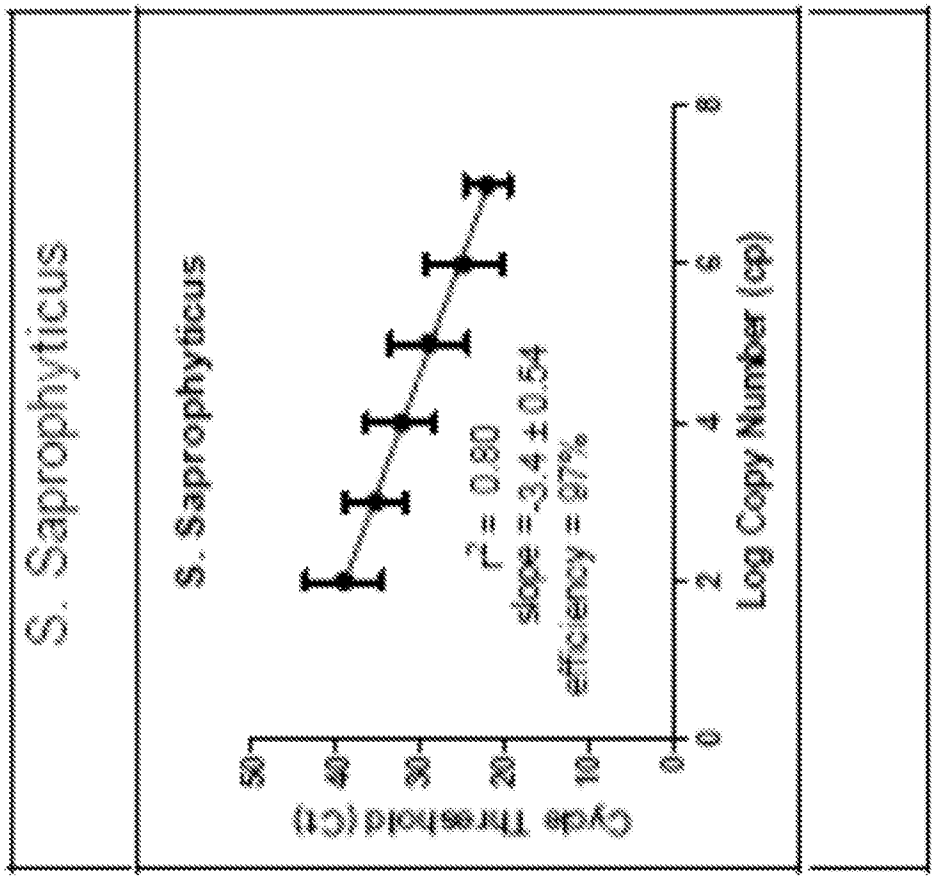

S. Saprophyticus
S. Saprophyticus
Cycle Threshold (Ct)
Log Copy Number (cpl)
r^2 = 0.80
slope = -3.4 ± 0.54
efficiency = 97%

FIG. 7D

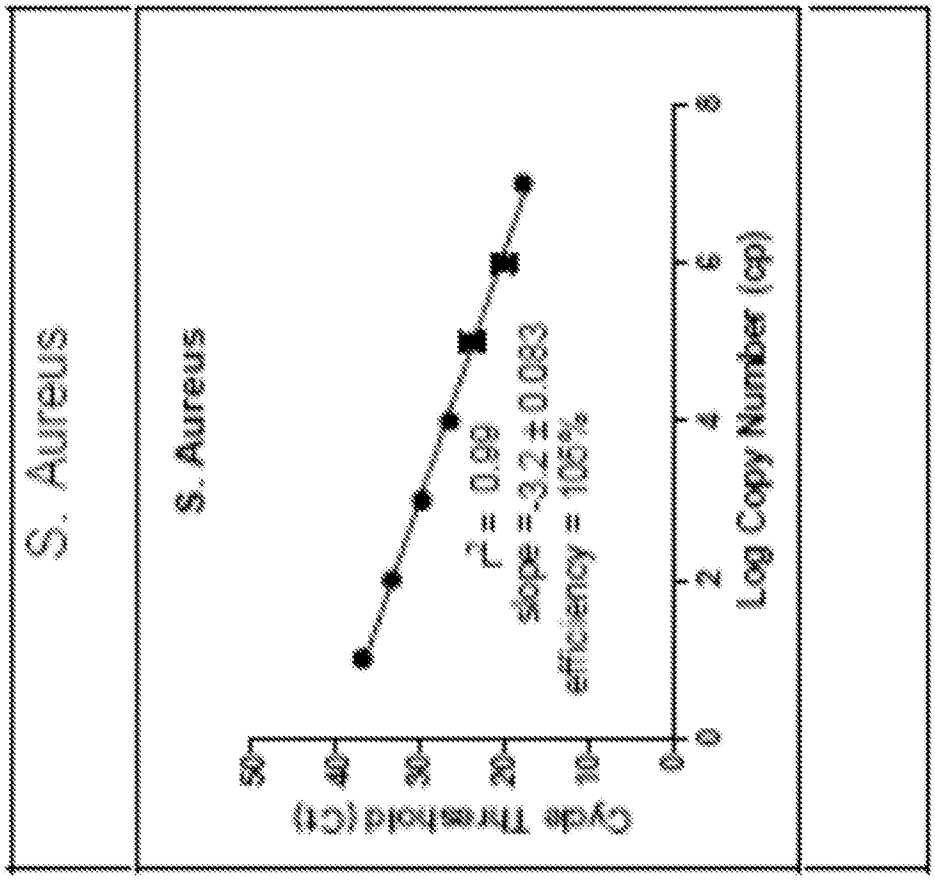

S. Aureus
S. Aureus
Cycle Threshold (Ct)
Log Copy Number (cpl)
r^2 = 0.99
slope = -3.2 ± 0.083
efficiency = 106%

FIG. 7C

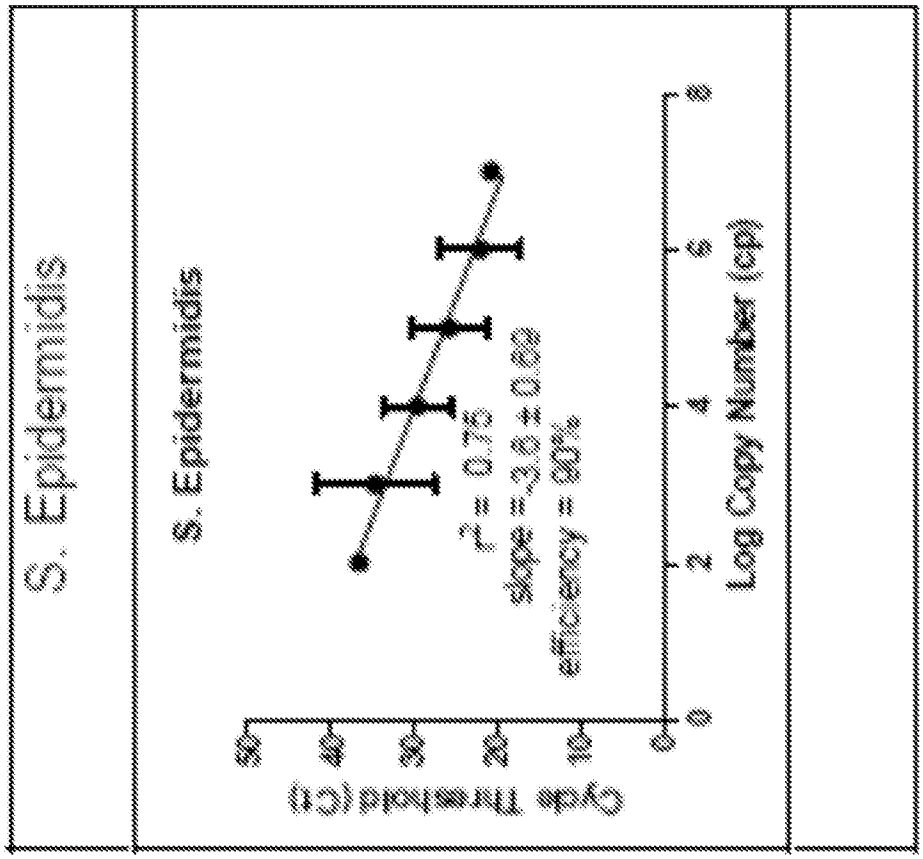

S. Epidermidis
S. Epidermidis
r^2 = 0.75
slope = -3.6 ± 0.08
efficiency = 90%
Cycle threshold (Ct)
Log Copy Number (cp)
0
2
4
6
8
10
12
14
16
18

FIG. 7E

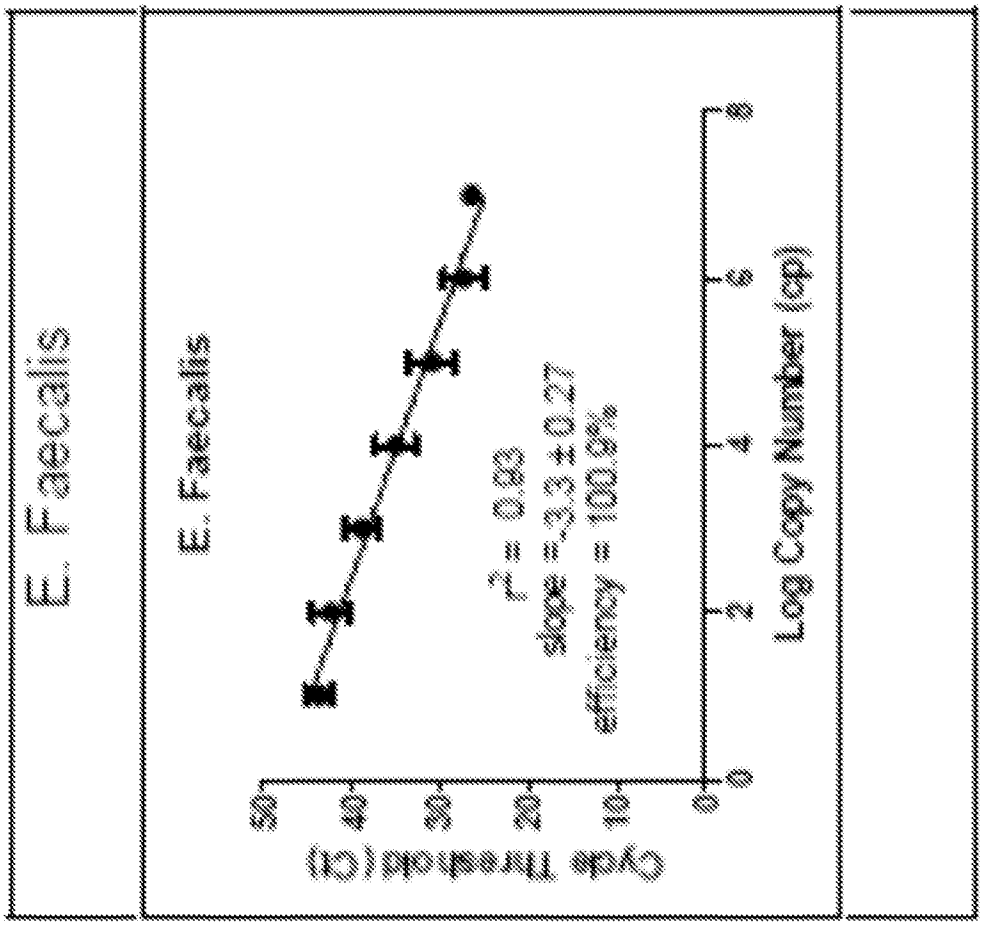

E. Faecalis
E. Faecalis
r^2 = 0.93
slope = -3.3 ± 0.27
efficiency = 100.9%
Cycle Threshold (Ct)
Log Copy Number (cp)
0
10
20
30
40
50
0
2
4
6
8

FIG. 7G

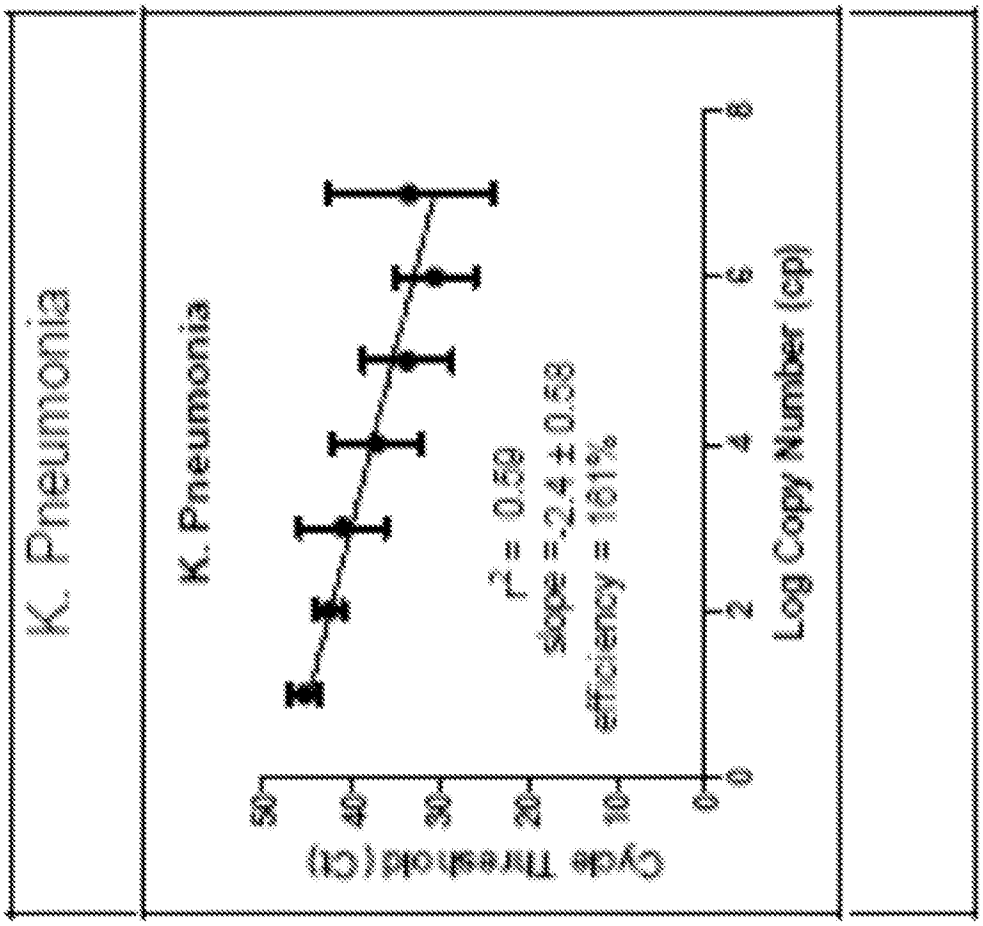

K. Pneumonia
K. Pneumonia
r^2 = 0.99
slope = -2.4 ± 0.03
efficiency = 161%
Cycle Threshold (Ct)
Log Copy Number (cp)
0
10
20
30
40
50
0
2
4
6
8

FIG. 7F

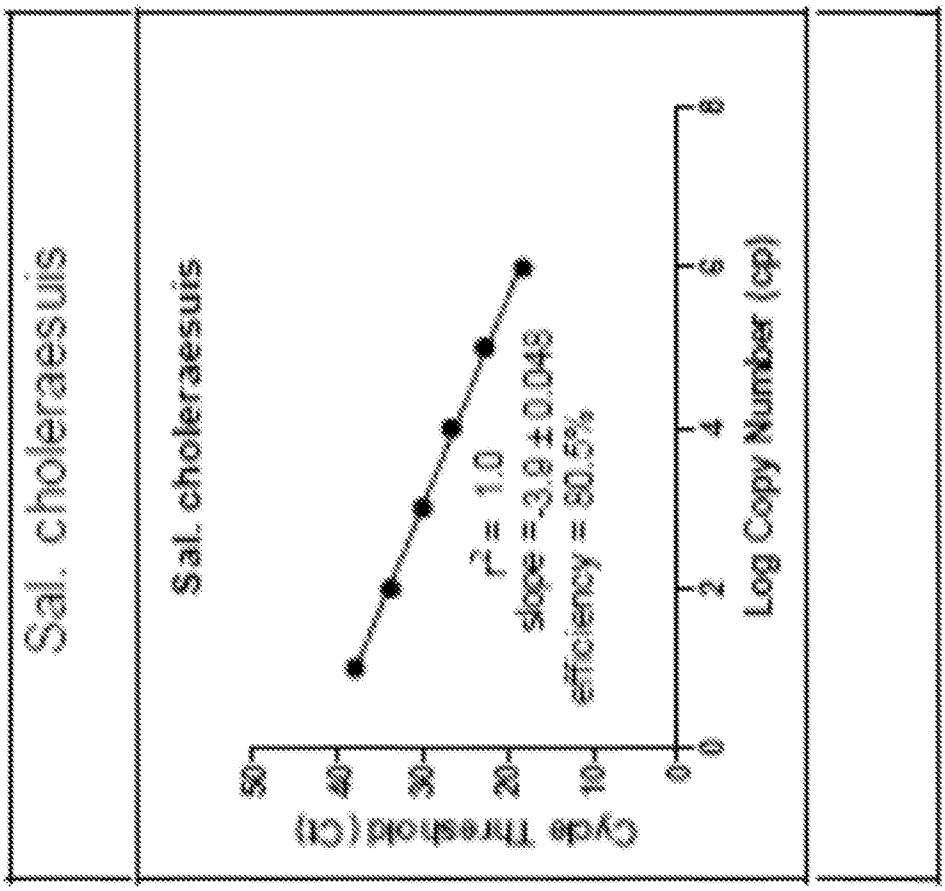

Sal. choleraesuis
Sal. choleraesuis
Cycle Threshold (Ct)
Log Copy Number (cp)
r^2 = 1.0
slope = -3.9 ± 0.048
efficiency = 80.6%

FIG. 7I

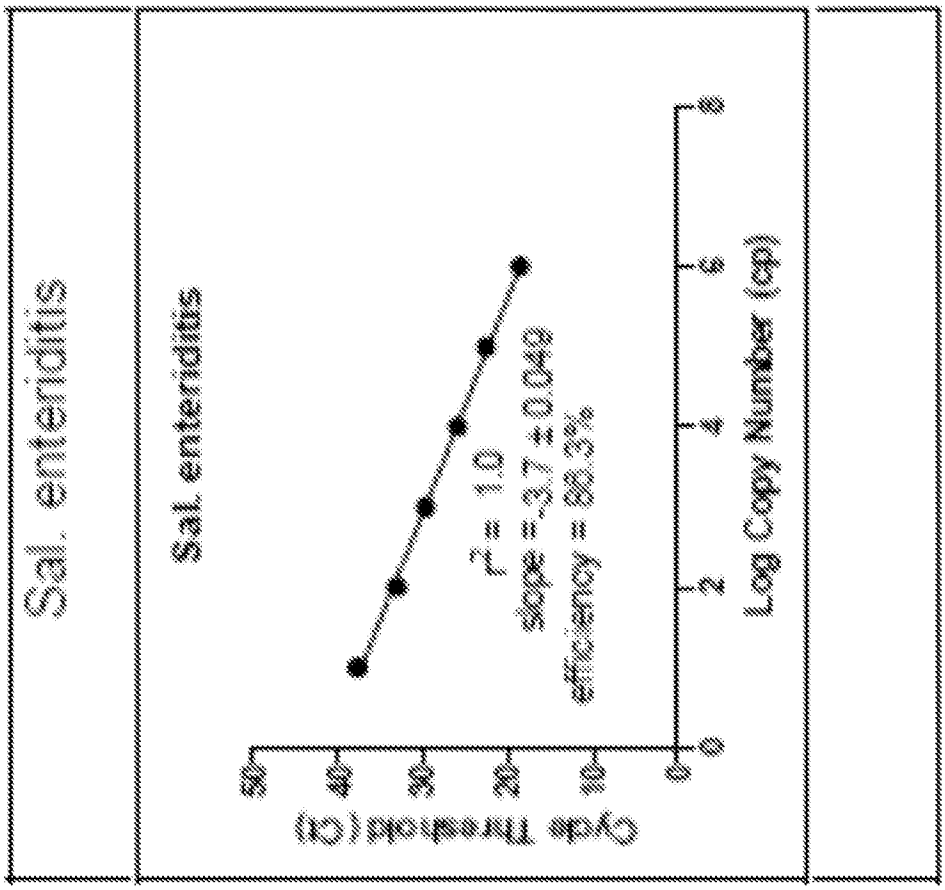

Sal. enteritidis
Sal. enteritidis
Cycle Threshold (Ct)
Log Copy Number (cp)
r^2 = 1.0
slope = -3.7 ± 0.049
efficiency = 88.3%

FIG. 7H

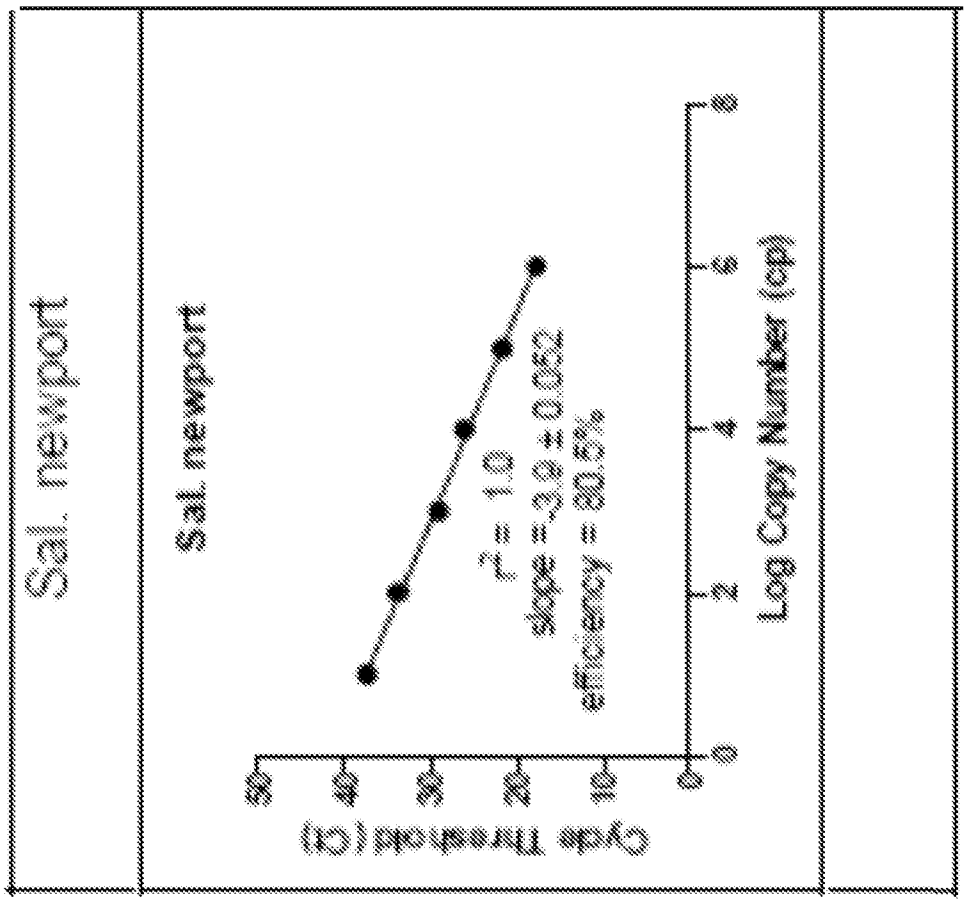

Sal. newport
Sal. newport
Cycle Threshold (Ct)
40
35
30
25
20
0
2
4
6
Log Copy Number (cp)
r² = 1.0
slope = -3.9 ± 0.052
efficiency = 80.6%

FIG. 7J

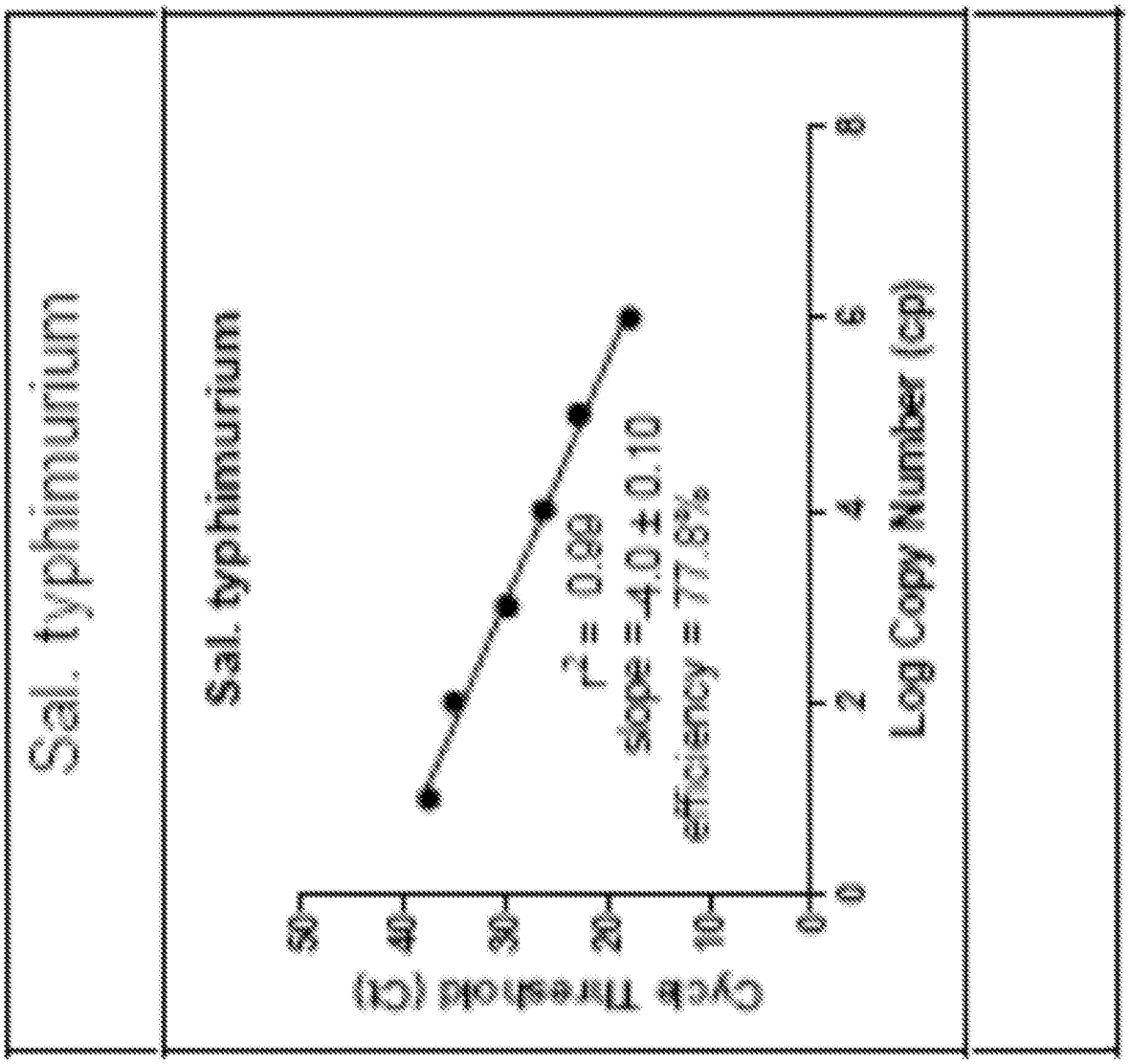

Sal. typhimurium
Sal. typhimurium
r^2 = 0.99
slope = -4.0 ± 0.10
efficiency = 77.6%
Cycle Threshold (Ct)
Log Copy Number (cp)
0
2
4
6
8
35
30
25
20
15

FIG. 7K

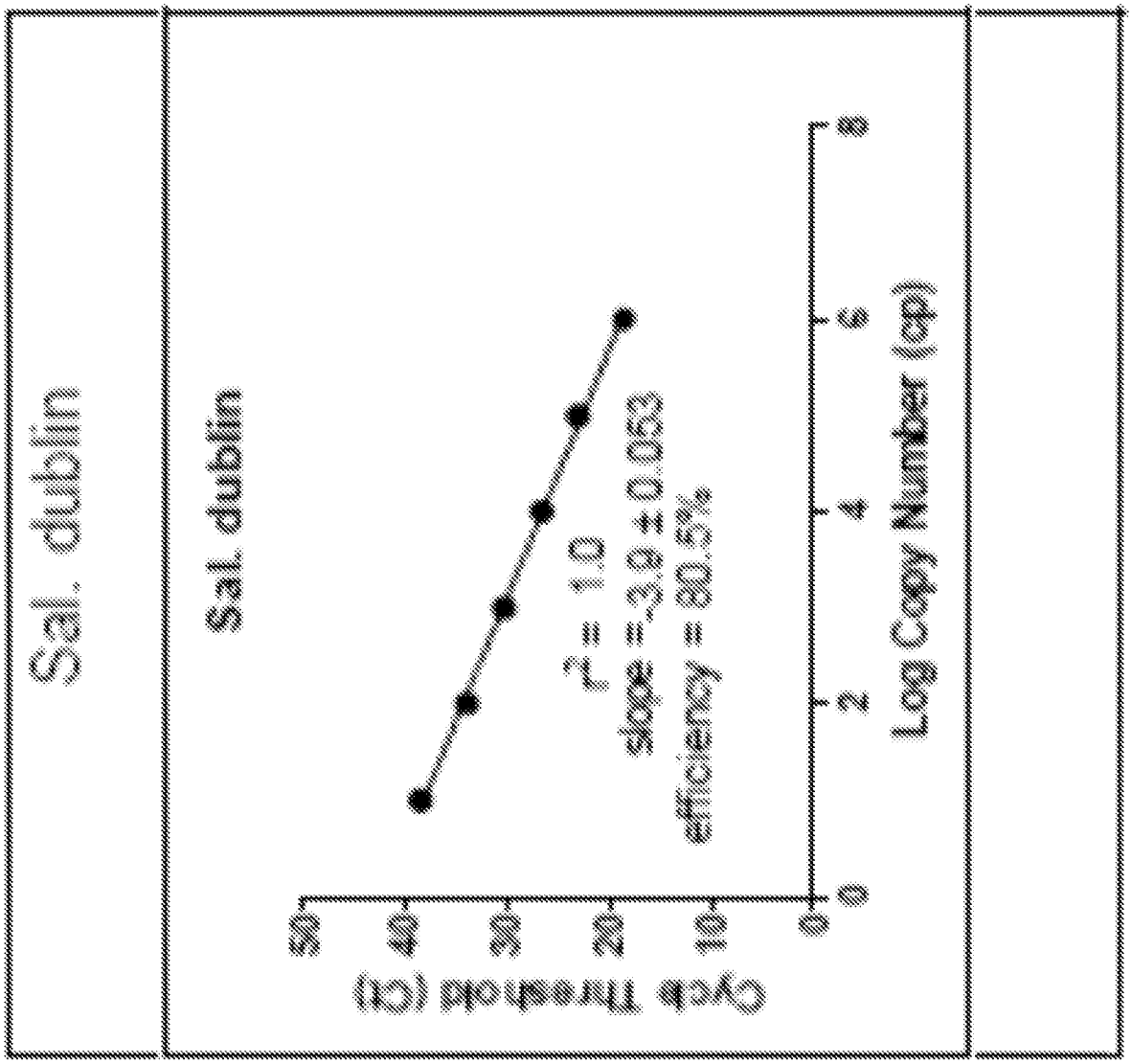

Sal. dublin
Sal. dublin
Cycle Threshold (Ct)
50
40
30
20
10
0
0
2
4
6
8
Log Copy Number (cp)
r^2 = 1.0
slope = -3.9 ± 0.053
efficiency = 80.5%

FIG. 7L

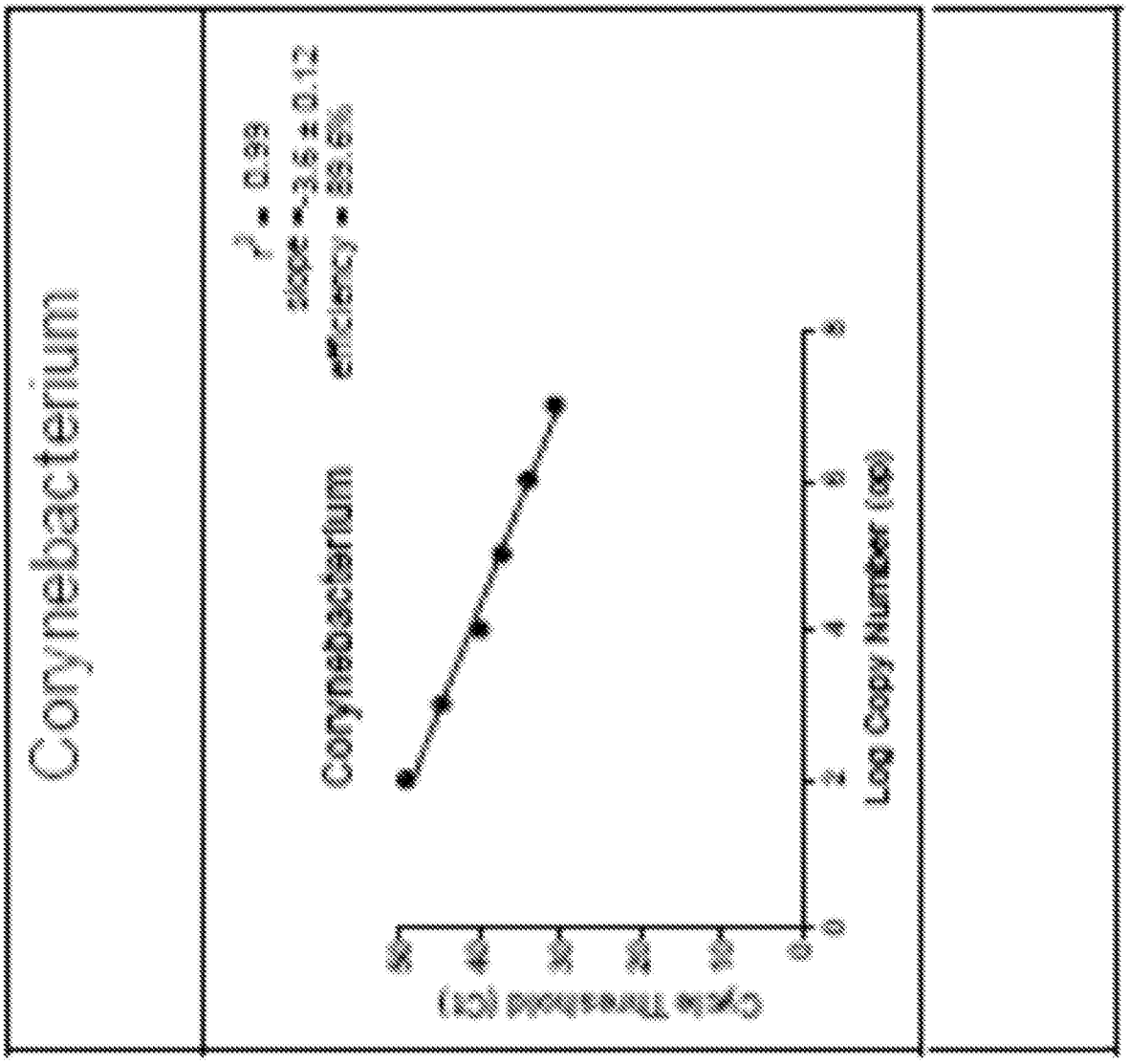

Corynebacterium
Corynebacterium
r^2 = 0.99
slope = -3.6 ± 0.12
efficiency = 89.6%
Log Copy Number (log)
Cycle Threshold (Ct)

FIG. 7M

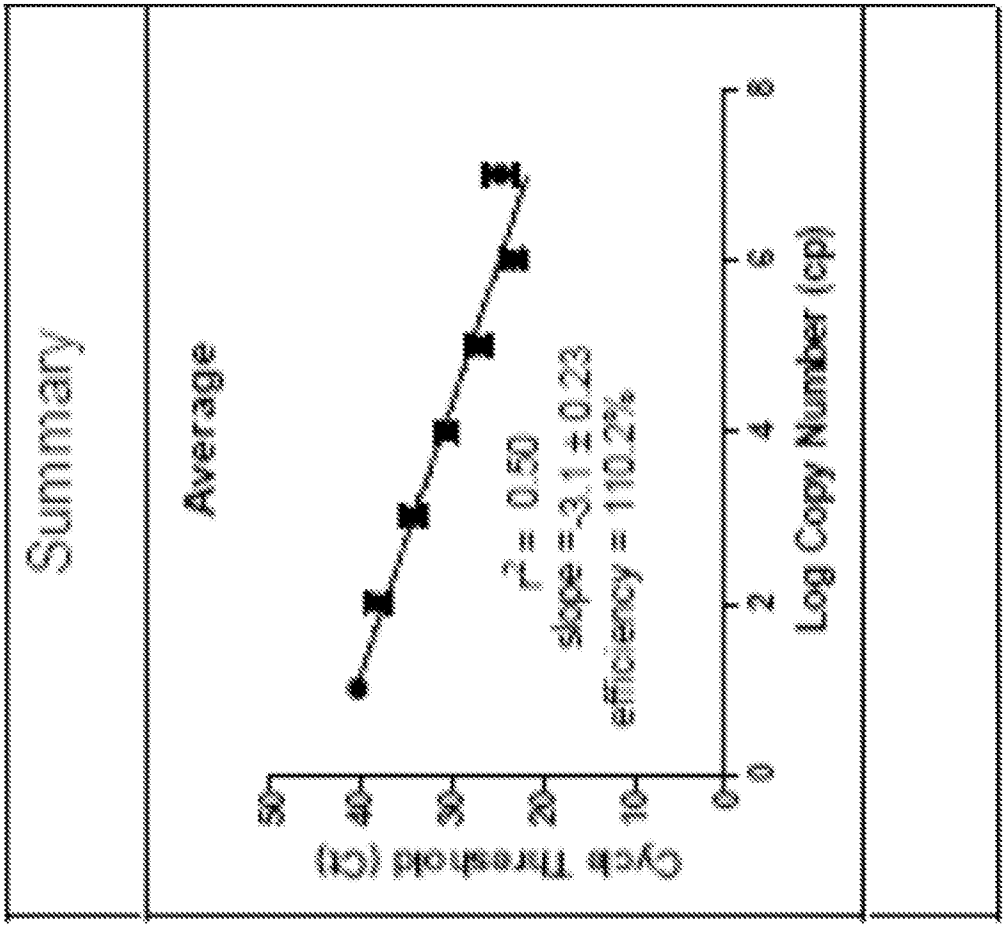

Summary
Average
Cycle Threshold (Ct)
Log Copy Number (cpi)
r² = 0.50
slope = -3.1 ± 0.23
efficiency = 110.2%

FIG. 7O

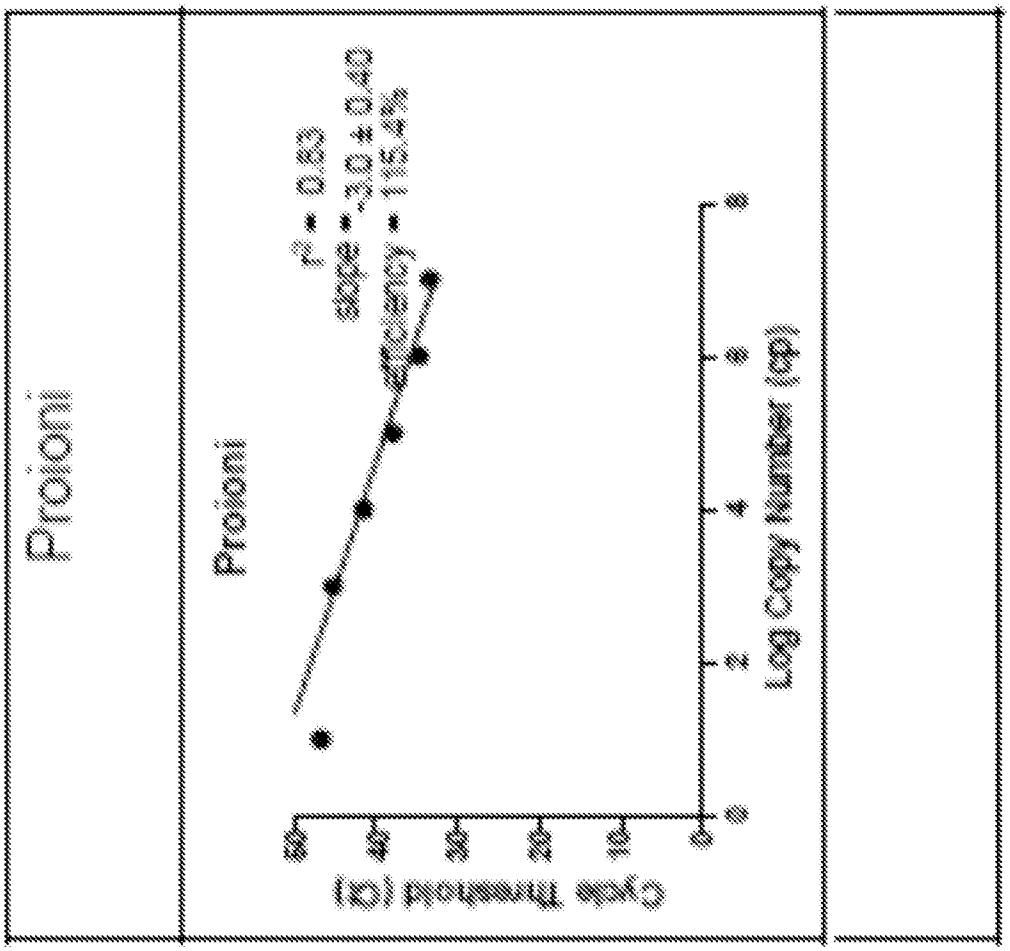

Proioni
Proioni
Cycle Threshold (Ct)
Log Copy Number (cpi)
r² = 0.53
slope = -3.0 ± 0.40
efficiency = 115.4%

FIG. 7N

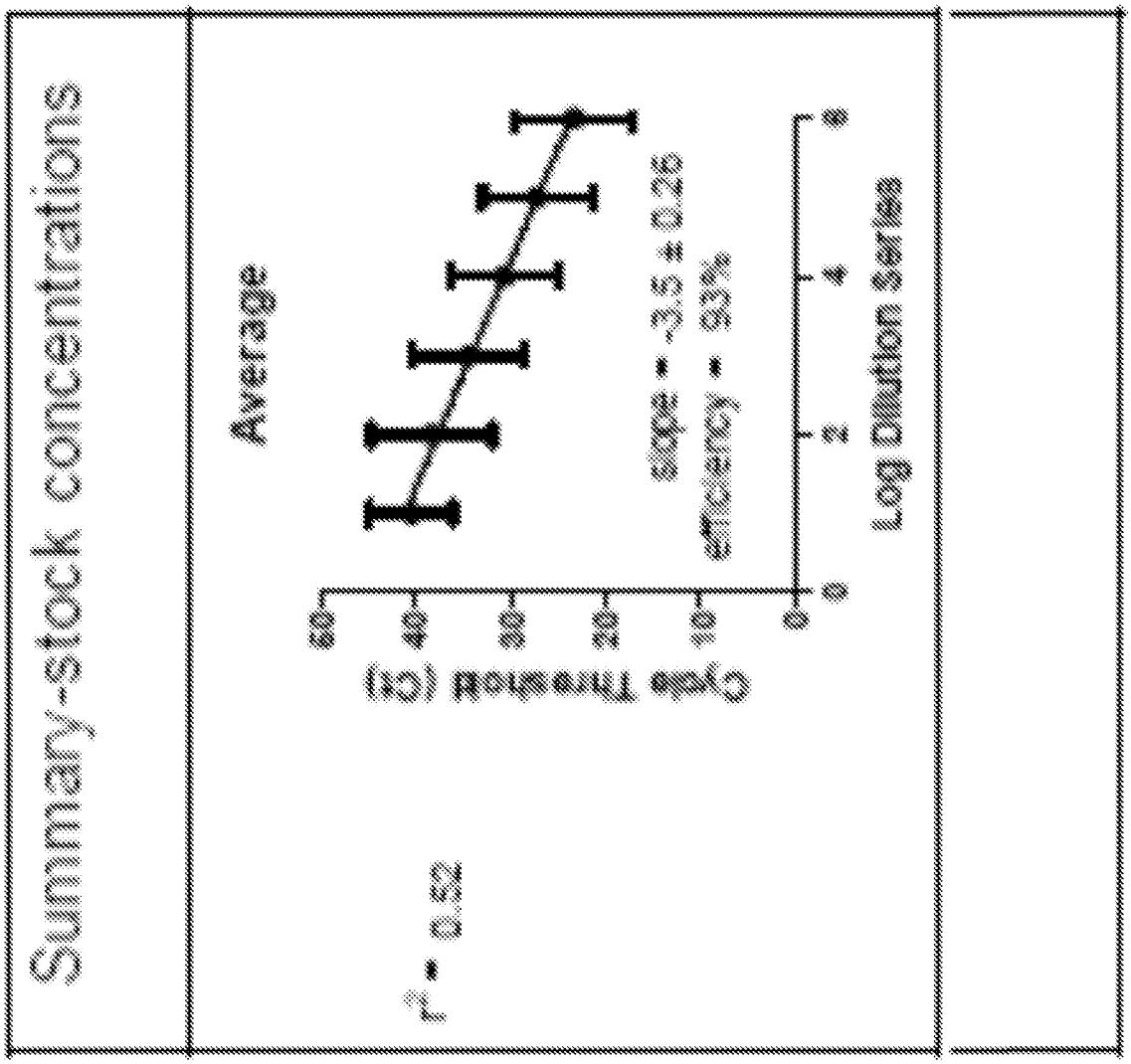

Summary-stock concentrations
Average
r² = 0.82
Cycle Threshold (Ct)
60
40
30
20
10
0
0
2
4
6
Log Dilution Series
slope = -3.5 ± 0.26
efficiency = 93%

FIG. 7P