Programming bacteria for multiplexed DNA detection

Yu-Yu Cheng[1†], Zhengyi Chen[1†], Xinyun Cao[1], Tyler D. Ross[1],
Tanya G. Falbel[2], Briana M. Burton[2], and Ophelia S. Venturelli[1,2,3*]

[1]Department of Biochemistry, University of Wisconsin - Madison, WI, United States
[2]Department of Bacteriology, University of Wisconsin - Madison, WI, United States
[3]Department of Chemical & Biological Engineering, University of Wisconsin-Madison

[†]These authors contributed equally to this work.
[*]To whom correspondence should be addressed: venturelli@wisc.edu

ABSTRACT

DNA is a universal and programmable signal of living organisms. Here we developed cell-based DNA sensors by engineering the naturally competent bacterium *Bacillus subtilis* (*B. subtilis*) to detect specific DNA sequences in the environment. The DNA sensor strains can identify diverse bacterial species including major human pathogens with high specificity and sensitivity. Multiplexed detection of genomic DNA from different species in complex samples can be achieved by coupling the sensing mechanism to orthogonal fluorescent reporters. We also demonstrate that the DNA sensors can detect the presence of species in the complex samples without requiring DNA extraction. The modularity of the living cell-based DNA sensing mechanism and simple detection procedure could enable programmable DNA sensing for broad applications.

INTRODUCTION

Next-generation engineered bacteria hold tremendous promise for a wide range of applications in human health, environment and agriculture by sensing key environmental signals, performing computation on these signals to regulate a response that modulates specific environmental parameters[1]. Developing specific and selective sensors of key environmental signals is a critical feature of next-generation engineered bacteria. For example, bacteria have been engineered to detect physical and chemical signals such as light, ultrasound, and quorum-sensing molecules[2–4]. These signals can be exploited to control the collective growth or gene expression of the bacterial population or mediate interactions between constituent community members[4–6]. In addition, we can exploit their sensing ability to achieve real-time monitoring of natural environments. For example, synthetic genetic circuits have been designed in *Escherichia coli* (*E. coli*) to sense signals produced by pathogens and use this information to regulate the production of antimicrobials that inhibit the target pathogen[7,8]. However, there are limited well-characterized and orthogonal signals that can be exploited to sense different bacterial species in a microbial community[9,10].

DNA provides the blueprint for living organisms and is prevalent in natural environments[11]. Therefore, extracellular DNA (eDNA) could be exploited as a biomarker for identifying different species. Naturally competent bacteria have the ability to take up DNA from the environment and integrate imported sequences onto the genome based on sequence homology requirements. Horizontal gene transfer (HGT) via natural transformation has been shown to have a variety of benefits such as nutrient utilization, DNA repair, or acquisition of genes[12]. Since homologous recombination of imported DNA requires sequences of sufficient length and homology[13], natural transformation could be exploited to build a selective cell-based DNA sensor.

We constructed a living cell-based DNA sensor by engineering the naturally competent bacterium *B. subtilis*. This circuit controls *B. subtilis* growth and fluorescence reporter genes in response to specific input DNA sequences. We demonstrate that the cell-based DNA sensor is

sensitive and highly specific to species harboring the target DNA sequence. In addition, we demonstrate that our cell-based DNA sensor can perform multiplexed DNA detection in complex samples. The cell-based DNA sensors can detect DNA released from pre-treated donor cells (i.e. crude samples). Our detailed characterization of the cell-based DNA sensors *in vitro* provides a foundation for future *in vitro* DNA detection and *in situ* sense-and-respond DNA applications.

RESULTS

Construction of a living DNA sensor strain

To build the living cell-based DNA sensor, we exploited the natural competence ability of the well-characterized soil bacterium *B. subtilis*[14]. The natural competence ability of *B. subtilis* enables uptake of environmental DNA and integration of specific sequences with sufficient homology into genome via homologous recombination[15]. The efficiency of homologous recombination depends stringently on the sequence percent identity and length[13,16], which can be exploited to build a highly specific DNA sensor.

To detect eDNA sequences in a programmable fashion, we constructed a synthetic genetic circuit in *B. subtilis* that implements a growth selection function based on the presence of target DNA sequence in the environment. The circuit consists of a xylose-inducible master regulator of competence *comK*[17] and IPTG-inducible toxin-antitoxin system *txpA-ratA*[18] and GFP regulated by the repressor LacI (**Fig. 1A,S1**). The target sequences were introduced to the flanking regions (upstream and downstream) of *txpA-ratA* and *lacI* and referred to as landing pads for homologous recombination.

In the presence of xylose, ComK activates competence genes for DNA uptake and homologous recombination (**Fig. 1A**). Bistability and stochastic processes in the regulation of natural competence can yield a sub-population that can be transformed with extracellular DNA[19]. This naturally competent sub-population forms competence pili which bind to double stranded DNA outside the cell. The DNA is cleaved into single-stranded DNA (ssDNA) outside the cell membrane, and transported into the cell[12]. Inside the cell, RecA binds the ssDNA sequences and searches the *B. subtilis* genome for a region with sufficient homology. If the target DNA sequence is present, homologous recombination removes the toxin-antitoxin *txpA-ratA* and repressor *lacI*. In the presence of the chemical inducer Isopropyl β-D-1-thiogalactopyranoside (IPTG), cell growth and GFP expression is enabled in the transformed sub-population. Growth of the non-transformed subpopulation is inhibited by the activity of TxpA, which blocks cell wall synthesis[18] (**Fig. 1A**).

We constructed a sensor for *E. coli* (EC sensor) by introducing the *xdhABC* operon onto the *B. subtilis* genome (landing pad region), which encodes genes for purine catabolism[20]. The

xdhABC operon is a representative sequence that can detect a wide range of *E. coli* strains. This sequence is highly conserved such that 99% of 5000 *E. coli* genomes in the NCBI database contain this sequence with >95% coverage (the degree of alignment of the query sequence with a reference sequence) and >95% identity similarity (the percentage of bases that are identical to the target sequence within the aligned region).

To characterize the homology length needed for robust DNA sensing, we varied the homology length of the *xdhABC* operon in each landing pad (0.5 to 2.5 kb). We performed time-series measurements of transformation efficiency (number of colonies for transformed *B. subtilis* cells divided by the number total *B. subtilis* colonies) with 100 ng/mL *E. coli* genomic DNA (gDNA). The transformation efficiency is defined as the ratio of the number of transformed *B. subtilis* to the total number of *B. subtilis* based on colony forming units (CFU). Transformation efficiency plateaued at approximately 10 hr and the colonies expressed GFP (**Fig. 1B**). In addition, transformation efficiency increased with homology length at 10 hr (**Fig. 1C**). A homology length of 1 kb or greater was required to robustly sense the target sequence over the background frequency of escape mutants (10^{-7}-10^{-6} frequency) that displayed heterogenous GFP expression (**Fig. S2A,B**). To achieve high performance of the DNA sensor (>10^2 increase in transformation efficiency above background), we used a landing pad homology length of 2.5 kb (transformation efficiency of 10^{-5}-10^{-4}). In the transformed sub-population, homologous recombination was confirmed by sequencing to occur at the expected location with the elimination of *txpA*-*ratA* and *lacI* (**Fig. S2C,D**). The moderate number of escape mutants that displayed growth in the absence of gDNA had mutations in *txpA* or *lacI*, which reduced the growth inhibitory activity of TxpA (**Fig. S2E,F**). In sum, the synthetic genetic circuit enabled *B. subtilis* to sense specific DNA sequences present in the environment.

Building living DNA sensors to sense human pathogens

Exploiting the modularity of the DNA sensing circuit, we replaced the landing pad region with specific sequences targeting different bacterial strains (**Fig. S1C**). To this end, we constructed DNA sensors to detect sequences harbored in human intestinal pathogens *Salmonella typhimurium*[21] (*S. typhimurium*), *Clostridium difficile*[22] (*C. difficile*), or the skin pathogen *Staphylococcus aureus*[23] (*S. aureus*). We selected two 2.5 kb sequences in the pathogenicity island *sipBCDA* of *S. typhimurium* (ST sensor), the heme biosynthesis pathway *hemEH* in *S. aureus* (SA sensor), and the phenylalanyl-tRNA synthetase *pheST* in *C. difficile* (CD sensor)[24–26].

The selected set of target DNA sequences are highly conserved within a given species (ST sensor: 94%, SA sensor: 96%, and CD sensor: 96% all with >95% coverage and >95%

106 identity similarity). In addition, some of the sequences are linked to virulence activities of the
107 pathogen or encode enzymes that are critical for fitness[24–26]. To further explore the conservation
108 of the target sequences across different strains, we performed nucleotide BLAST using the NCBI
109 Database to quantify the homology coverage and sequence similarity across species. The
110 pathogenicity island *sipBCDA* in *S. typhimurium* was found only in *Salmonella enterica* species
111 and infrequently observed in other species (**Fig. S3A**). Homologs in other species have low
112 coverage and identity similarity, suggesting that the pathogenicity island could be a good target
113 sequence for this species (**Fig. S3A**). The heme biosynthesis pathway *hemEH* in *S. aureus* and
114 phenylalanyl-tRNA synthetase *pheST* in *C. difficile* are conserved in some closely related strains
115 with varying degrees of similarity and coverage (**Fig. 2I,K**). The *E. coli* MG1655 *xdhABC* purine
116 catabolism operon is found in other closely related bacteria such as *Shigella* with high coverage
117 and identity similarity (**Fig. S3B**). Although the target sequences for building the different DNA
118 sensor strains varied in the degree of specificity based on bioinformatic analyses, a detailed
119 characterization of circuit performance could guide the design of optimized cell-based DNA
120 sensors for future applications.

121 The four sensors robustly detected the presence of 100 ng/mL target gDNA over
122 background (0 ng/mL gDNA) based on transformation efficiency (**Fig. 2A**). We evaluated the
123 sensitivity of each DNA sensor strain by performing time-series GFP measurements in liquid
124 culture after being transformed with a wide range of gDNA concentrations (0-1500 ng/mL gDNA)
125 from single species (**Fig. 2B,C,S4**). We evaluated the time required for each culture to display a
126 fluorescence level higher than a threshold (i.e. detection time) (**Fig. 2C,S4**). The sensitivity of the
127 circuit was evaluated as the lowest gDNA concentration that yielded a statistically significant
128 difference in the detection time in the presence versus absence of gDNA (**Fig. 2D-G**).

129 The relationship between the log transformed gDNA concentration and detection time is
130 linear due to the exponential growth of the fluorescent *B. subtilis* sub-population successfully
131 transformed with the input target sequence (**Fig. S5**). Therefore, to assess the range of gDNA
132 concentrations that can be accurately sensed, we inferred the parameters of a linear function fit
133 to the log transformed gDNA concentration versus detection time (**Fig. 2D-G**). The inferred slope
134 of the linear function is determined by the cell doubling time (~0.5 hour) and intercept is
135 determined by the background mutation frequency (**Fig. 2D-G**). The EC, SA, and CD DNA sensor
136 strains displayed high sensitivity of 1-16 ng/mL (10^5-10^6 chromosome copy number/mL), whereas
137 the ST sensor displayed a lower sensitivity (62.5 ng/mL, 10^7 chromosome copy number/mL).
138 While the DNA sensor strains displayed lower sensitivities than quantitative real-time polymerase
139 chain reaction (qPCR) reported for *E. coli* (3.5×10^3 CFU/mL in pure culture[27]), the observed

140 sensitivities are within the range of the sensitivities reported for the lateral flow immunoassay
141 ($1.8x10^5$ CFU/mL for a pure culture of *E. coli*[28]) .
142 To characterize the specificity of each DNA sensor strain to the target sequence, we
143 performed time-series fluorescence measurements in liquid culture in response to all individual
144 species gDNA. The fluorescence signal was observed at a substantially earlier time (6.1-7.1 hr)
145 in the presence of the corresponding species' gDNA than in the presence of a non-target species'
146 gDNA or in the absence of DNA (9.1-10.7 hr) (**Fig. 2H, S6**). This demonstrates that the cell-based
147 DNA sensors were highly specific to the target sequence.
148 In the context of microbial communities, the cell-based DNA sensors may need to
149 distinguish between closely related species. Therefore, we evaluated the ability of the DNA
150 sensors to distinguish between closely related species with similar target sequences. To this end,
151 we measured the transformation frequency of the SA sensor in the presence of gDNA (100 ng/mL)
152 derived from *S. epidermidis*. *S. epidermidis* is a closely related human skin commensal bacterium
153 that harbors a similar *hemEH* sequence to the SA sensor landing pad region (89% coverage and
154 77% identity similarity) (**Fig. 2I**). Since the number of total *B. subtilis* colonies was similar in the
155 presence and absence of gDNA, we quantified the number of transformed colonies as opposed
156 to transformation efficiency (**Fig. 2A,S7**). The number of colonies in the presence of *S.*
157 *epidermidis* gDNA (SE) was substantially lower than in the presence of *S. aureus* gDNA (SA) and
158 similar to the absence of DNA (**Fig. 2J**). Similarly, we characterized the ability of the CD sensor
159 to detect gDNA derived from a human gut commensal bacterium *C. hiranonis*, a close relative of
160 *C. difficile* that contains a similar *pheST* sequence in its genome to the landing pad region in CD
161 sensor (87% coverage and 75% identity similarity). The number of colonies in the presence of
162 100 ng/mL *C. hiranonis* gDNA (CH) was substantially lower than in the presence of *C. difficile*
163 gDNA (CD) and also similar to the absence of DNA (**Fig. 2K,L**). These data demonstrate that the
164 cell-based DNA sensors are highly specific to species that harbor an exact match to the target
165 sequence. Therefore, the sensors do not display false positives in the presence of closely related
166 species that harbor similar target sequences. The high specificity of the sensors is due to the
167 stringent requirements for homologous recombination in *B. subtilis*[13,29].
168
169 *Multiplexed detection of pathogen DNA in complex samples*
170 Since certain future applications may require sensing of more than one organism, we tested the
171 ability of the DNA sensors to detect more than one species within mixed DNA samples. To this
172 end, we constructed individual sensors with orthogonal fluorescent reporters to achieve
173 multiplexed DNA detection. Exploiting the modularity of the circuit, we constructed an RFP-

174 labeled ST sensor (ST-R) and a BFP-labeled SA sensor (SA-B), in addition to the GFP-labeled
175 EC sensor (EC-G) (**Fig. S1C**). We introduced gDNA (200 ng/mL) extracted from each of the three
176 target strains into a culture containing EC-G, ST-R, and SA-B sensors and determined the number
177 of fluorescent colonies for each reporter (**Fig. 3A**). The sensors accurately reported the
178 presence/absence of all combinations of species' gDNA reliably (**Fig. 3B,S8**). Therefore, a
179 mixture of DNA sensor strains each individually labeled with a unique fluorescent reporter enabled
180 multiplexed detection of gDNA derived from different species.

181 To investigate if multiplexed detection can be achieved for samples derived from a
182 complex microbial community, we constructed a four-member human gut community composed
183 of diverse commensal bacteria from three major phyla in human gut – *Anaerostipes caccae* (AC,
184 Firmicutes), *Bacteroides thetaiotaomicron* (BT, Bacteroidetes), *Bifidobacterium longum* (BL,
185 Actinobacteria), and *Clostridium asparagiforme* (CG, Firmicutes). This community also contained
186 the target pathogens *S. typhimurium* (ST) and *S. aureus* (SA). We tested whether the DNA
187 sensors could accurately report the relative abundance of the two pathogens during community
188 assembly. The 6-member community was inoculated in equal initial species proportions based on
189 absorbance at 600 nm (OD600, Day 0) and cultured anaerobically for 24 hr (Day 1). An aliquot of
190 the community was transferred to fresh media and community composition was characterized
191 following an additional 24 hr (Day 2). Based on 16S rRNA gene sequencing, the abundance of *S.*
192 *typhimurium* was similar as a function of time whereas the abundance of *S. aureus* decreased
193 over time (**Fig. 3C**).

194 We characterized the ability of the DNA sensors to accurately track the temporal trends in
195 species abundance by introducing purified community gDNA collected at different times into a
196 mixed culture of the ST-R and SA-G sensors (**Fig. 3D**). Due to the low abundance of target
197 species in the sample, a higher amount of DNA (1 μg/mL) was used for transformation. Consistent
198 with the trends based on 16S rRNA gene sequencing, the number of GFP fluorescent colonies of
199 the SA-G sensor decreased at sequential time points, whereas the number of RFP fluorescent
200 colonies of ST-R sensor were similar at sequential time points (**Fig. 3D-F**). The SA sensor
201 displayed better performance in mirroring the trend from 16S rRNA gene sequencing than the ST
202 sensor, consistent with its higher sensitivity than other sensors (**Fig. 2A,3E,F**). For the community
203 lacking *S. typhimurium* and *S. aureus*, a much smaller number of background colonies was
204 detected than in the 6-member community. This implies that the ST and SA sensors were specific
205 to the target species gDNA and did not generate false positives in the presence of the other
206 constituent community member gDNA (**Fig. S9**). In sum, our results show that accurate

207 multiplexed DNA detection can be achieved in samples derived from multi-species microbial
208 communities.
209
210 *Detection of target species without DNA extraction*
211 Specific bacterial species have been shown to release eDNA in response to environmental
212 stimuli[11], suggesting that the DNA sensor could detect species without requiring prior gDNA
213 purification. To test this possibility, we co-cultured individual DNA sensor strains with the
214 corresponding donor species with an initial OD600 0.1 of the target strain (1.22×10^8 CFU/mL,
215 1.07×10^8 CFU/mL, 3.2×10^8 CFU/mL, and 1.1×10^7 CFU/mL for *E. coli*, *S. typhimurium*, *S. aureus*,
216 *and C. difficile*, respectively) (**Fig. 4A**). Since the other species could compete with *B. subtilis*, we
217 introduced specific antibiotics (ABX) to inhibit the growth of the donor cells and enhance donor
218 eDNA release. The DNA sensor strains are resistant to the antibiotics since they harbor the
219 appropriate antibiotic resistance genes. In the presence of 100 μg/mL spectinomycin, the DNA
220 sensors displayed robust detection of *E. coli*, *S. typhimurium*, and *S. aureus* (**Fig. 4B**). The
221 addition of spectinomycin was not required for *C. difficile* detection since the growth of *C. difficile*
222 is negatively impacted by the presence of oxygen[30].
223 To confirm the transformation was mediated by the eDNA released from the target strain,
224 DNase I (1 unit/mL) was added into the co-culture. The number of transformed cells was
225 substantially lower in the presence of DNase I, indicating that DNA detection occurred via natural
226 transformation in the co-cultures (**Fig. 4B**). Antibiotic resistance is prevalent in microbiomes and
227 may not be used universally as a treatment for the donor cells. Therefore, we tested if heat
228 treatment could be used to efficiently release donor cell DNA. Incubation of *E. coli* at 90°C for 10
229 minutes substantially enhanced the EC sensor detection limit (5×10^6 CFU/mL) compared to the
230 addition of spectinomycin (**Fig. 4C**). In sum, detection of target DNA sequences directly from
231 crude samples in the absence of DNA purification could enable the deployment of the DNA
232 sensors for different future applications.
233 To evaluate the robustness of the DNA sensing function, we characterized the
234 performance of the DNA sensors for multiplexed detection of spike-in bacteria in the presence of
235 cecal contents derived from germ-free mice that were orally gavaged with a defined bacterial
236 consortium (Methods). Mouse ceca contain other bacterial species, host cells and other chemical
237 compounds (e.g. dietary factors), and thus can be used to evaluate the robustness of the DNA
238 sensor. To this end, we introduced varying amounts of *E. coli* and *S. typhimurium* into 10 mg of
239 mouse ceca, incubated these samples at 90°C for 10 minutes, and then transferred the samples
240 into a mixed culture containing the EC-G and ST-R sensors (**Fig. 4D**). Our results demonstrated

241 that both sensors can detect *E. coli* and *S. typhimurium* cells in ceca without DNA extraction. In
242 particular, the EC and ST sensors displayed a detection limit of 10^7 CFU/mL (**Fig. 4D**). In samples
243 containing a single donor species, high density of *E. coli* or *S. typhimurium* (10^8 CFU/mL) yielded
244 infrequent false positives for the multiplexed DNA detection. This suggests that further
245 optimization of the DNA sensors may be needed in complex samples containing high donor cell
246 densities that are heat treated (**Fig. S10**). In sum, we show that the cell-based DNA sensors can
247 robustly perform DNA detection of heat-treated samples that contain the target sequence.
248

249 **DISCUSSION**

250 Here we engineered the naturally competent bacterium *B. subtilis* to sense and respond to
251 specific DNA sequences. DNA sensing can be achieved for purified DNA or eDNA released from
252 pre-treated samples containing donor cells harboring the target sequence. We demonstrate that
253 DNA sensing can be sensitive and specific, and multiplexed sensing can be achieved by
254 engineering sensors with orthogonal reporter genes. Detection of species using a living cell-based
255 DNA sensor strain opens avenues for future research for versatile sensing of species and does
256 not rely on chemical or physical signals[31,32]. Since our circuit design is modular, customized
257 sensors could be constructed in the future for the detection of sequences derived from diverse
258 organisms including viruses, fungi and mammalian cells.

259 The DNA detection limit is impacted by the frequency of background mutations and could
260 be improved by reducing the background genetic mutation rate. For example, counter-selectable
261 markers[33] that can achieve lower background mutation rate could be used to optimize the strength
262 of negative growth selection. The mutation rate can also be reduced by deleting endogenous
263 genes in *B. subtilis* that promote mutagenesis such as the transcription conflict factor *mfd*[34]. The
264 reduction of mutation rates is also critical for long-term implementation of living DNA sensors in
265 the environments[35].

266 The time required for DNA detection using the cell-based DNA sensors is relatively slow
267 compared to other diagnostic methods[36]. In particular, the total time required for DNA detection
268 including transformation and selection is approximately one day, which is not suitable for certain
269 applications that require a rapid response. To reduce the detection time, directed evolution or
270 rational design of the natural competence pathway could be used for enhancing the
271 transformation efficiency of *B. subtilis*[37]. This in turn would reduce the time required for
272 transformation and detection of GFP in liquid media. In addition, some naturally competent
273 bacteria such as *Streptococcus pneumoniae* can achieve 50% transformation efficiency[38]. This
274 transformation efficiency is substantially higher than our constructed DNA sensors (10^{-5}-10^{-4}).

With a suitable chassis with a high transformation ability, GFP expression could be observed more rapidly following transformation, which requires a few hours.

The living DNA sensors have potential for *in vitro* DNA detection applications. One unique feature of the cell-based DNA sensor is the long homology within the landing pad region of the circuit, distinct from PCR-based methods that use short recognition sequences. Therefore, the target DNA sequence can be specified at the level of genes or pathways (i.e. biosynthetic gene clusters), and the DNA sensor could be used to mine such sequences from metagenomic DNA[39]. In addition, the access to NGS sequencing or multiplexed qPCR may not be widely available[36,40]. By contrast, the cell-based DNA detection is relatively simple and cost-effective. The DNA sensors may be suitable for large-scale screening with limited experimental resources. Further, *B. subtilis* sensors could be stored as spores for easy and long-term storage[41]. The metrics used in this study (sequence identity similarity and coverage) should be systematically examined using existing sequencing data and experimental characterization to elucidate sequence design rules for homologous recombination. In addition, tools from machine learning could be used to predict the impact of landing pad sequences on the fitness of *B. subtilis* to minimize any negative effects on growth rate[42].

One of the most unique aspects of this system is the potential for *in situ* DNA detection. *B. subtilis* has been shown to colonize or reside temporarily in diverse environments including soil and the mammalian gastrointestinal tract[43,44], enabling *in situ* DNA monitoring. For example, living DNA sensors could be introduced into gastrointestinal tract or plant-associated environment to monitor microbiome dynamics by sensing and recording in real-time[45]. The sensing mechanisms could be coupled to the release of antimicrobials to target specific pathogens[46]. A recent study demonstrated that the naturally competent bacterium *Acinetobacter baylyi* (*A. baylyi*) can be engineered to detect tumor DNA in the mouse colon[47], demonstrating a potential application of *in situ* DNA detection. In their study, the native CRISPR system in *A. baylyi* was exploited to detect a single mutation in the *KRAS* gene in cancer cells. Similar CRISPR systems could be incorporated into our current circuit design in the future to discriminate between single-nucleotide differences. In sum, we believe that engineering DNA-sensing bacteria could open new avenues for both *in vitro* and *in situ* applications in the future.

MATERIALS AND METHODS

Plasmid and strain construction

All DNA sensor strains were derived from *B. subtilis* PY79. Plasmids constructed in this work are listed in **Table S1**. The pAX01-comK plasmid was purchased from Bacillus Genetic

309 Stock Center (BGSC ID: ECE222) to introduce P_{xylA}-*comK* at the *lacA* locus in *B. subtilis* PY79 by
310 the selection of MLS (1 µg/mL erythromycin from Sigma-Aldrich and 25 µg/mL lincomycin from
311 Thermo Fisher Scientific) to enhance the transformation efficiency in LB[17,48]. Genes of fluorescent
312 protein GFP(Sp), mCherry, and mTagBFP were cloned from plasmid pDR111_GFP(Sp)[10] (BGSC
313 ID: ECE278), plasmid mCherry_Bsu[11] (BGSC ID: ECE756), and plasmid mTagBFP_Bsu[11] (BGSC
314 ID: ECE745) to construct fluorescent reporter plasmids pOSV00170, pOSV00455 and
315 pOSV00456, respectively. The fluorescent reporter was introduced at the *ycgO* locus by the
316 selection of 5 µg/mL chloramphenicol (MilliporeSigma). The null DNA detection plasmid
317 pOSV00157 was composed of Repressor *lacI* and IPTG-inducible toxin-antitoxin system
318 $P_{hyperspank}$-*txpA*-*ratA* and can be introduced at the *amyE* locus by the selection of 100 µg/mL
319 spectinomycin (Dot Scientific). The toxin-antitoxin system *txpA*-*ratA* was PCR amplified from *B.*
320 *subtilis* 168 gDNA.
321 *B. subtilis*, *E. coli*, *S. typhimurium*, *S. aureus* and *S. epidermidis* were all cultured at 37°C
322 in Lennox LB medium (MilliporeSigma). *C. difficile* and gut bacterial strains *A. caccae*, *B.*
323 *thetaiotaomicron*, *C. asparagiforme*, *C. hiranonis*, *and B. longum* were cultured at 37°C in YBHI
324 medium in an anaerobic chamber (Coy Laboratory). YBHI medium is Brain-Heart Infusion Medium
325 (Acumedia Lab) supplemented with 0.5% Bacto Yeast Extract (Thermo Fisher Scientific), 1
326 mg/mL D-Cellobiose (MilliporeSigma), 1 mg/mL D-maltose (MilliporeSigma), and 0.5 mg/mL L-
327 cysteine (MilliporeSigma). The gDNA of each species was extracted using DNeasy Blood &
328 Tissue Kit (Qiagen). For *S. aureus* gDNA extraction, 0.1 mg/mL Lysostaphin (MilliporeSigma) was
329 added in the pre-treatment step in combination with enzymatic lysis buffer (Qiagen). Bacterial
330 strains are listed in **Table S2**.
331 The target sequences *xdhABC* were PCR amplified from *E. coli* MG1655 gDNA (NCBI
332 Reference Sequence: NC_000913.3; Location: 3001505-3004004 and 3004005-3006504),
333 *sipBCDA* from *Salmonella enterica* serovar Typhimurium LT2 ATCC 700720 (NCBI Reference
334 Sequence: NC_003197.2; Location: 3025979-3028478 and 3028479-3030978), *hemEH* from *S.*
335 *aureus* DSM 2569 (GenBank: LHUS02000002.1; Location: 553-2770 and 2864-5638), and
336 *pheST* from *C. difficile* DSM 27147 (GenBank: FN545816.1; Location: 770923-773144 and
337 773157-775686) to construct a set of plasmids (pOSV00169, pOSV00205, pOSV00206,
338 pOSV00207, pOS00208, pOSV00292, pOSV00459 and pOSV00475) using restriction enzymes
339 BamHI-HF (New England Biolabs) and EcoRI-HF (New England Biolabs) or Golden Gate
340 Assembly Mix (New England Biolabs). DNA sequences of genetic parts are listed in **Table S3**.
341 All plasmids were constructed using *E. coli* DH5α and transformed into *B. subtilis* using
342 MC medium[49]. MC medium is composed of 10.7 g/L potassium phosphate dibasic (Chem-Impex

343 International), 5.2 g/L potassium phosphate monobasic (MilliporeSigma), 20 g/L glucose
344 (MilliporeSigma), 0.88 g/L sodium citrate dihydrate (MilliporeSigma), 0.022 g/L ferric ammonium
345 citrate (MilliporeSigma), 1 g/L Oxoid casein hydrolysate (Thermo Fisher Scientific), 2.2 g/L
346 potassium L-glutamate (MilliporeSigma), and 20 mM magnesium sulfate (MilliporeSigma).
347 Plasmids were extracted from *E. coli* DH5α using Plasmid Miniprep Kit (Qiagen) for transformation
348 of *B. subtilis*. *B. subtilis* was inoculated into MC medium and incubated at 37°C for 2 hours.
349 Extracted plasmids were added into B. subtilis cell culture and incubated 37°C for another 4 hours.
350 Transformed *B. subtilis* were selected on the LB agar plate with selective antibiotics. Double
351 crossover was verified for colonies by the replacement of a different antibiotic resistance gene at
352 the integration locus.
353
354 *DNA detection using cell-based sensors*
355 DNA sensor strain was inoculated from the -80°C glycerol stock into LB medium with 100
356 µg/mL spectinomycin and incubated at 37°C with shaking (250 rpm) for 14 hours. On the next
357 day, the OD600 of overnight culture was measured by NanoDrop One (Thermo Fisher Scientific)
358 and diluted to OD0.1 in 1 mL LB in 14 mL Falcon™ Round-Bottom Tube (Thermo Fisher Scientific)
359 supplemented with 50 mM xylose (Thermo Fisher Scientific) and 100 µg/mL spectinomycin (Dot
360 Scientific). Xylose was added to induce the competence. Spectinomycin was added to avoid
361 contamination from other bacteria but it was not required. The sample gDNA was quantified by
362 the Quant-iT dsDNA Assay Kit (Thermo Fisher Scientific) and supplemented in sensor culture
363 with known concentration. The DNA sensor culture was incubated at 37°C with shaking (250 rpm)
364 for 10 hours for transformation. Culture of transformed sensors (5 µL) was plated onto a 12-well
365 plate (Thermo Fisher Scientific) for selection. In these plates, each well contained 1 mL LB agar
366 supplemented with 2 mM IPTG (Bioline), 5 µg/mL chloramphenicol (MilliporeSigma), and MLS (1
367 µg/mL erythromycin from Sigma-Aldrich and 25 µg/mL lincomycin from Thermo Fisher Scientific).
368 Antibiotics were used in agar to avoid contamination. GFP-expressing colonies were imaged
369 using Azure Imaging System 300 (Azure Biosystems) using Epi Blue LED Light Imaging with 50
370 millisecond exposure time. CFU was counted manually. Transformation efficiency is defined as
371 the ratio of CFU on selective plates (transformed *B. subtilis* with GFP expression) to the CFU on
372 non-selective plate (total *B. subtilis*). To count CFU of total *B. subtilis*, cell culture was serially
373 diluted in phosphate-buffered saline (PBS) (Dot Scientific) and plated onto LB agar plate
374 supplemented with 5 µg/mL chloramphenicol and MLS. Since the total *B. subtilis* CFU was similar
375 for most conditions, CFU of transformed cells was used to indicate detection efficiency.

376 To quantify the sensitivity or specificity of sensors, cell culture was transferred to liquid LB
377 medium with a 1:20 dilution after transformation with serially diluted DNA from 1500 ng/mL to 1
378 ng/mL. To test the specificity towards gDNA from different strains, 100 ng/mL purified DNA was
379 used for transformation. LB medium was supplemented with 2 mM IPTG (Bioline), 5 µg/mL
380 chloramphenicol (MilliporeSigma), and MLS. Diluted cell culture was transferred to a 96-well black
381 and clear-bottom CELLSTAR® microplate with 100 uL volume in each well (Greiner Bio-One).
382 Plate was sealed with Breathe-Easy Adhesive Microplate Seals (Thermo Fisher Scientific) and
383 incubated in the SPARK Multimode Microplate Reader (TECAN) at 37°C with shaking for time-
384 series OD600 and GFP measurements. A threshold of GFP fluorescence 400 was used to
385 determine the detection time for different DNA or different concentrations. The threshold was
386 placed in the region of exponential amplification of GFP across all of the conditions where the
387 difference in the detection time between different conditions was not affected much by the choice
388 of threshold. Unpaired *t*-test was performed to determine if the detection time of specific DNA
389 concentration is different than the background without DNA (N=4). The detection limit is the lowest
390 DNA concentration with a statistical difference. A straight line was fitted to the mean detection
391 time of the four technical replicates with statistical difference versus the logarithmic DNA
392 concentration. DNA mass per mL in detection limit (1 ng, 62.5 ng, 4 ng, and 16 ng) and genome
393 size (4639675 bp, 4857450 bp, 2827820 bp, and 4153430 bp) were converted to chromosome
394 copy number by NEBioCalculator for *E. coli* (2.10×10^5), *S. typhimurium* (1.25×10^7), *S. aureus*
395 (1.38×10^6), and *C. difficile* (3.75×10^6), respectively.
396
397 *Bioinformatic analysis of the specificity of target DNA*
398 To analyze if the target sequence is conserved for the target species, we searched the 5000 bp
399 of *E. coli xdhABC*, *S. typhimurium sipBCDA*, *S. aureus hemEH*, and *C. difficile pheST* DNA
400 sequence within the same species in NCBI Nucleotide Collection Database. Nucleotide BLAST
401 was optimized for somewhat similar sequences (blastn). The search was specified to taxid 561
402 for *E. coli*, taxid 28901 for *S. enterica*, taxid 1280 for *S. aureus*, and taxid 1496 for *C. difficile*.
403 Accession date of data is 2022-08-19. For the same strains with target sequence with more than
404 95% identity similarity and 95% coverage, the percentage is 99% (N=3462), 94.3% (N=2078),
405 95.6% (N=1488), and 96.4% (N=139) for *E. coli*, *S. typhimurium*, *S. aureus*, and *C. difficile*,
406 respectively. To analyze if the target sequence is conserved in other species, the same search
407 was performed excluding the same species. Homologs with varying identity similarity and
408 coverage were found in different species for the target sequences *E. coli xdhABC* (N=5000), *S.*
409 *typhimurium sipBCDA* (N=117), *S. aureus hemEH* (N=2993), and *C. difficile pheST* (N=5000).

410 Dot plot of homology coverage and identity similarity of each hit were shown for each target
411 sequence.
412
413 *Multiplexed DNA detection using cell-based sensors*
414 Overnight cultures of DNA sensor strains EC-G, ST-R and SA-B were diluted to OD0.1, OD0.1,
415 and OD0.01, respectively, in one single culture containing 1 mL LB supplemented with 50 mM
416 xylose (Thermo Fisher Scientific) and 100 µg/mL spectinomycin (Dot Scientific). Different
417 combinations of gDNA of *E. coli*, *S. typhimurium*, and *S. aureus* (200 ng/mL each) were added
418 into the mixed culture containing three DNA sensor strains in separate 14 mL Falcon tubes
419 (Thermo Fisher Scientific). The DNA sensor strains were incubated at 37°C with shaking (250
420 rpm) for 10 hours and 5 µL of cell culture was plated onto 12-well plates (Thermo Fisher Scientific).
421 In these plates, each well contained 1 mL LB agar supplemented with 2 mM IPTG, 5 µg/mL
422 chloramphenicol, and MLS. Plates were incubated at 37°C overnight for bacterial growth.
423 On the next day, each well was imaged using Nikon Eclipse Ti-E Microscope. Brightfield
424 images were collected at 4X magnification using the built-in transilluminator of the microscope.
425 Fluorescence images were collected with the epifluorescence light source X-Cite 120 (Excelitas)
426 and standard band filter cubes including BFP (Excitation: 395/25 nm, Emission: 460/50 nm,
427 Chroma), GFP (Excitation: 470/40 nm, Emission: 525/50 nm, Nikon) and Texas Red (Excitation:
428 560/40 nm, Emission: 630/70 nm, Nikon) to image BFP, GFP, and RFP, respectively. Pixels were
429 processed with 8x8 binning when taking images. The exposure times for BFP, GFP, RFP were
430 1.5 ms, 1.5 ms, and 7 ms, respectively. Complete images of each LB agar well were generated
431 from multipoint images after scanning a 24x24 mm area. Once the full images were assembled,
432 each of the four channels were mapped to unity by the minimum and maximum pixel values of all
433 images for that channel using ImageJ. Colonies of different colors were counted manually.
434
435 *Multiplexed detection in complex DNA samples using NGS or cell-based DNA sensors*
436 Bacterial species *A. caccae*, *B. thetaiotaomicron*, *B. longum*, *C. asparagiforme*, *S. typhimurium*,
437 and *S. aureus* were inoculated from the -80°C glycerol stock into YBHI medium and incubated at
438 37°C anaerobically overnight. On the next day, cell culture of each strain was diluted to OD0.01
439 in YBHI (Passage 0) and incubated for 24 hours (Passage 1). Cell culture was diluted to OD0.1
440 into a fresh YBHI and incubated for another 24 hours (Passage 2). At each passage, cell pellet
441 was collected for DNA extraction using DNeasy Blood & Tissue Kit (Qiagen). To extract *S. aureus*
442 gDNA, 0.1 mg/mL Lysostaphin (MilliporeSigma) was added in the pre-treatment step in

443 combination with enzymatic lysis buffer (Qiagen). Purified DNA was stored at -20°C before further
444 processing for NGS or cell-based detection.
445 To use NGS for characterizing microbial community composition, the V3-V4 region of the
446 16S rRNA gene was PCR amplified from extracted DNA using custom dual-indexed primers on a
447 96-well PCR plate (detailed method described in Clark et al. Nat. Comm., 2021[50]). PCR products
448 from each well were pooled and purified using DNA Clean & Concentrator kit (Zymo). The
449 resulting library was sequenced on an Illumina MiSeq using a MiSeq Reagent Kit v2 (500-cycle)
450 to generate 2×250 paired-end reads. Sequencing data were demultiplexed using Basespace
451 Sequencing Hub's FastQ Generation program. Custom python scripts were used for further data
452 processing and sequences were mapped to the 16S rRNA reference database created using
453 consensus sequences from Sanger sequencing data of monospecies cultures. Relative
454 abundance was calculated as the read count mapped to each species divided by the total number
455 of reads of all species.
456 To use cell-based DNA sensors to detect *S. typhimurium* and *S. aureus* gDNA in the
457 community DNA, 1000 ng/mL extracted community DNA was added to the mixture of SA-G
458 (OD0.1) and ST-R sensors (OD0.1) for the multiplexed detection of *S. typhimurium* and *S. aureus*.
459 The DNA sensor strains were incubated at 37°C with shaking (250 rpm) for 10 hours and 5 µL of
460 cell culture was plated onto 12-well plates (Thermo Fisher Scientific) containing 1 mL LB agar
461 supplemented with 2 mM IPTG, 5 µg/mL chloramphenicol, and MLS. Plates were incubated at
462 37°C overnight for bacterial growth. On the next day, each well was imaged using Nikon Eclipse
463 Ti-E Microscope. Brightfield and fluorescence images of GFP and RFP were collected at 4X
464 magnification using the same procedure described in the previous section for the multiplexed
465 DNA detection. Colony numbers were counted manually. The fold change of mean colony
466 numbers compared to Day 1 was compared with the fold change of relative abundance by NGS
467 results.
468
469 *Direct detection of target species in the co-culture using cell-based sensors*
470 *E. coli*, *S. typhimurium*, and *S. aureus* and the corresponding DNA sensor strains were inoculated
471 from the -80°C glycerol stock into LB medium and incubated at 37°C with shaking (250 rpm) for
472 14 hours. *C. difficile* was separately inoculated in YBHI medium and incubated in an anaerobic
473 chamber (Coy Laboratory). On the next day, cell culture of sensor and target strain were diluted
474 to an OD0.1 each in a single culture containing 1 mL LB supplemented with 50 mM xylose with
475 or without 100 µg/mL spectinomycin in 14 mL Falcon tubes (Thermo Fisher Scientific). Overnight
476 culture of target strains (14 hr) was diluted in PBS (Dot Scientific) and plated onto LB agar plates

or YBHI agar plates to determine the initial CFU of target bacteria in the co-culture (OD0.1), which was 1.22×10^8 CFU/mL, 1.07×10^8 CFU/mL, 3.2×10^8 CFU/mL, and 1.1×10^7 CFU/mL for *E. coli*, *S. typhimurium*, *S. aureus*, *and C. difficile*, respectively. Sensor and target strains were co-cultured at 37°C with shaking (250 rpm) for 10 hours, and 5 µL of cell culture was plated onto 12-well plates (Thermo Fisher Scientific). Each well contained 1 mL LB agar supplemented with 2 mM IPTG, 5 µg/mL chloramphenicol, and MLS for the selection of transformed *B. subtilis*. The 12-well plates were incubated overnight at 37°C. On the next day, fluorescent colonies were imaged by Azure Imaging System 300 (Azure Biosystems) using the Epi Blue LED Light Imaging with 50 millisecond exposure time. Colonies with GFP expression were counted manually. To determine if the detection of target bacteria was via transformation, 1 µL (1 unit) DNase I (Thermo Fisher Scientific) was added to the 1 mL co-culture. One unit of DNase I can completely degrade 1 µg of plasmid DNA in 10 min at 37°C according to the manufacturer's specification.

To improve the detection efficiency, overnight culture of *E. coli* was incubated at 90°C in digital dry baths/block heaters (Thermo Fisher Scientific) for 10 min and placed on ice for 3 min before being transferred to the sensor culture containing 1 mL LB, 50 mM xylose, and OD0.1 of sensor strain for detection. Spectinomycin was not used in the sensor culture for heat-treated samples. To test the multiplexed detection of *E. coli* and *S. typhimurium* in mice cecal samples, 10 mg cecal samples were first resuspended with 100 µL LB and 50 mM xylose in 1.7 mL Eppendorf tubes (Dot Scientific). Different amounts of overnight culture of *E. coli* and *S. typhimurium* were spiked into cecal samples. The cecal samples were then incubated at 90°C for 10 min and sat on ice for 3 min before being transferred to the mixed culture of EC-G and ST-R sensors (1 mL LB, 50 mM xylose, OD0.1 of EC-G, and OD0.1 of ST-R) for multiplexed detection. Cecal samples were collected from germ-free mouse experiments following protocols approved by the University of Wisconsin-Madison Animal Care and Use Committee. Briefly, 8-week-old C57BL/6 gnotobiotic male mice (wild-type) were inoculated with 8 human gut bacteria - *Dorea formicigenerans*, *Coprococcus comes*, *Anaerostipes caccae*, *Bifidobacterium longum*, *Bifidobacterium adolescentis*, *Bacteroides vulgatus*, *Bacteroides caccae*, and *Bacteroides thetaiotaomicron* via oral gavage. After four weeks of colonization, mice were euthanized for cecal sample collection. Cecal samples were stored at -80°C and thawed for the use in multiplexed detection of spike-in *E. coli* and *S. typhimurium*.

ACKNOWLEDGEMENTS

We would like to thank Dr. Scott Coyle and Zhejing Xu for the use of Nikon Eclipse Ti2-E Microscope. We also want to thank Yiyi Liu for providing the mice cecal samples. This work was

supported by the Defense Advanced Research Projects Agency (DARPA) Grant HR0011-19-2-0002.

AUTHOR CONTRIBUTIONS

Y.Y.C., B.M.B. and O.S.V. conceived the research. Y.Y.C., and Z.C. performed the experiments. Y.Y.C., Z.C., and X.C. constructed the sensors. T.D.R. processed the microscopic images. T.G.F. assisted with the strain construction. Y.Y.C. and O.S.V. wrote the manuscript. O.S.V. secured the funding.

COMPETING INTERESTS

O.S.V., Y.Y.C., and Z.C. have filed a provisional application with the US Patent and Trademark Office on this work (U.S. Provisional Patent No. 63/290,442). The other authors declare that they have no competing interests.

REFERENCES

1. Brophy, J. A. N. & Voigt, C. A. Principles of genetic circuit design. *Nature Methods* **11**, (2014).
2. Levskaya, A. *et al.* Engineering Escherichia coli to see light. *Nature* **438**, 441–442 (2005).
3. Bourdeau, R. W. *et al.* Acoustic reporter genes for noninvasive imaging of microorganisms in mammalian hosts. *Nature* **553**, (2018).
4. You, L., Cox, R. S., Weiss, R. & Arnold, F. H. Programmed population control by cell-cell communication and regulated killing. *Nature* **428**, (2004).
5. Danino, T., Mondragón-Palomino, O., Tsimring, L. & Hasty, J. A synchronized quorum of genetic clocks. *Nature* **463**, (2010).
6. Liao, M. J., Din, M. O., Tsimring, L. & Hasty, J. Rock-paper-scissors: Engineered population dynamics increase genetic stability. *Science (80-.).* **365**, (2019).
7. Saeidi, N. *et al.* Engineering microbes to sense and eradicate Pseudomonas aeruginosa, a human pathogen. *Mol. Syst. Biol.* **7**, (2011).
8. Borrero, J., Chen, Y., Dunny, G. M. & Kaznessis, Y. N. Modified lactic acid bacteria detect and inhibit multiresistant enterococci. *ACS Synth. Biol.* **4**, (2015).
9. Scott, S. R. & Hasty, J. Quorum Sensing Communication Modules for Microbial Consortia. *ACS Synth. Biol.* **5**, (2016).
10. Sexton, J. T. & Tabor, J. J. Multiplexing cell-cell communication. *Mol. Syst. Biol.* **16**, (2020).
11. Ibáñez de Aldecoa, A. L., Zafra, O. & González-Pastor, J. E. Mechanisms and regulation of extracellular DNA release and its biological roles in microbial communities. *Front. Microbiol.* **8**, 1390 (2017).
12. Johnston, C., Martin, B., Fichant, G., Polard, P. & Claverys, J.-P. Bacterial transformation: distribution, shared mechanisms and divergent control. *Nat. Rev. Microbiol.* **12**, 181–196 (2014).
13. Carrasco, B., Serrano, E., Martín-González, A., Moreno-Herrero, F. & Alonso, J. C. Bacillus subtilis MutS modulates RecA-mediated DNA strand exchange between divergent DNA sequences. *Front. Microbiol.* (2019). doi:10.3389/fmicb.2019.00237
14. Popp, P. F., Dotzler, M., Radeck, J., Bartels, J. & Mascher, T. The Bacillus BioBrick Box 2.0: Expanding the genetic toolbox for the standardized work with Bacillus subtilis. *Sci. Rep.* **7**, (2017).
15. Dubnau, D. Genetic competence in Bacillus subtilis. *Microbiological Reviews* (1991).
16. Ji, M. *et al.* Engineering Bacillus subtilis ATCC 6051a for the production of recombinant catalases. *J. Ind. Microbiol. Biotechnol.* **48**, (2021).

562 17. Zhang, X.-Z. & Zhang, Y.-H. P. Simple, fast and high-efficiency transformation system for
563 directed evolution of cellulase in Bacillus subtilis. *Microb. Biotechnol.* **4**, 98–105 (2011).
564 18. Silvaggi, J. M., Perkins, J. B. & Losick, R. Small untranslated RNA antitoxin in Bacillus
565 subtilis. *J. Bacteriol.* **187**, (2005).
566 19. Maamar, H. & Dubnau, D. Bistability in the Bacillus subtilis K-state (competence) system
567 requires a positive feedback loop. *Mol. Microbiol.* (2005). doi:10.1111/j.1365-
568 2958.2005.04592.x
569 20. Xi, H., Schneider, B. L. & Reitzer, L. Purine catabolism in Escherichia coli and function of
570 xanthine dehydrogenase in purine salvage. *J. Bacteriol.* **182**, (2000).
571 21. McClelland, M. *et al.* Complete genome sequence of Salmonella enterica serovar
572 Typhimurium LT2. *Nature* **413**, (2001).
573 22. Stabler, R. A. *et al.* Comparative genome and phenotypic analysis of Clostridium difficile
574 027 strains provides insight into the evolution of a hypervirulent bacterium. *Genome Biol.*
575 **10**, (2009).
576 23. Soni, I., Chakrapani, H. & Chopra, S. Draft genome sequence of methicillin-sensitive
577 Staphylococcus aureus ATCC 29213. *Genome Announc.* **3**, (2015).
578 24. Tucker, S. C. & Galán, J. E. Complex function for SicA, a Salmonella enterica serovar
579 typhimurium type III secretion-associated chaperone. *J. Bacteriol.* **182**, (2000).
580 25. Lobo, S. A. L. *et al.* Staphylococcus aureus haem biosynthesis: Characterisation of the
581 enzymes involved in final steps of the pathway. *Mol. Microbiol.* **97**, (2015).
582 26. Beyer, D. *et al.* New Class of Bacterial Phenylalanyl-tRNA Synthetase Inhibitors with High
583 Potency and Broad-Spectrum Activity. *Antimicrob. Agents Chemother.* **48**, (2004).
584 27. Ibekwe, A. M. & Grieve, C. M. Detection and quantification of Escherichia coli O157:H7 in
585 environmental samples by real-time PCR. *J. Appl. Microbiol.* **94**, (2003).
586 28. Jung, B. Y., Suk, C. J. & Chang, H. K. Development of a rapid immunochromatographic
587 strip for detection of Escherichia coli O157. *J. Food Prot.* **68**, (2005).
588 29. Serrano, E., Ramos, C., Alonso, J. C. & Ayora, S. Recombination proteins differently
589 control the acquisition of homeologous DNA during Bacillus subtilis natural chromosomal
590 transformation. *Environ. Microbiol.* **23**, (2021).
591 30. Giordano, N. *et al.* Cysteine desulfurase IscS2 plays a role in oxygen resistance in
592 Clostridium difficile. *Infect. Immun.* **86**, (2018).
593 31. Papagiannopoulou, C., Parchen, R., Rubbens, P. & Waegeman, W. Fast Pathogen
594 Identification Using Single-Cell Matrix-Assisted Laser Desorption/Ionization-Aerosol
595 Time-of-Flight Mass Spectrometry Data and Deep Learning Methods. *Anal. Chem.* **92**,
596 (2020).
597 32. Nißler, R. *et al.* Remote near infrared identification of pathogens with multiplexed
598 nanosensors. *Nat. Commun.* **11**, (2020).
599 33. Dong, H. & Zhang, D. Current development in genetic engineering strategies of Bacillus
600 species. *Microb. Cell Fact.* **13**, (2014).
601 34. Ragheb, M. N., Merrikh, C., Browning, K. & Merrikh, H. Mfd regulates RNA polymerase
602 association with hard-to-transcribe regions in vivo, especially those with structured RNAs.
603 *Proc. Natl. Acad. Sci. U. S. A.* **118**, (2020).
604 35. Renda, B. A., Hammerling, M. J. & Barrick, J. E. Engineering reduced evolutionary
605 potential for synthetic biology. *Molecular BioSystems* **10**, (2014).
606 36. Law, J. W. F., Mutalib, N. S. A., Chan, K. G. & Lee, L. H. Rapid methods for the detection
607 of foodborne bacterial pathogens: Principles, applications, advantages and limitations.
608 *Front. Microbiol.* **5**, (2014).
609 37. Rahmer, R., Heravi, K. M. & Altenbuchner, J. Construction of a super-competent Bacillus
610 subtilis 168 using the PmtlA-comKS inducible cassette. *Front. Microbiol.* **6**, 1431 (2015).
611 38. Kurushima, J. *et al.* Unbiased homeologous recombination during pneumococcal
612 transformation allows for multiple chromosomal integration events. *Elife* **9**, (2020).

39. Burian, J. *et al.* High-throughput retrieval of target sequences from complex clone libraries using CRISPRi. *Nat. Biotechnol. 2022* 1–5 (2022). doi:10.1038/s41587-022-01531-8
40. Gu, W., Miller, S. & Chiu, C. Y. Clinical Metagenomic Next-Generation Sequencing for Pathogen Detection. *Annu. Rev. Pathol. Mech. Dis.* **14**, (2019).
41. Liu, F., Li, J., Zhang, T., Chen, J. & Ho, C. L. Engineered Spore-Forming Bacillus as a Microbial Vessel for Long-Term DNA Data Storage . *ACS Synth. Biol.* (2022). doi:10.1021/ACSSYNBIO.2C00291/SUPPL_FILE/SB2C00291_SI_001.PDF
42. Gomes, A. L. C. *et al.* Genome and sequence determinants governing the expression of horizontally acquired DNA in bacteria. *ISME J.* **14**, (2020).
43. Beauregard, P. B., Chai, Y., Vlamakis, H., Losick, R. & Kolter, R. Bacillus subtilis biofilm induction by plant polysaccharides. *Proc. Natl. Acad. Sci. U. S. A.* **110**, (2013).
44. Tam, N. K. M. *et al.* The intestinal life cycle of Bacillus subtilis and close relatives. *J. Bacteriol.* (2006). doi:10.1128/JB.188.7.2692-2700.2006
45. Sheth, R. U. & Wang, H. H. DNA-based memory devices for recording cellular events. *Nature Reviews Genetics* (2018). doi:10.1038/s41576-018-0052-8
46. González, L. M., Mukhitov, N. & Voigt, C. A. Resilient living materials built by printing bacterial spores. *Nat. Chem. Biol.* **16**, (2020).
47. Cooper, R. M. *et al.* Engineered bacteria detect tumor DNA in vivo. *bioRxiv* 2021.09.10.459858 (2021). doi:10.1101/2021.09.10.459858
48. Cheng, Y.-Y., Papadopoulos, J. M., Falbel, T., Burton, B. M. & Venturelli, O. S. Efficient plasmid transfer via natural competence in a synthetic microbial community. *bioRxiv* 2020.10.19.342733v2 (2020). doi:10.1101/2020.10.19.342733
49. Konkol, M. A., Blair, K. M. & Kearns, D. B. Plasmid-encoded comi inhibits competence in the ancestral 3610 strain of Bacillus subtilis. *J. Bacteriol.* **195**, (2013).
50. Clark, R. L. *et al.* Design of synthetic human gut microbiome assembly and butyrate production. *Nat. Commun.* **12**, (2021).

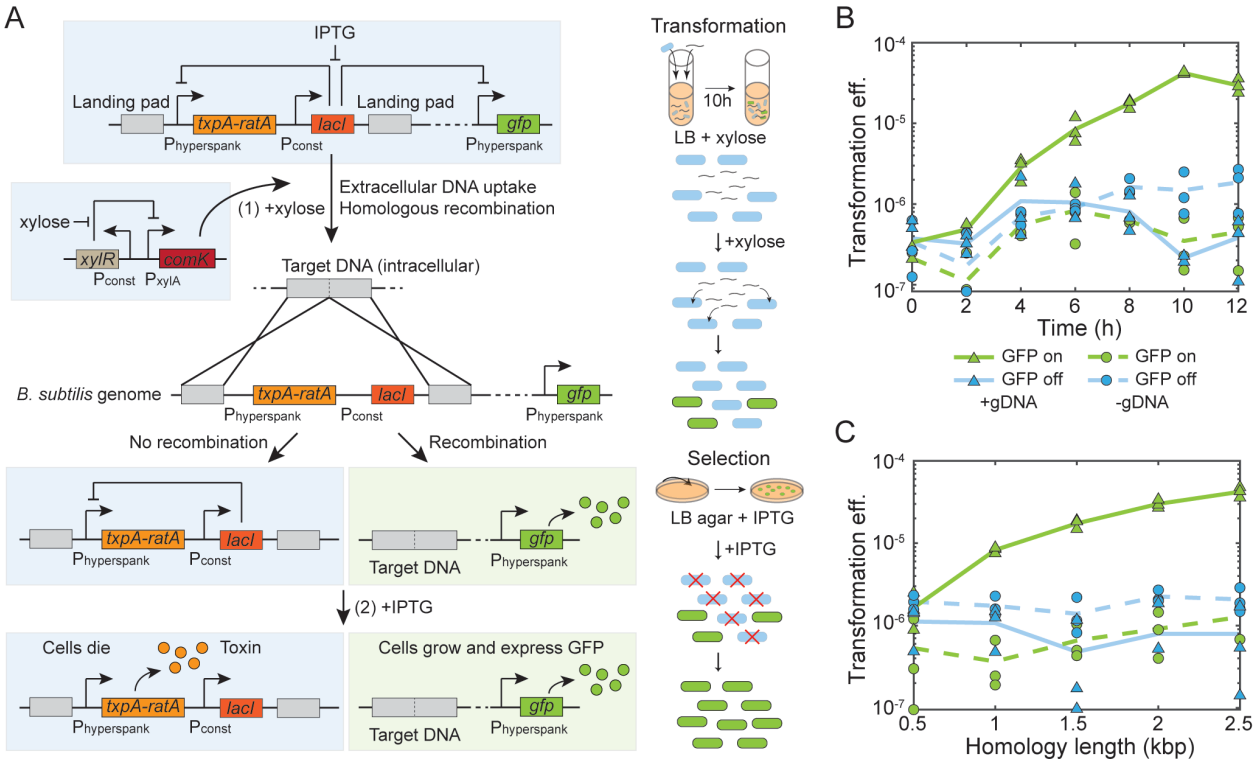

Figure 1. Construction and characterization of a living cell-based DNA sensor. (A) Schematic of a synthetic genetic circuit constructed in *B. subtilis* that allows the recognition of specific extracellular DNA. The xylose-inducible master regulator of competence *comK* can induce B. subtilis to take up eDNA and undergo homologous recombination. The toxin-antitoxin system *txpA-ratA* and a distal fluorescent reporter *gfp* are regulated by the repressor *lacI*. The target sequences are split into two and introduced to the flanking region of toxin and repressor as landing pad. In the presence of target DNA, homologous recombination can remove the toxin and the repressor, so a sub-population of transformed cells can express GFP. To select the transformed cells, IPTG can induce the toxin to kill the non-transformed cells, so the fluorescence can be enhanced by the growth of transformed cells. **(B)** Line plot of transformation efficiency versus transformation time for the *E. coli* DNA sensor in the presence (solid lines) or absence (dashed lines) of 100 ng/mL purified *E. coli* gDNA as input. The colonies on agar plate can be fluorescent (green) or not (blue), coming from background escape mutation or transformation. Transformation efficiency as output is optimal at 10 hr. Data points represent three biological replicates and lines are the average of the replicates. **(C)** Line plot of transformation efficiency versus homology length of *E. coli xdhABC* operon on each side of flanking region in the circuit in the presence (solid lines) or absence (dashed lines) of 100 ng/mL purified *E. coli* gDNA as input. The colonies on agar plate can be fluorescent (green) or not (blue). The transformation efficiency as output is above the background escape mutation when the homology length is equal or greater than 1 kbp and increases as the homology increases. Data points represent three biological replicates and lines are the average of replicates.

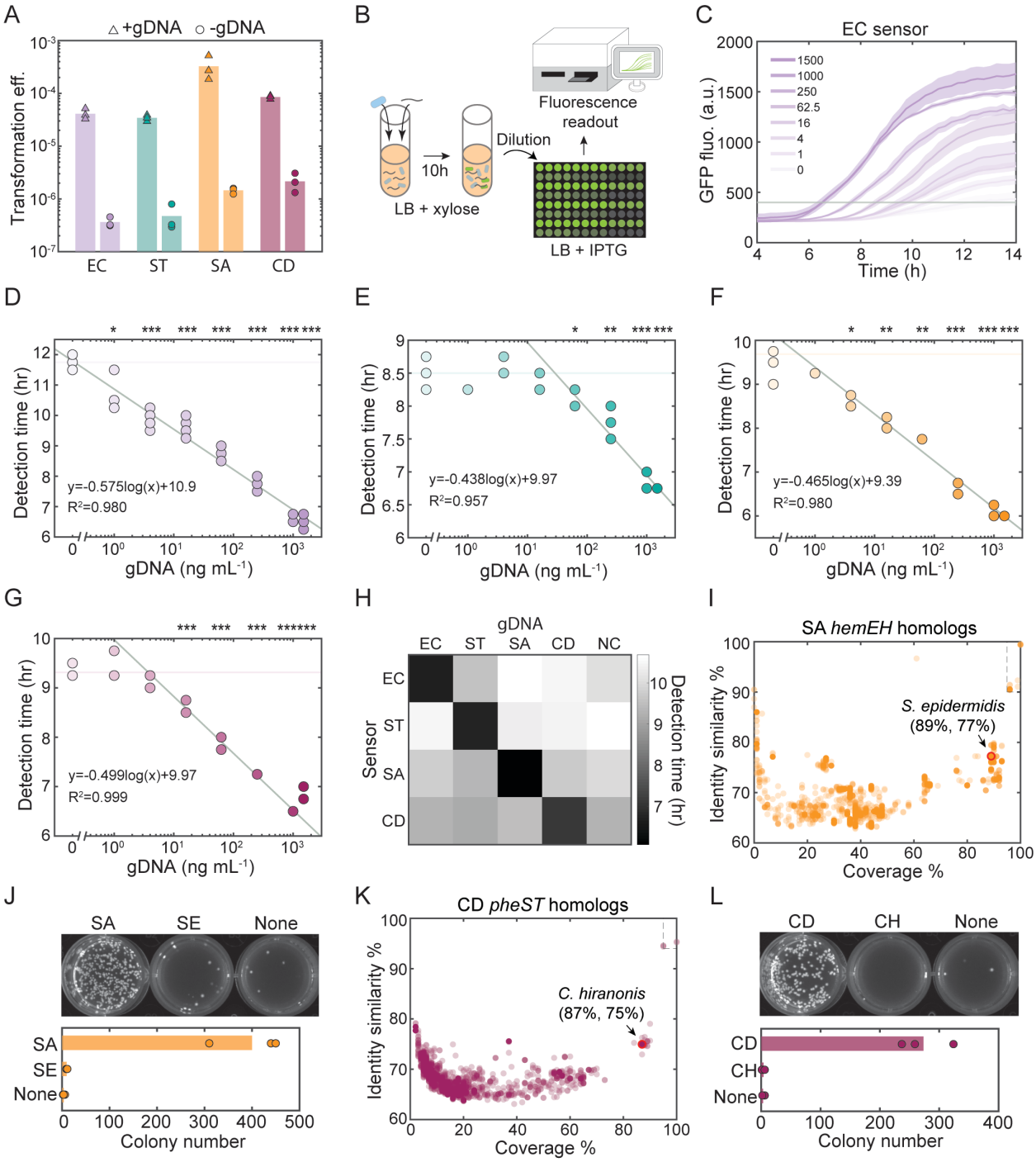

Figure 2. Cell-based sensors can detect DNA from diverse bacteria with high sensitivity and specificity. (A) Transformation efficiency of DNA sensors that can detect *E. coli* (EC sensor), *S. typhimurium* (ST sensor), *S. aureus* (SA sensor), or *C. difficile* (CD sensor) in the presence of 100 ng/mL extracted gDNA (triangles) or no gDNA (circles) at 10 hr. Bar represents the average of three biological replicates. **(B)** Schematic of experimental procedure for quantifying GFP expression after transformation. **(C)** Time-series measurements of GFP expression of EC sensor in liquid medium after the transformation of varying *E. coli* gDNA concentrations (ng/mL). A threshold of GFP 400 was used to determine the detection time for each gDNA concentration.

675 Line is the average of four technical replicates and the shaded region represents one standard
676 deviation from the average. Detection time versus gDNA concentration for **(D)** EC sensor, **(E)** ST
677 sensor, **(F)** SA sensor, and **(G)** CD sensor. Horizontal line (pale color) is the background GFP
678 fluorescence in the absence of gDNA. Unpaired *t*-test was performed to determine if the detection
679 time with specific DNA concentration is different from the background fluorescence, and *, **, and
680 *** denote *p*-values < 0.05, 0.01 and 0.001, respectively. A straight line was fitted to the
681 transformation efficiencies versus logarithmic gDNA concentrations with statistical differences.
682 The slope of the fitted line was determined by the cell growth of *B. subtilis*, while the intercept was
683 determined by the background escape mutation. The coefficient of determination R^2 shows the
684 goodness of the fit. **(H)** Detection time of each DNA sensor after the transformation of 100 ng/mL
685 gDNA of a given donor species or no gDNA representing the negative control (NC). Each sensor
686 expressed GFP 3~4 hours earlier in the presence of gDNA from its target strain than gDNA from
687 other strains. Data are the average of four technical replicates. **(I)** Nucleotide BLAST search of
688 5000 bp *S. aureus hemEH* in the NCBI database. Each circle represents a homolog found in
689 species other than *S. aureus* and its coverage and identity similarity. A closely related human
690 commensal strain *S. epidermidis* with high coverage and similarity was selected for specificity test
691 using SA sensor. The region within dashed lines indicates *Staphylococcus* species that could be
692 recognized by the SA sensor based on their high similarity of homologous sequences. **(J)**
693 Comparison of colony numbers of SA sensor (with GFP expression) on selective agar plate after
694 the transformation of 100 ng/mL *S. aureus* gDNA, 100 ng/mL *S. epidermidis* gDNA or no gDNA.
695 Bar represents the average of three technical replicates. SA sensor can distinguish between *S.*
696 *aureus* and closely related *S. epidermidis*. **(K)** Nucleotide BLAST search of 5000 bp *C. difficile*
697 *pheST* in the NCBI database. Each circle represents a homolog found in species other than *C.*
698 *difficile* and its coverage and identity similarity. A closely related human commensal strain *C.*
699 *hiranonis* with high coverage and similarity was selected for specificity test using CD sensor. The
700 region within dashed lines indicates *Clostridium* species that could be recognized by the CD
701 sensor based on their high similarity of homologous sequences. **(L)** Comparison of colony
702 numbers of CD sensor (with GFP expression) on selective agar plate after the transformation of
703 100 ng/mL *C. difficile* gDNA, 100 ng/mL *C. hiranonis* gDNA or no gDNA. Bar represents the
704 average of three technical replicates. CD sensor can distinguish between *C. difficile* and closely
705 related *C. hiranonis*.

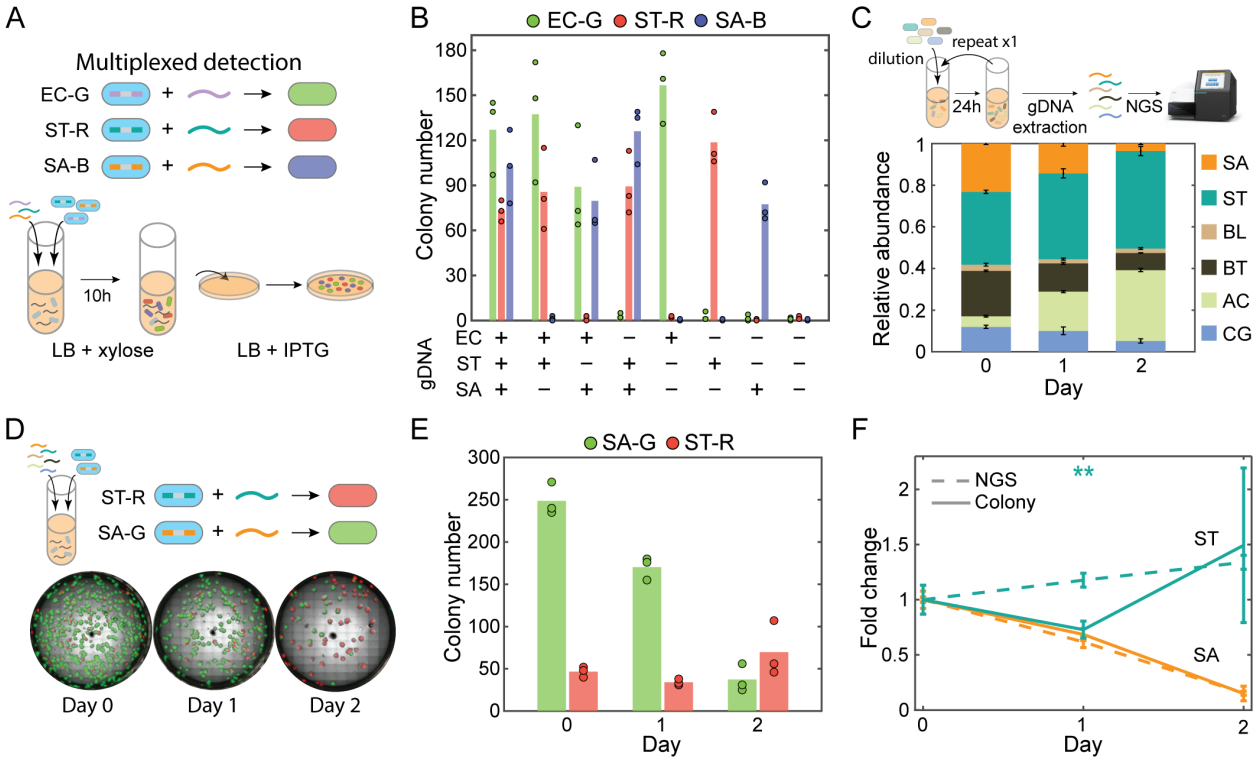

Figure 3. DNA sensors can perform multiplexed detection in complex DNA samples. (A) Schematic of experimental procedure using the living cell-based DNA sensors for multiplexed detection. *E. coli*, *S. typhimurium*, and *S. aureus* DNA sensors were labeled GFP (EC-G), RFP (ST-R) and BFP (SA-B) for detection, respectively. Three sensors and gDNA extracted from different strains were mixed into liquid medium for transformation. Transformed cells were selected on agar and colonies expressing GFP, RFP and BFP can indicate the presence of target DNA. **(B)** Numbers of GFP, RFP, or BFP-expressing colonies on agar plate after the transformation of different combinations of gDNA extracted from *E. coli*, *S. typhimurium*, or *S. aureus*. Sensors can report the presence of target DNA for all 8 different combinations. Bar represents the average of three technical replicates. **(C)** Relative abundance of six species in a synthetic gut microbial community composed of *S. aureus* (SA), *S. typhimurium* (ST), *Bifidobacterium longum* (BL), *Bacteroides thetaiotaomicron* (BT), *Anaerostipes caccae* (AC), and *Clostridium asparagiforme* (CG) over two days of growth. The six bacteria were co-cultured in liquid medium anaerobically for 24 hours and cell culture was diluted in fresh medium once for species to continue their competition. 16S rRNA gene of each strain was PCR amplified and sequenced using NGS to determine the relative abundance over time. Bar represents the average of three technical replicates of 16S rRNA sequencing. *S. typhimurium* abundance remained similar while *S. aureus* abundance decreased over time. **(D)** Representative fluorescence images of transformed SA-G and ST-R sensors on selective agar plate after the transformation of gDNA extracted from the bacterial community in a mixture of SA-G and ST-R sensors. The numbers of green and red colonies indicate the abundance of *S. aureus* and *S. typhimurium* in gut microbial community, respectively. **(E)** Numbers of GFP or RFP-expressing colonies on agar plate after the transformation of gDNA extracted from the community at different time. Green colonies (SA-G) decreased over time while red colonies (ST-R) remained similar numbers. Bar represents the average of three technical replicates. **(F)** Fold change of *S. aureus* and *S. typhimurium* abundance compared to day 1 characterized by NGS (dashed line) or cell-based detection method (solid line).

Measurements of *S. aureus* (orange) by the two methods were similar, while characterization of *S. typhimurium* (green) by the cell-based detection method had larger variability than NGS and the data at Day 1 were statistically different determined by Unpaired *t*-test (*p*-values = 0.0015).

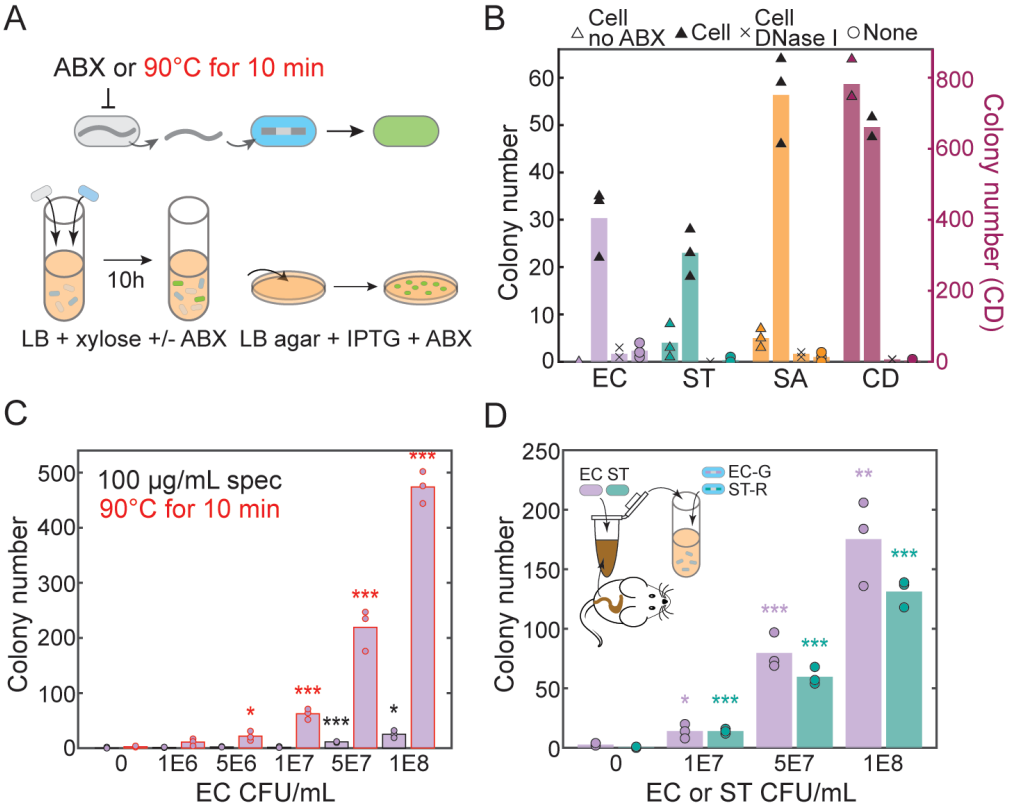

Figure 4. DNA sensors can directly detect target species without DNA extraction. **(A)** Schematic of experimental design for the detection in the co-culture. Sensor and target strain were co-cultured in liquid medium with or without selective antibiotics (ABX) to determine the effect of antibiotic for detection. Heat-killed target strain was also tested without the use of antibiotics in the co-culture. The cell culture was plated on IPTG and ABX (1 µg/mL erythromycin and 25 µg/mL lincomycin) agar plates to select for transformed *B. subtilis*. **(B)** Colony numbers of transformed EC, ST, SA, and CD sensors co-cultured with (1) target cell, (2) target cell and 100 µg/mL spectinomycin (spec), (3) target cell, spec, and 1 unit/mL DNase I, and (4) spec only. Spectinomycin can enhance the detection of *E. coli*, *S. typhimurium*, and *S. aureus,* but it is not required for *C. difficile* detection. Addition of DNase I in the co-culture reduced the numbers of transformed sensors significantly. **(C)** Colony number of transformed EC sensor co-cultured with *E. coli* using spectinomycin or co-cultured with heat-treated *E. coli*. Heat treatment of *E. coli* can significantly increase the number of transformed EC sensor. **(D)** Colony number of transformed EC-G and ST-R sensors co-cultured with heat-treated mice cecal samples spiked in with different amounts of *E. coli* and *S. typhimurium*. Both sensors can detect the presence of target strains in cecal samples. Unpaired *t*-test was performed to determine if the colony number is different from no cell condition, and *, **, and *** denote *p*-values < 0.05, 0.01, and 0.001, respectively. Bar represents the average of three technical replicates.