A broad-host-range event detector: expanding and quantifying performance across bacterial species

Nymul Khan[1], Enoch Yeung[2], Yuliya Farris[1], Sarah J. Fansler[1], Hans C. Bernstein[3,4]*

[1]Biological Sciences Division, Pacific Northwest National Laboratory, Richland, WA, USA; [2] Department of Mechanical Engineering, University of California, Santa Barbara, CA, USA; [3]Faculty of Biosciences, Fisheries and Economics, UiT - The Arctic University of Norway, Tromsø, Norway; [4]The Arctic Centre for Sustainable Energy, UiT - The Arctic University of Norway, Tromsø, Norway

***Correspondence**: Hans C. Bernstein, UiT-The Arctic University of Norway, BFE-fak., Postboks 6050 Langnes, 9037 Tromsø, NORWAY*;Phone: +47 776 46 114; Emial:* Hans.C.Bernstein@uit.no

Running Title: Building and quantifying multi-host synthetic gene devices.

Keywords: Genetic Device, Integrase, Genetic Memory, Chassis, Pseudomonas, Escherichia coli Nissle

ABSTRACT

Modern microbial biodesign relies on the principle that well-characterized genetic parts can be reused and reconfigured for different functions. However, this paradigm has only been successful in a limited set of hosts, mostly comprised from common lab strains of *Escherichia coli*. It is clear that new applications – such as chemical sensing and event logging in complex environments – will benefit from new host chassis. This study quantitatively compared how a chemical event logger performed across multiple microbial species. An integrase-based sensor and memory device was operated by two representative soil Pseudomonads – *Pseudomonas fluorescens* SBW25 and *Pseudomonas putida* DSM 291. Quantitative comparisons were made between these two non-traditional hosts and two bench-mark *Escherichia coli* chassis including the probiotic Nissle 1917 and common cloning strain DH5α. The performance of sensor and memory components changed according to each host, such that a clear chassis effect was observed and quantified. These results were obtained via fluorescence from reporter proteins that were transcriptionally fused to the integrase and down-stream recombinant region and via data-driven kinetic models. The *Pseudomonads* proved to be acceptable chassis for the operation of this event logger, which outperformed the common *E. coli* DH5α in many ways. This study advances an emerging frontier in synthetic biology that aims to build broad-host-range devices and understand the context by which different species can execute programmable genetic operations.

INTRODUCTION

Synthetic biology is built on the concept that complex biological behaviors can be programmed using relatively simple modules of biological parts. While the field of microbial biodesign has seen major advances, the overwhelming majority of parts have only been tested in model organisms. To date, we know little about how even our most standard genetic devices will perform in microbial hosts beyond common laboratory strains of *Escherichia coli* or *Saccharomyces cerevisiae*. This represents a major knowledge gap and limitation in the field. While useful for the development and demonstration of capabilities under stable laboratory conditions, these species do not survive well in many real-world applications. Most traditional microbial hosts have limited metabolic potential, preferring substrates such as simple sugars that are typically not available in environments relevant to the next generation of synthetic biology applications such as event detection within soils, built environments or the human gut. Therefore, programmable genetic devices must be expanded into new, non-traditional chassis that are already evolved to operate in complex, dynamic environments.

One of the most common biodesign principles is that well-characterized genetic parts – e.g., promoters, UTRs and transcription factors – can be reused and reconfigured to program different functions. Some of the benchmark examples are given by the toggle switch (1), repressilator (2) and previous demonstrations of integrase-based recording devices (3-6); all of which were exclusively demonstrated in *E. coli.* These devices have laid the foundation for more applied microbial sensor-regulator-actuator devices that have been developed to detect/report signals from the mammalian gut (7,8) and chemical threats (9,10); yet even these advanced examples relied solely on the genetic tractability of *E. coli*. Synthetic biologists are keen to harness new non-traditional hosts such as *Pseudomonads* (11-13); yet, successful transplantation of broad-host-range genetic devices across multiple bacterial species has remained elusive, until now.

Here we present a study that demonstrates how a relatively simple chemical event logger performs across multiple microbial hosts. We chose to comparatively quantify each component of an integrase-based sensor/memory device between two *Pseudomonas* species – *Pseudomonas fluorescens* SBW25 (*Pf*) and *Pseudomonas putida* DSM 291(*Pp*) – along with two more standard *Escherichia coli* strains including the probiotic Nissle 1917 (*EcN*) and common cloning strain DH5α (*Ec*). The event detector was expressed from each host as the same sequence on an identical broad-host range expression vector. Here, we show that a genetically identical chemical event logging device can be ported across multiple species, including two Pseudomonads that open new chemical sensing/logging applications soil and

66 plant-associated environments. The performance for each component of the device depended on the host – it was subject
67 to a strong chassis effect. Hence, study presents a new broad-host range event logging system, which advances to a
68 rapidly growing frontier in synthetic biology aimed at engineering devices that can function across multiple species and
69 environments (14,15).
70
71 **MATERIALS AND METHODS**
72 **Bacterial strains and cultivation.** The bacterial strains used in this study includes *Escherichia coli* DH5α (New England
73 Biolabs), *Escherichia coli* Nissle 1917 (isolated from probiotic Mutaflor capsule), *Pseudomonas fluorescens* SBW25
74 and *Pseudomonas putida* DSM 291 (DSMZ). All bacteria were cultured in Lauria-Bertani medium at 30 ^{o}C.
75
76 **Transformation.** *E. coli DH5α* was transformed using standard chemical transformation protocol. For the other species,
77 electrocompetent cells were prepared as follows: overnight cultures were diluted 1:100 into 200 mL LB medium and
78 grown to OD600 nm of about 0.3-0.4 (mid-log phase); cultures were harvested and spun down in four 50 mL centrifuge
79 tubes at 5000 x g and the supernatant was discarded; cell pellets were resuspended in 15% glycerol, combined into one
80 50 mL centrifuge tube and collected via centrifugation at 5000 x g. This wash cycle was repeated twice and the final cell
81 pellet was resuspended in 1 mL 15% glycerol for electroporation. The cells were then transformed by electroporation at
82 12500 V/cm (200 Ω and 25 μF) in 1 mm cuvettes. The entire protocol was successfully carried out at room temperature
83 as outlined in the Tu et al. 2016 study (16). The efficiency, in general, was found to be higher in the room-temperature
84 methods than in conventional ice-cold methods.
85
86 **Plate reader and cytometry assays.** For each bacterial species, three positive transformants were grown, passaged twice
87 at 30 ^{o}C and then assayed in a 24-well plate at eight different IPTG concentrations (0, 0.01, 0.05, 0.1, 0.2, 0.5, 0.8 and 1
88 mM). Each well of the plate contained 1.8 mL LB + kanamycin (50 μg/mL) + IPTG. A Synergy H1 (Biotek, Winooski,
89 VT) was used as the fluorescent plate reader for all assays. 1 μL samples were collected from each of the wells at various
90 stages of growth and analyzed via flow cytometry (Novocyte, ACEA biosciences, San Diego, CA). Simultaneously, 100
91 μL samples were also collected and frozen at -80 ^{o}C. Plasmid DNA was later extracted from the frozen samples and used
92 for real time qPCR assays to measure the fraction of device flipped.

93

94 **Real time quantitative PCR.** Real time PCR was performed by ARQ Genetics LLC (Bastrop, TX) on the BioRad
95 CFX384 Real Time System (BioRad, Hercules, CA) using assays specific for each plasmid. All of the plasmid DNA was
96 extracted with a Zyppy – 96 Plasmid Miniprep kit (Zymo, Irvine, CA) following the next steps: all of the stains were
97 grown at 30 °C and harvested at 3-5 h intervals; the DNA was quantified by performing PicoGreen assay on the Biotek
98 Synergy H1 (Biotek, Winooski, VT) and reactions were diluted to matching concentrations. Each reaction within multi-
99 plate wells contained 5 μL of TaqMan Universal Master Mix II (Applied Biosystems, Waltham, MA), 2 μL of each
00 sample template and 0.5 μL of each specific plasmid assay in a reaction volume of 10 μL. Cycling conditions were as
01 follows: 95 °C for 10 min for polymerase activation, followed by 40 cycles of 95 °C for 15 seconds and 63 °C for 1 min.
02 Data analysis was performed using CFX Manager software from BioRad, version 3.1. The experimental Cq (cycle
03 quantification) was calibrated against the standard curve for each plasmid orientation.

04

05 **Numerical simulation.** The system of ODEs were solved numerically using the 'deSolve' package (17) in R (18). The
06 results were fitted to experimental data to estimate the five kinetic rate constants (P_A, P_B, P_C, D and k_{flip}) for each of the
07 species in the model. The specific growth rate (μ) was calculated from measured OD600 nm at each time point and given
08 as an input to the numerical solver.

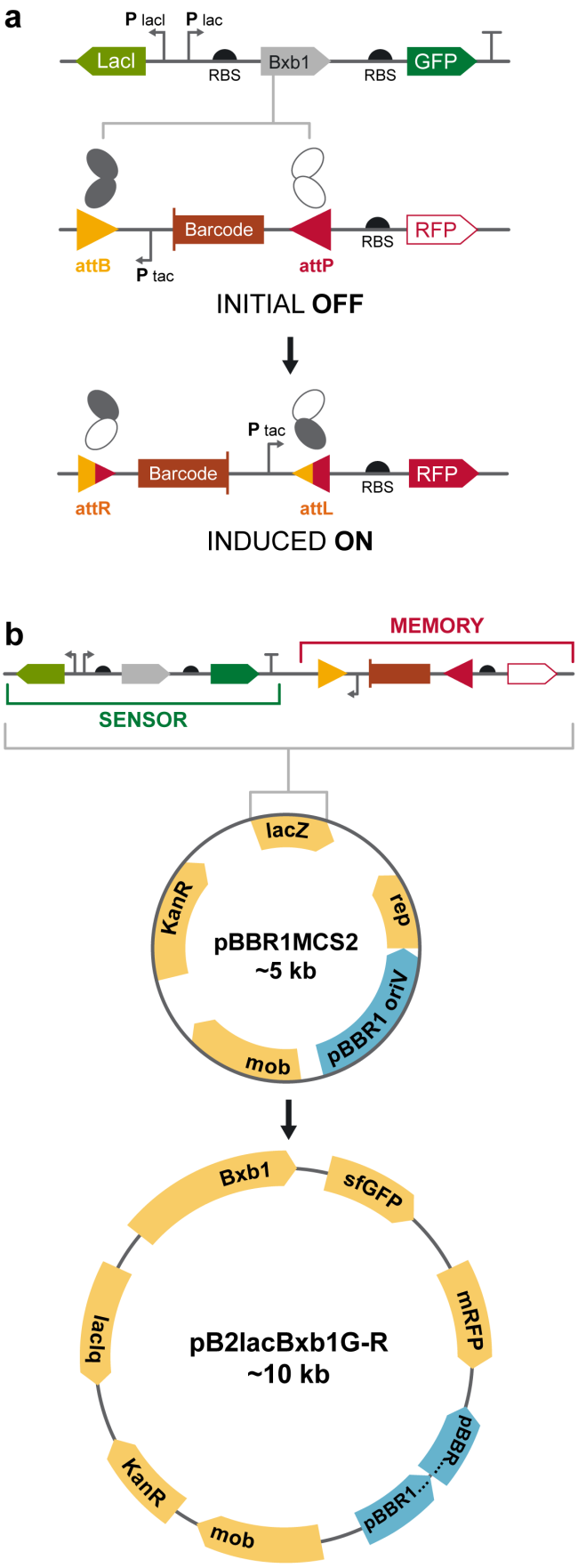
09

10 **RESULTS**

11 **The broad-host range device and its components.** A two-state chemical event logger was built and quantitatively
12 compared across microbial hosts to determine the chassis effect on performance. The device was built with specific
13 sensor and memory components (Fig. 1a). The sensor apparatus consisted of an IPTG inducible P_{lac} promoter (with *lac*
14 operator) driving the expression of the Bxb1 serine-integrase. The Bxb1 gene was transcriptionally fused to a green
15 fluorescent protein (GFP) to monitor the sensor's output. The $lacI^q$ transcription factor, controlling the induction of P_{lac},
16 was driven by the constitutive promoter, P_{lacI^q}. The memory element had two potential states that depended on the Bxb1
17 integrase. The "*off state*" was the initial or unchanged state of the plasmid where the constitutive P_{tac} promoter (without
18 a *lac* operator) was in the reverse orientation from its intended open reading frame, followed by a unique barcode DNA
19 sequence. This construct was enclosed by the *attB* and an *attP* recombination sites recognized by Bxb1. The induced

20 state or "*on state*" was controlled by the formation of a mature Bxb1 dimer, DNA binding and tetramer formation (19),
21 and the respective recombination of *attB* and *attP*. This process re-oriented the constitutive P_{tac} to drive expression of a
22 red fluorescent protein (RFP). A permanent digital memory output was stored by the orientation of the barcode. The
23 performance of the sensor and logger elements were measured by the respective GFP and RFP signals. The entire device
24 was built into a single contig and cloned into the broad-host-range vector pBBR1MCS2 (20) (Fig. 1b).

Figure 1. The broad-host range event logger and its modes of operation. (a) Expression of the Bxb1 integrase was controlled by the Plac promoter and respective IPTG concentrations. Mature Bxb1 proteins dimerize and bind to DNA recombination sites (attB and attP). The dimers on the two ends of the recording element join to form a tetramer in a synaptic event that folds the DNA in the process. DNA strands are exchanged when Bxb1 monomers trade positions, flipping the internal region, which contains both a barcoded digital recorder and a constitutive Ptac promotor. The new sites, attL and attR – formed in the process of DNA flipping – can no longer stay attached to the Bxb1 dimers because of altered sequence, therefore release the dimers in an irreversible digital recording process. GFP and RFP reporter genes were transcriptionally fused onto the IPTG inducible sensor and recording recombination sites. (b) Map of the device with the sensor and memory components cloned in place of the lacZ gene on the pBBR1MCS2 broad-host-range vector.



Quantifying the chassis effect. The device was operational across each of the four species. Performance of each component – sensor and logger – was assayed at eight different IPTG concentrations (0 – 1 mM) by measuring the respective mean GFP and RFP signals (Fig. 2). Total growth was also measured simultaneously via optical density (OD600 nm). The specific growth rates (μ) showed that each species – except *EcN* ($\mu = 1.087 \pm 0.017\ h^{-1}$) – had very similar growth rates under these conditions: *Ec* ($0.427 \pm 0.016\ h^{-1}$), *Pf* ($0.508 \pm 0.019\ h^{-1}$) and *Pp* ($0.439 \pm 0.0436\ h^{-1}$) (Supplemental Fig. S1a). The standard deviations represent the variation in growth rates across all IPTG treatments and ranged from approximately 1.5% to 10% of mean values. The induction strength of the device – controlled by IPTG concentration – showed minimal effect on the specific growth rate. While this indicates that there was little additional metabolic load with respect to IPTG induction, the plasmid encoded device itself imposed a significant metabolic burden on both *Pseudomonas* hosts. This was apparent from the wild-type growth rates of these species, measured at $0.759 \pm 0.017\ h^{-1}$ and $1.127 \pm 0.012\ h^{-1}$ for *Pf* and *Pp* respectively (Supplemental Fig. S1b); much higher than the respective engineered strains. In contrast, there was very little change in specific growth rate of the engineered *E. coli* hosts compared to their respective wild-types ($0.412 \pm 0.012\ h^{-1}$ and $0.979 \pm 0.022\ h^{-1}$ for *Ec* and *EcN* respectively).

The sensor apparatus did not show the tight transcriptional control that was expected from the P_{lac}/*lacI* system in any of the hosts (Fig. 2a). All species had significant level of basal GFP fluorescence (at 0 mM IPTG), indicating leaky expression. Among all the hosts, the sensor component performed the best in *EcN* which had high fluorescence and low basal expression, indicating higher response and tighter regulation than others. *Pf* and *Pp* had similar but high basal GFP expression. While natural fluorescence of the Pseudomonads in the GFP emission spectrum was initially thought to influence the reporter signals, detailed examination revealed that wildtype fluorescence was insignificant compared to the devices´ GFP signal. *Ec* had the lowest GFP fluorescence of all the species but also the lowest basal expression. In general, however, the *E. coli* species showed tighter transcriptional control of the sensor apparatus (ratio of maximum to basal GFP fluorescence around 3) compared to the Pseudomonads (ratio of maximum to basal GFP fluorescence around 1.5).

While the sensor portion of the device behaved similarly in each chassis, the memory apparatus – measured via RFP fluorescence – performed quite differently and showed significant chassis effect. Both *Pseudomonas* hosts performed significantly better than the *E. coli* counterparts (Fig. 2a). They had good transcriptional control and strong

53 RFP fluorescence – *Pp* and *Pf* respectively 3- and 6-times more than the next best *E. coli* host, *EcN*. Interestingly, *Ec*
54 was the lowest performing host from this study, both in terms of dynamic range and maximum fluorescence. In fact, it
55 had a maximum RFU about 24-times lower than the best performing host, *Pf*.

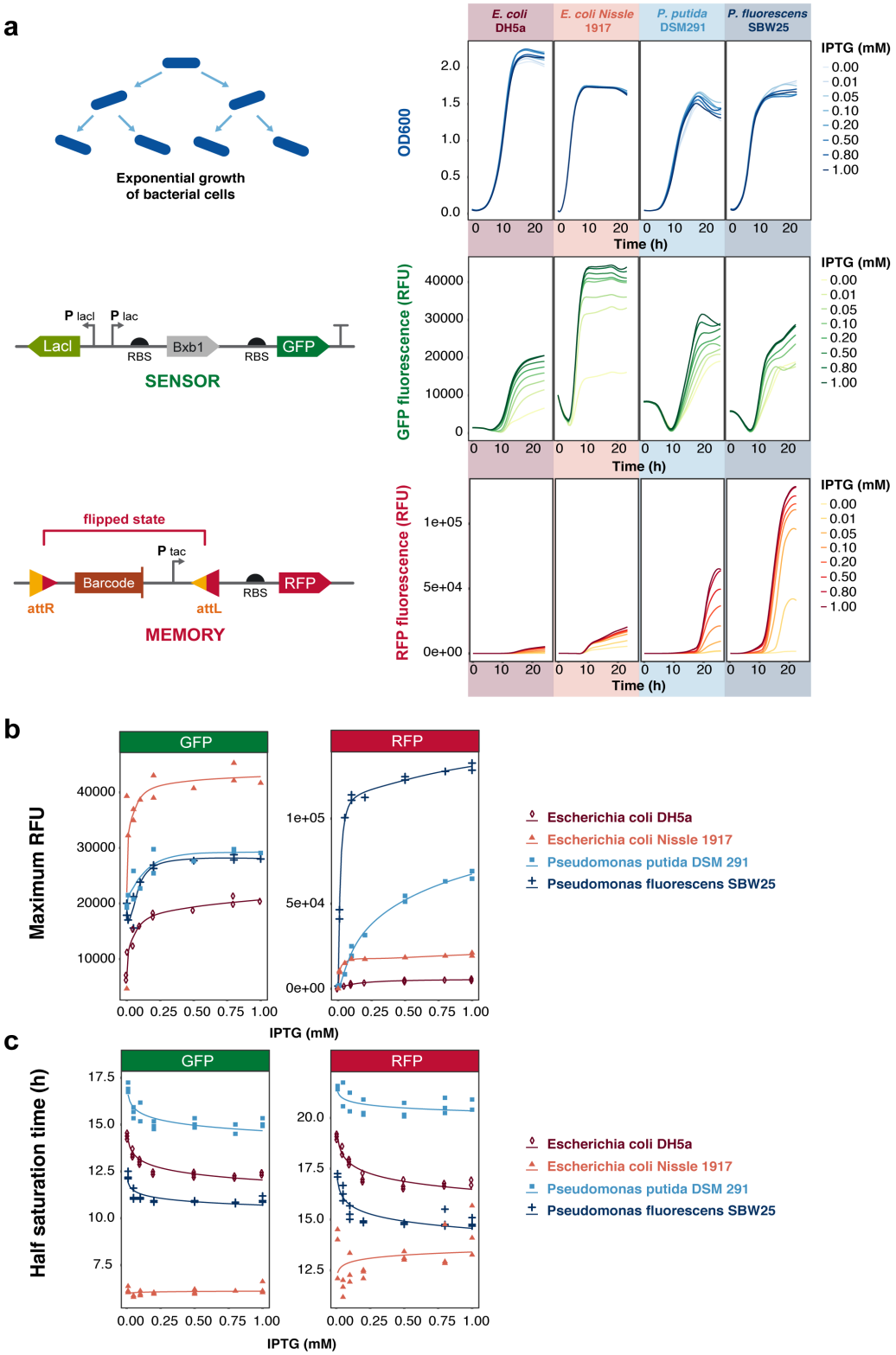


Figure 2. Quantifying the chassis effect. (a) The top four panels represent cell growth (OD600 nm), at different IPTG concentrations, of E. coli DH5α, E. coli Nissle 1917, P. putida DSM 291 and P. fluorescens SBW25 from left to right respectively. The middle row of graphs represents the fluorescence output of the GFP reporter that was transcriptionally fused to the sensor element; the bottom row represents the fluorescence output from the RFP that was fused to the memory component. (b) The maximum GFP and RFP fluorescence at different IPTG concentrations. (c) The activation coefficient shown as half-saturation times of GFP and RFP at different IPTG concentration (> 0 mM IPTG).

57 Another important performance metric for cross-chassis comparison was the time scale of induction (Fig. 2b).
58 This was quantified by the activation coefficient, which is the time for fluorescence to reach half maximum at a given
59 IPTG concentration (21). These had very different profiles than maximum relative fluorescence measurements. *EcN*
60 showed the fastest induction time, but was largely flat with respect to IPTG concentration. This effect was quantified by
61 comparing half saturation times at 0.01 mM IPTG (6.17 ± 0.17 h) and 1 mM IPTG (6.26 ± 0.32 h). For others, the
62 activation coefficient decreased with increasing IPTG concentration and approached a minimum value at the highest
63 IPTG concentrations. Minimum GFP half saturation times were 11 ± 0.17 h, 12.34 ± 0.096 h and 15.09 ± 0.22 h for *Pf,*
64 *Ec* and *Pp* respectively at 1 mM IPTG. RFP fluorescence followed a similar pattern, corresponding directly to induction
65 of P_{lac} by IPTG. This result is evidence that the expression of Bxb1 – thus P_{lac} strength – was the rate limiting step in the
66 process from induction by IPTG to DNA flipping.
67
68 **Population-based comparisons.** Population-level measurements of the device was also carried out by flow-cytometry
69 at five time points during the growth of the microbes (Fig. 3). All populations showed a portion with slight RFP
70 fluorescence (slight tailing of the peaks) from the initial point of induction until about 4-6 hr time period. However, at
71 low IPTG concentrations (0 and 0.01 mM), the RFP populations decreased in most of the species (observed by a
72 sharpening of the peaks) until about 8-16 h. This observed effect evinced an initial and small population harboring a
73 flipped memory element, that were quickly overtaken by the unflipped population because of low level of induction. At
74 later time points and at higher IPTG concentrations, the populations shifted towards more flipped state.
75 These measurements also show that transcriptional control and/or stability of the device was better in the *E. coli*
76 hosts. The 0 mM IPTG treatments showed that the population distributions remained essentially constant for *Ec* and *EcN*,
77 but both *Pp* and *Pf* shifted significantly at the final time point to favor an increasing number of cells with flipped memory
78 element. This represents a false trigger of the event logger after extended periods (likely during stationary phase of
79 growth) and is consistent with the interpretation of leaky expression of the P_{lac} promoter (Fig. 2a).
80

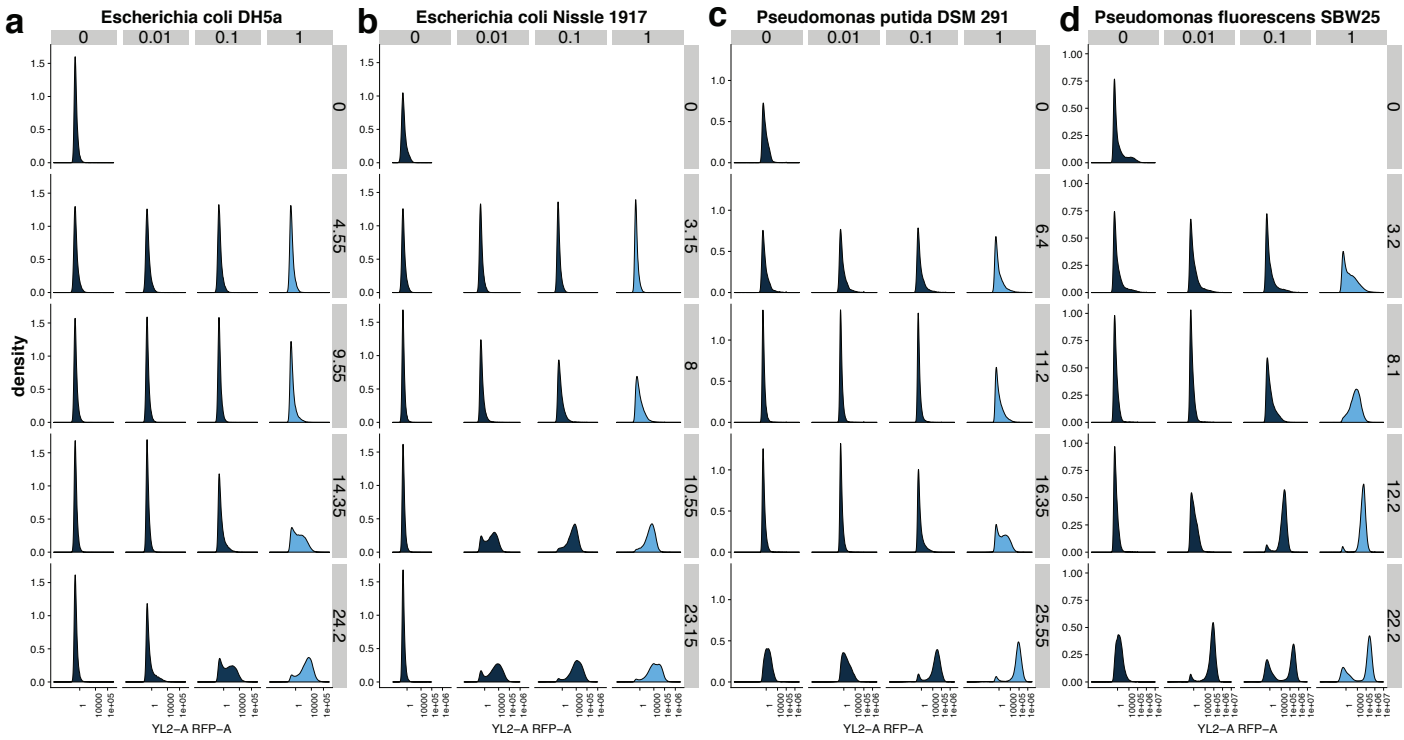


Figure 3. Population-based measurements of RFP reporting of signal recording across hosts. (a) Escherichia coli DH5α, (b) Escherichia coli Nissle 1917, (c) Pseudomonas putida DSM 291 and (d) Pseudomonas fluorescence SBW25. The rows in each of the panels correspond with sampling times in hours and columns are the IPTG concentrations given in mM. The first row in each plot shows data from the induction state from each experiment (0 h).

Simulations and performance metrics across chassis. A kinetic model was formulated to help quantify how individual components of the event logger performed across each chassis. Like the physical construction of the device the model was broken down into the respective sensor and memory component categories. The sensor part of model included expressions that accounted for IPTG induction of the Bxb1 integrase (given as *I*) and GFP as shown by Equations 1 and 2.

$$\frac{dI}{dt} = P - \mu * I - D * I \qquad \text{(Eq. 1)}$$

$$\frac{dGFP}{dt} = P - \mu * GFP - D * GFP \qquad \text{(Eq. 2)}$$

$$P = P_A + P_B ln([IPTG] + P_C) \qquad \text{(Eq. 3)}$$

92 The relationship between IPTG concentration and promoter activity is given by Equation 3, where P represents the
93 activity of P_{lac}. It was found that this type of behavior adequately described the fluorescence output from the device. D
94 is the protein degradation constant and μ is the specific growth rate, which accounts for dilution effects incurred by cell
95 growth.

96 The memory component of the device was modeled by Equations 4 and 5, where PB is the fraction of un-flipped
97 DNA and LR is the fraction of flipped DNA; k_{flip} is the rate constant for integrase-mediated recombination (flipping). We
98 assumed that the plasmid copy number of each host was equivalent and that there existed an un-flipped induction state
99 at the beginning of the experiment.

00

01 $$\frac{dPB}{dt} = -k_{flip} * PB * I^4 \quad \text{(Eq. 4)}$$

02 $$LR = 1 - PB \quad \text{(Eq. 5)}$$

03

04 The rate constant, k_{flip} encompasses three-time steps: 1) the time required for the integrase (I) to form a tetramer; 2) the
05 time required for binding of the tetramer to the DNA (PB) and 3) the DNA flipping event. The overall readout from the
06 memory component of the device was given by expression of RFP (Equation 6), which is analogous to Eq. 2 describing
07 GFP, with the exception of the non-inducible *tac* promoter (P_{RFP}).

08

09 $$\frac{dRFP}{dt} = P_{RFP} - \mu * RFP - D * RFP \quad \text{(Eq. 6)}$$

10

11 This model adequately explained the operation of the genetic device and fitted distinct parameters for each respective
12 host. This was especially evident by the degree to which the model could be fit to the GFP and RFP time series data;
13 each respective output from the sensor and memory components of the device (Fig. 4ab). However, the model's ability
14 to capture the dynamics of DNA flipping was variable between each of the hosts (Fig. 4c). We observed significant
15 scatter derived from the qPCR assays that were designed to measure the orientation of the barcoded DNA associated
16 with the digital memory read-out. These data were collected to determine the fraction of flipped DNA. The noise in the
17 measurement likely resulted from plasmid degradation, variability of plasmid recovery and purification from each host.

Yet, the model still conveyed the overall the pattern for which IPTG induction instigated barcode flipping for each of the hosts and did a reasonably good job at fitting most of the data derived from each time series measurement.

Overall, the model helped show that the *Pseudomonas* hosts had favorable kinetics for operating this device despite the fact the genetic parts have been largely developed and optimized in *E. coli*. They had similar GFP promoter strength (P) as *Ec* (Fig. 4d) but much higher RFP promoter strength (P_{RFP}) and lower degradation constant (D) than both the *E. coli* species (Fig. 4e). However, the rate of DNA flipping, k_{flip}, was quite low in *Pp* which indicated a slightly reduced performance in this host. To our surprise, the simulations suggest that *Ec* – the benchmark chassis – had the most unfavorable kinetics (low P, P_{RFP} and high D). The probiotic strain, *EcN* showed the strongest ability to operate the sensor component of the device and was unique with respect to the suite of hosts tested in this study (Fig. 4d). This could be ascertained by combining the modeled predictions with experimental measurements. For instance, the simulated promoter strength of P_{lac} (see Eq. 3) showed that *Ec, Pp* and *Pf* are aligned with similar profiles, while the values of P for *EcN* were estimated to be about 4-times higher for any given concentration of IPTG. This result was consistent with independent and direct measurements of specific growth rates (Supplemental Fig. S1); *EcN* showed the fastest specific growth rate and should therefore have the highest dilution of expressed GFP protein leading to decreased fluorescence. Yet, *EcN* also showed the highest GFP fluorescence. Thus, the strength of P_{lac} would need to be much higher to account for these opposing effects – as found by the model.

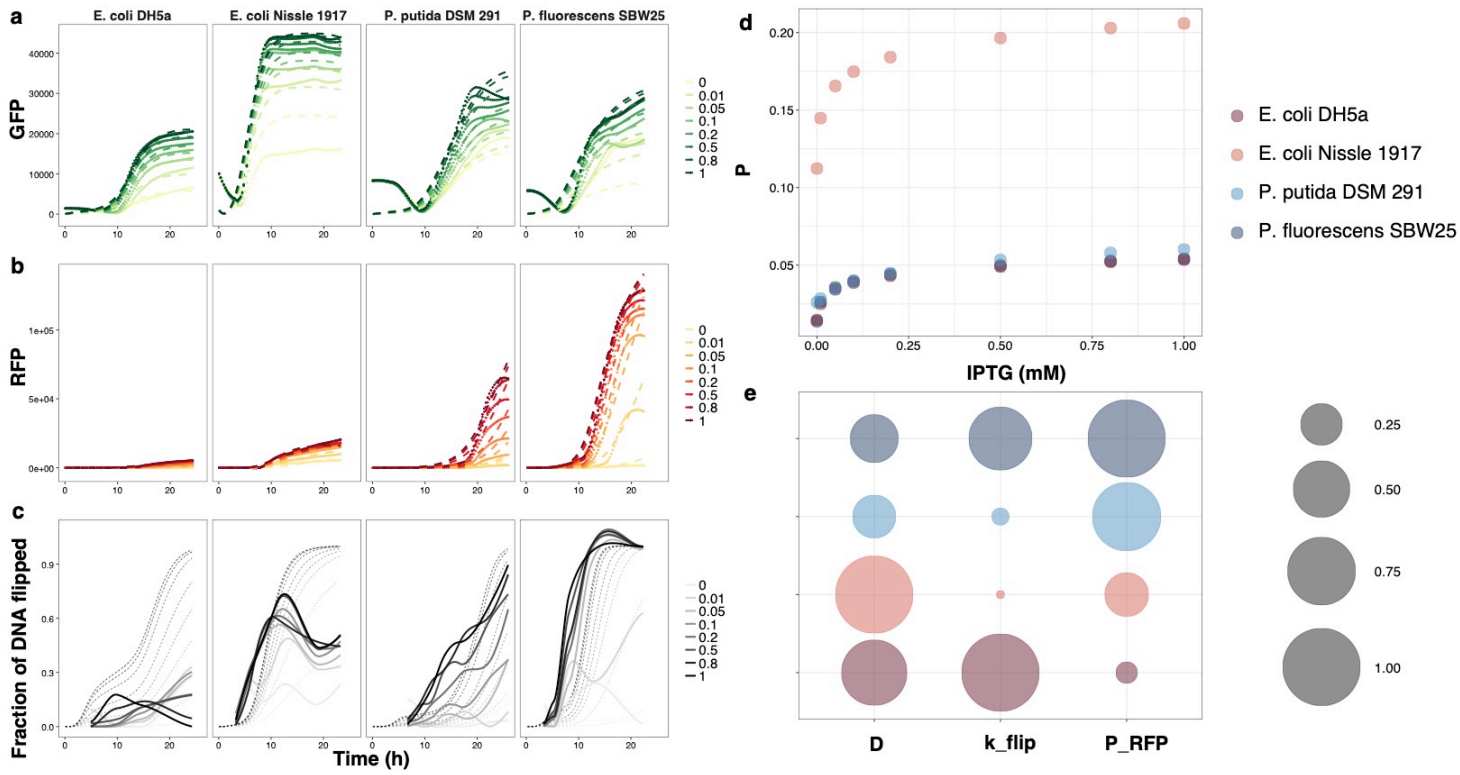


Figure 4. Comparative kinetics show chassis-dependent performance. The model outputs (dotted lines) of (a) the sensor component as fit by GFP fluorescence data; (b) the memory component as fit by RFP fluorescence; and (c) the memory component as fit by the fraction of DNA flipped (qPCR measurements). Each simulated time trace (panels a through c) are overlaid on the respective data typed used to parameterize the model (solid lines). (d) The promoter strength of Plac (given in the model as P) calculated from estimated model parameters, PA, PB and PC, plotted against IPTG concentration. (e) Comparison of simulated kinetic parameters for the protein degradation constant (D); flipping constant (kflip) and strength of the constitutive Ptac promoter (PRFP) compared across each host.

DISCUSSION

Synthetic biologists commonly (re-)discover that even the most well-characterized genetic parts often will not function in a predictable manner when taken out of the context from which they were originally characterized. For any given host, this unpredictability can arise from interference between genetic parts that have been introduced as well as cellular noise inherent to the native biological system (22,23). Yet, the degree to which these factors are influenced by the biology of any given microbial species requires that the same genetic parts be used and compared across multiple hosts. Here, we showed that an identical genetic device can be constructed in a broad-host-range vector and ported across multiple microbial species and that its performance is host-dependent.

We chose to deploy a relatively simple event logger and experimental design, which has enabled this study to demonstrate emerging capability of broad-host-range genetic devices. In fact, the Bxb1 serine integrase was originally

45 harnessed by synthetic biologists (24) in part because it does not require host cofactors, which is a feature enabling reuse
46 across multiple hosts (25,26). The results from this current study confirm the suitability of Bxb1 as a broad-host-amenable
47 genetic part by comparatively quantifying its performance with our broad-host event detector expressed from four
48 species. While it is clear that more species and devices need to be tested before more extensive broad-host-range parts
49 libraries can mature, this early step is an important contribution towards alleviating our current dependency and
50 limitations on small subset of model microbial hosts.

51 Despite the fact that *E. coli* DH5α is a common tool for the design and implementation of modern genetic devices,
52 we found that in many ways it was the least ideal host for the implementation of a device for actual application as tested
53 in this study. In fact, a primary finding was that – compared to *Ec* – the two *Pseudomonas* species (*Pp* and *Pf*) showed
54 reasonable potential as chassis for chemical event logging even though the majority of previously published reports on
55 the parts used to build the device have only considered *E. coli* (3,4,6,27). We regard this as a promising result because
56 these and closely related *Pseudomonads* are known to have tremendous metabolic potential for the synthesis of novel
57 compounds (28,29), consuming complex substrates (30,31) and persisting in a wide range of habitats that include soils,
58 plant tissues and marine ecosystems (32-35). Expanding the synthetic parts list for *Pseudomonas* species will
59 undoubtedly enable new biotechnological applications that should include chemical sensing and event logging in
60 complex natural environments.

61 *E. coli* Nissle 1917 was also chosen for comparative analysis in this study because it served as an intra-species
62 comparator to *Ec*. It was also chosen because of its growing importance in the bio-design community based on the fact
63 that it is a commonly used probiotic(36) and highly genetically tractable. Researchers are rapidly uncovering many
64 exciting opportunities to use *EcN* and other probiotic-hosts as programable therapeutic agents and/or diagnostic tools for
65 human health(37-39). In some cases, differences in intra-species performance – within *E. coli* strains – exceeded inter-
66 species variability. This was somewhat unexpected and specifically evident from comparisons made on the sensor
67 component of the device, which performed better in *EcN* as compared to *Ec* and both *Pseudomonas* hosts. This was
68 specifically evident by comparing the kinetics associated with the sensor apparatus and indicates that *EcN* maintained
69 the tightest control and largest dynamic ranges of the IPTG inducible components of the device.

70 Kinetic parameters estimated from the model provided a good quantitative comparison of biological properties
71 that cannot be easily measured (Fig. 4e). For instance, the degradation constant, *D*, is found to be fairly similar in all the

72 hosts, which is hardly surprising considering the standardized growth conditions and similar growth rates. Estimates of
73 the flipping rate constant (k_{flip}), however, were highly variable. In contrast to D, which is more indicative of cellular
74 physiology, k_{flip} is more representative of the device-specific kinetics. This parameter depends on a number of biological
75 factors such as codon usage, transcription, translation, protein folding as well as the efficiency of Bxb1-mediated
76 recombination. Based on the model-enabled predictions, *EcN* stood out with a very small k_{flip} values; about 89 times
77 smaller than the largest value attributed to *Pf*. Its high transcription rate (given by estimates of P) and low k_{flip} account
78 for the observation of fast DNA flipping after initial induction followed by relatively immediate saturation (Fig. 4c). The
79 k_{flip} values of Ec and *Pp* are moderate but the reason for the higher value of *Ec* relative to *Pp* is still somewhat uncertain
80 since the fraction of DNA flipped is higher in *Pp* than *Ec*. The fourth parameter, P_{RFP}, is the measure of the strength of
81 P_{tac} promoter and varies in the same way as RFP fluorescence. This promoter was actually found to work better – as
82 assayed by the strength or RFP fluorescence – in the *Pseudomonads* than *E. coli* species.

83 Integrated data and kinetic modelling approaches are useful for quantifying and comparing performance across
84 hosts. One limitation, however, was that our approach contained few species-specific physiological parameters. The
85 exception to this is the specific growth rate (μ). Although the hosts in this study all showed similar growth rates, the
86 specific growth rate should prove to be an important consideration when evaluating the performance of a device as hosts
87 and growth conditions change. It was also interesting that we were able to observe and simulate dynamics in the device's
88 performance while the cells were in stationary phase. Often, experimental observations made on engineered devices are
89 only contextualized during log growth phases. However, future applications such as chemical event logging in dynamic
90 environments will be better served by understanding how chassis/device pairs may function through lag, log and
91 stationary phases of growth. This is a point that shall require more deserving attention in future studies.

92 The field of microbial bio-design is keen to harness new, non-traditional hosts for synthetic biology applications.
93 Some significant advancements towards programing genetic devices – including sensors – have already been shown in
94 other non-traditional microbial hosts. Of specific note are previous success shown in a human gut microbe *Bacteroides*
95 *thetaiotaomicron* (40) and a suite of proteobacteria isolated from a bee gut microbiome (41). Here in this current study,
96 we have advanced an emerging concept of broad-host-range genetic devices. While this is certainly a new frontier, some
97 notable examples have preceded this current report including a study by Kushwana and Salis that that presented the
98 concept of "portable power supplies" between species and demonstrated that some genetic parts can ported between *E.*

99 *coli*, *P. putida* and *Bacillus subtilis* (14). Another important avenue has been the pursuit of broad-spectrum genetic parts
00 such as the promoters presented in a study from Yang et al. that are operational between *E. coli*, *B. subtilis* and *S.*
01 *cerevisiae* (15). The efforts to date – including our current study – have only considered a relatively small set of microbes.
02 Future developments on cross-chassis devices may encounter new technical hurdles as the taxonomic diversity of hosts
03 are expanded. Once harnessed, the concept of broad-host-range genetic devices should also bring new species-specific-
04 applications. The major technical hurdle that will need to be overcome for developing chemical sensing capabilities will
05 be the discovery or engineering of genetic components with specificity for analytes of real-world interest. The current
06 suite of commonly used transcriptional factors and inducible promoters are clearly limited. New parts discovery and
07 characterization efforts are sorely needed to advance the current state of microbial biodesign.
08
09 **Conclusions.** We quantified the chassis effect of an integrase-based chemical event logger across multiple species and
10 two different Genera – *Pseudomonas* and *Escherichia*. The performance of sensor and memory components changed
11 according to each host as ascertained via integrated experimental measurements and outputs from kinetic models.
12 Specifically, *EcN* – a common probiotic bacterium – showed the tightest control and most stability of the sensor apparatus
13 that regulated expression of the Bxb1 integrase. Both *Pseudomonas* hosts showed greater RFP output signals that
14 corresponded with Bxb1-mediated recombination in the memory component of the device. A primary finding of this
15 study was that – compared to *Ec* – the two *Pseudomonas* species (*Pp* and *Pf*) showed reasonable potential as chemical
16 event logging chassis. This study advances an emerging frontier in synthetic biology that aims to build broad-host-range
17 devices and understand the context by which different microbial species can execute programmable genetic operations
18
19 **SUPPLEMENTARY DATA**
20 The raw data for this study along with and the sequence information for the genetic construct pB2lacBxb1G-R (shown
21 in Fig. 1b) are available from the Open Science Framework (OSF) under the name "A broad-host-range event detector:
22 data and models" (DOI 10.17605/OSF.IO/J295C) at https://osf.io/j295c/. This OSF project also contains the R Markdown
23 scripts that can be used to reproduce all of the analyses, statistics and technical graphs presented in this manuscript.
24
25

CONFLICT OF INTEREST

The authors have no conflict of interest to declare.

ACKNOWLEDGMENTS

This research was supported by funding from the National Security Directorate Seed Initiative, a Laboratory Directed Research and Development (LDRD) Program of Pacific Northwest National Laboratory (PNNL). PNNL is operated for the DOE by Battelle under contract no. DE-AC05-76RLO-1830. Specific acknowledgments are given to Rose Perry at PNNL and John Repass at Genetics for technical assistance with graphics and real time PCR. The authors would also like to thank Drs. Victoria Hsiao and Richard Murray for kindly sharing materials and knowledge.

REFERENCES CITED

1. Gardner, T.S., Cantor, C.R. and Collins, J.J. (2000) Construction of a genetic toggle switch in Escherichia coli. *Nature*, **403**, 339.
2. Elowitz, M.B. and Leibler, S. (2000) A synthetic oscillatory network of transcriptional regulators. *Nature*, **403**, 335.
3. Hsiao, V., Hori, Y., Rothemund, P.W. and Murray, R.M. (2016) A population-based temporal logic gate for timing and recording chemical events. *Mol Syst Biol*, **12**, 869.
4. Yang, L., Nielsen, A.A., Fernandez-Rodriguez, J., McClune, C.J., Laub, M.T., Lu, T.K. and Voigt, C.A. (2014) Permanent genetic memory with> 1-byte capacity. *Nat Meth*, **11**, 1261.
5. Siuti, P., Yazbek, J. and Lu, T.K. (2013) Synthetic circuits integrating logic and memory in living cells. *Nat Biotech*, **31**, 448.
6. Shur, A. and Murray, R.M. (2018) Proof of concept continuous event logging in living cells. *bioRxiv*, 225151.
7. Riglar, D.T., Giessen, T.W., Baym, M., Kerns, S.J., Niederhuber, M.J., Bronson, R.T., Kotula, J.W., Gerber, G.K., Way, J.C. and Silver, P.A. (2017) Engineered bacteria can function in the mammalian gut long-term as live diagnostics of inflammation. *Nat Biotech*, **35**, 653.
8. Kotula, J.W., Kerns, S.J., Shaket, L.A., Siraj, L., Collins, J.J., Way, J.C. and Silver, P.A. (2014) Programmable bacteria detect and record an environmental signal in the mammalian gut. *Proc Nat Acad Sci*, **111**, 4838-4843.
9. Wang, B., Barahona, M. and Buck, M. (2013) A modular cell-based biosensor using engineered genetic logic circuits to detect and integrate multiple environmental signals. *Biosens Bioelectron*, **40**, 368-376.
10. Stocker, J., Balluch, D., Gsell, M., Harms, H., Feliciano, J., Daunert, S., Malik, K.A. and Van der Meer, J.R. (2003) Development of a set of simple bacterial biosensors for quantitative and rapid measurements of arsenite and arsenate in potable water. *Environ Sci Technol*, **37**, 4743-4750.
11. Tan, S.Z., Reisch, C.R. and Prather, K.L. (2018) A Robust CRISPRi Gene Repression System in Pseudomonas. *J Bacteriol*, JB. 00575-00517.
12. Nikel, P.I., Martínez-García, E. and De Lorenzo, V. (2014) Biotechnological domestication of pseudomonads using synthetic biology. *Nat Rev Microbiol*, **12**, 368.
13. Nikel, P.I., Chavarría, M., Danchin, A. and de Lorenzo, V. (2016) From dirt to industrial applications: Pseudomonas putida as a synthetic biology chassis for hosting harsh biochemical reactions. *Curr Opin Chem Biol*, **34**, 20-29.
14. Kushwaha, M. and Salis, H.M. (2015) A portable expression resource for engineering cross-species genetic circuits and pathways. *Nat Commun*, **6**, 7832.
15. Yang, S., Liu, Q., Zhang, Y., Du, G., Chen, J. and Kang, Z. (2017) Construction and characterization of broad-spectrum promoters for synthetic biology. *ACS Synth Biol*, **7**, 287-291.
16. Tu, Q., Yin, J., Fu, J., Herrmann, J., Li, Y., Yin, Y., Stewart, A.F., Müller, R. and Zhang, Y. (2016) Room temperature electrocompetent bacterial cells improve DNA transformation and recombineering efficiency. *Sci Rep*, **6**, 24648.
17. Soetaert, K., Petzoldt, T. and Setzer, R.W. (2010) Solving differential equations in R: package deSolve. *J Stat Software*, **33**.
18. Team, R.C. (2013) R: A language and environment for statistical computing.
19. Ghosh, P., Pannunzio, N.R. and Hatfull, G.F. (2005) Synapsis in phage Bxb1 integration: selection mechanism for the correct pair of recombination sites. *J Mol Biol*, **349**, 331-348.
20. Kovach, M.E., Elzer, P.H., Hill, D.S., Robertson, G.T., Farris, M.A., Roop, R.M. and Peterson, K.M. (1995) Four new derivatives of the broad-host-range cloning vector pBBR1MCS, carrying different antibiotic-resistance cassettes. *Gene*, **166**, 175-176.
21. Alon, U. (2006) *An introduction to systems biology: design principles of biological circuits*. CRC press.
22. Hooshangi, S., Thiberge, S. and Weiss, R. (2005) Ultrasensitivity and noise propagation in a synthetic transcriptional cascade. *Proc Nat Acad Sci*, **102**, 3581-3586.
23. Lou, C., Stanton, B., Chen, Y.-J., Munsky, B. and Voigt, C.A. (2012) Ribozyme-based insulator parts buffer synthetic circuits from genetic context. *Nat Biotech*, **30**, 1137.
24. Bonnet, J., Subsoontorn, P. and Endy, D. (2012) Rewritable digital data storage in live cells via engineered control of recombination directionality. *Proc Nat Acad Sci*, **109**, 8884-8889.

25. Keravala, A., Groth, A.C., Jarrahian, S., Thyagarajan, B., Hoyt, J.J., Kirby, P.J. and Calos, M.P. (2006) A diversity of serine phage integrases mediate site-specific recombination in mammalian cells. *Mol Gen Genom*, **276**, 135.
26. Courbet, A., Endy, D., Renard, E., Molina, F. and Bonnet, J. (2015) Detection of pathological biomarkers in human clinical samples via amplifying genetic switches and logic gates. *Sci Trans Med*, **7**, 289ra283-289ra283.
27. Siuti, P., Yazbek, J. and Lu, T.K. (2014) Engineering genetic circuits that compute and remember. *Nat Protoc*, **9**, 1292.
28. Poblete-Castro, I., Becker, J., Dohnt, K., Dos Santos, V.M. and Wittmann, C. (2012) Industrial biotechnology of Pseudomonas putida and related species. *Appl Microbiol Biotechnol*, **93**, 2279-2290.
29. Di Gioia, D., Luziatelli, F., Negroni, A., Ficca, A.G., Fava, F. and Ruzzi, M. (2011) Metabolic engineering of Pseudomonas fluorescens for the production of vanillin from ferulic acid. *J Biotechnol*, **156**, 309-316.
30. Clarke, P.H. (1982) The metabolic versatility of pseudomonads. *Anton Van Leeuw*, **48**, 105-130.
31. Nelson, K., Weinel, C., Paulsen, I., Dodson, R., Hilbert, H., Martins dos Santos, V., Fouts, D., Gill, S., Pop, M. and Holmes, M. (2002) Complete genome sequence and comparative analysis of the metabolically versatile Pseudomonas putida KT2440. *Environ Microbiol*, **4**, 799-808.
32. Rainey, P.B. and Bailey, M.J. (1996) Physical and genetic map of the Pseudomonas fluorescens SBW25 chromosome. *Molec Microbiol*, **19**, 521-533.
33. Silby, M.W., Cerdeño-Tárraga, A.M., Vernikos, G.S., Giddens, S.R., Jackson, R.W., Preston, G.M., Zhang, X.-X., Moon, C.D., Gehrig, S.M. and Godfrey, S.A. (2009) Genomic and genetic analyses of diversity and plant interactions of Pseudomonas fluorescens. *Genome Biol*, **10**, R51.
34. Compeau, G., Al-Achi, B.J., Platsouka, E. and Levy, S. (1988) Survival of rifampin-resistant mutants of Pseudomonas fluorescens and Pseudomonas putida in soil systems. *Appl Environ Microb*, **54**, 2432-2438.
35. Isnansetyo, A. and Kamei, Y. (2009) Bioactive substances produced by marine isolates of Pseudomonas. *J Ind Microbiol Biot*, **36**, 1239-1248.
36. Schultz, M. and Burton, J. (2017), *The Microbiota in Gastrointestinal Pathophysiology*. Elsevier, pp. 59-69.
37. Slomovic, S., Pardee, K. and Collins, J.J. (2015) Synthetic biology devices for in vitro and in vivo diagnostics. *Proc Nat Acad Sci*, **112**, 14429-14435.
38. Danino, T., Prindle, A., Kwong, G.A., Skalak, M., Li, H., Allen, K., Hasty, J. and Bhatia, S.N. (2015) Programmable probiotics for detection of cancer in urine. *Sci Transl Med*, **7**, 289ra284-289ra284.
39. Sedlmayer, F. and Fussenegger, M. (2017) Synthetic biology: A probiotic probe for inflammation. *Nat Biomed Eng*, **1**, 0097.
40. Mimee, M., Tucker, A.C., Voigt, C.A. and Lu, T.K. (2015) Programming a human commensal bacterium, Bacteroides thetaiotaomicron, to sense and respond to stimuli in the murine gut microbiota. *Cell Syst*, **1**, 62-71.
41. Leonard, S.P., Perutka, J., Powell, J.E., Geng, P., Richhart, D.D., Byrom, M., Kar, S., Davies, B.W., Ellington, A.D. and Moran, N.A. (2018) Genetic engineering of bee gut microbiome bacteria with a toolkit for modular assembly of broad-host-range plasmids. *ACS Synth Biol*, **7**, 1279-1290.

Supplemental Information for: *A broad-host-range event detector: expanding and quantifying performance across bacterial species*

Table 1. Standard Curves for qPCR - preparation of plasmid DNA serial dilutions for standard curve reactions

Plasmid + Diluent* (µL)	**Dilution Factor**	**Plasmid copies per µL**	**Plasmid copies per reaction**
Dilute working stock**	None	1.0E+08	N/A
3 + 597	1/200	0.5E+06	1.0E+07
5 + 45	1/10	0.5E+05	1.0E+06
5 + 45	1/10	0.5E+04	1.0E+05
5 + 45	1/10	0.5E+03	1.0E+04
5 + 45	1/10	0.5E+02	1.0E+03
12.5 + 37.5	1/4	125	250
25 + 25	1/2	50	100
NTC (diluent only)	N/A	0	0

*Plasmid diluent solution was ultrapure water containing 20ng/µL of human genomic DNA.

** Flipped and un-flipped plasmids were diluted in above diluent to create a working master stock of 1.0E + 08 copies per µL (30 March 2018; stored frozen -80C). Single-use aliquots were prepared on 4April2018 by diluting the working stock to 1.0E+07 in diluent and stored frozen. Single-use aliquots were diluted on testing day to create standard curve samples. Determination of vector DNA concentration was only employed on the linear portion of the standard curve.

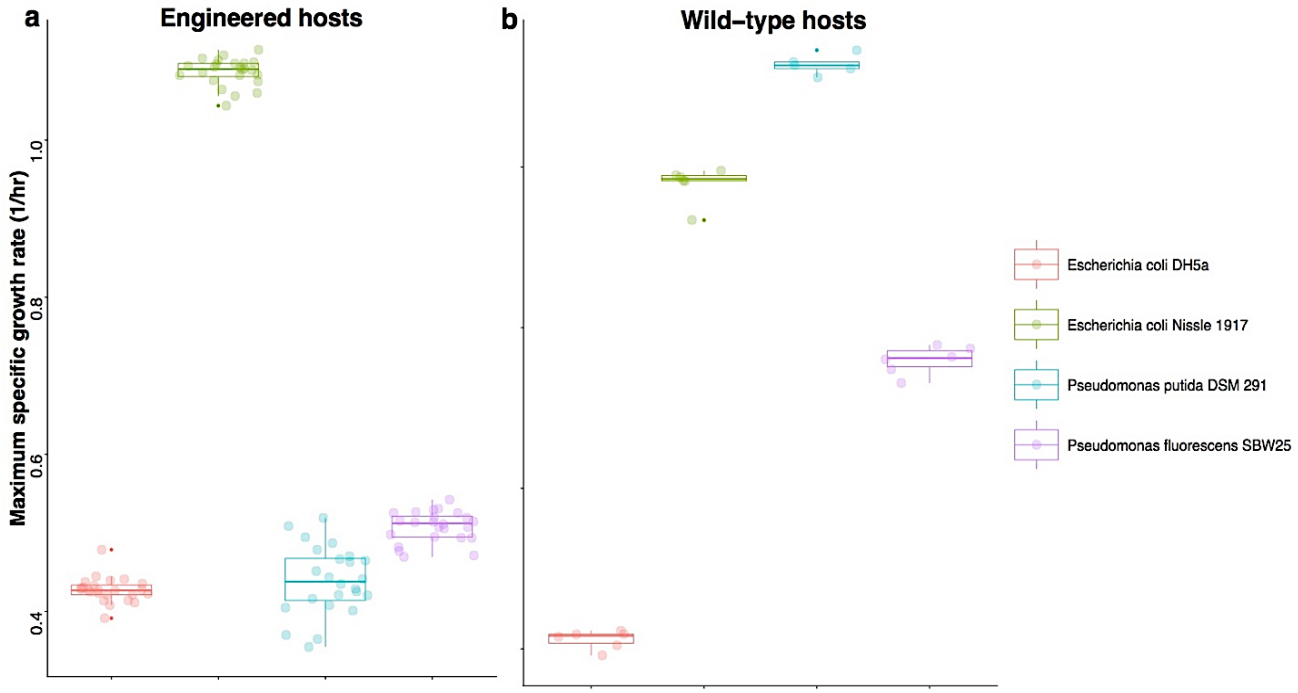
39 **Table 2. Sequences of major components**

Bxb1 integrase	GTGAGAGCCCTGGTAGTCATCCGCCTGTCCCGCGTCACCGATGCTACGACTTC ACCGGAGCGTCAGCTGGAGTCTTGCCAGCAGCTCTGCGCCCAGCGCGGCTGG GACGTCGTCGGGGTAGCGGAGGATCTGGACGTCTCCGGGGCGGTCGATCCGT TCGACCGGAAGCGCAGACCGAACCTGGCCCGGTGGCTAGCGTTCGAGGAGCA ACCGTTTGACGTGATCGTGGCGTACCGGGTAGATCGGTTGACCCGATCGATCC GGCATCTTCAGCAGCTGGTCCACTGGGCCGAGGACCACAAGAAGCTGGTCGT CTCCGCGACCGAAGCGCACTTCGATACGACGACGCCGTTTGCGGCGGTCGTC ATCGCGCTTATGGGAACGGTGGCGCAGATGGAATTAGAAGCGATCAAAGAGC GGAACCGTTCGGCTGCGCATTTCAATATCCGCGCCGGGAAATACCGAGGATC CCTGCCGCCGTGGGGATACCTGCCTACGCGCGTGGACGGGGAGTGGCGGCTG GTGCCGGACCCTGTGCAGCGAGAGCGCATCCTCGAGGTGTATCACCGCGTCG TCGACAACCACGAGCCGCTGCATCTGGTGGCCCACGACCTGAACCGGCGTGG TGTCCTGTCGCCGAAGGACTACTTCGCGCAGCTGCAAGGCCGCGAGCCGCAG GGCCGGGAGTGGTCGGCTACCGCGCTGAAGCGATCGATGATCTCCGGGGCGA TGCTCGGGTACGCGACTCTGAACGGTAAGACCGTCCGAGACGACGACGGAGC CCCGCTGGTGCGGGCTGAGCCGATCCTGACCCGTGAGCAGCTGGAGGCGCTG CGCGCCGAGCTCGTGAAGACCTCCCGGGCGAAGCCCGCGGTGTCTACCCCGT CGCTGCTGCTGCGGGTGTTGTTCTGCGCGGTGTGCGGGGAGCCCGCGTACAAG TTCGCCGGGGGAGGACGTAAGCACCCGCGCTACCGCTGCCGCTCGATGGGGT TCCCGAAGCACTGCGGGAACGGCACGGTGGCGATGGCCGAGTGGGACGCGTT CTGCGAGGAGCAGGTACTGGATCTGCTCGGGGACGCGGAGCGTCTGGAGAAA GTCTGGGTAGCGGGCTCGGACTCCGCGGTCGAACTCGCGGAGGTGAACGCGG AGCTGGTGGACCTGACGTCGCTGATCGGCTCCCCGGCCTACCGGGCGGGCTCT CCGCAGCGAGAAGCACTGGATGCCCGTATTGCGGCGCTGGCCGCGCGGCAAG AGGAGCTGGAGGGCCTGGAGGCTCGCCCGTCTGGCTGGGAGTGGCGCGAGAC CGGGCAGCGGTTCGGGGACTGGTGGCGGGAGCAGGACACCGCGGCAAAGAA CACCTGGCTTCGGTCGATGAACGTTCGGCTGACGTTCGACGTCCGCGGCGGGC TGACTCGCACGATCGACTTCGGGGATCTTCAGGAGTACGAGCAGCATCTCAG GCTCGGCAGCGTGGTCGAACGGCTACACACCGGGATGTCGTAA
sfGFP	ATGCGTAAAGGCGAAGAGCTGTTCACTGGTGTCGTCCCTATTCTGGTGGAACT GGATGGTGATGTCAACGGTCATAAGTTTTCCGTGCGTGGCGAGGGTGAAGGT GACGCAACTAATGGTAAACTGACGCTGAAGTTCATCTGTACTACTGGTAAACT GCCGGTACCTTGGCCGACTCTGGTAACGACGCTGACTTATGGTGTTCAGTGCT TTGCTCGTTATCCGGACCATATGAAGCAGCATGACTTCTTCAAGTCCGCCATG CCGGAAGGCTATGTGCAGGAACGCACGATTTCCTTTAAGGATGACGGCACGT ACAAAACGCGTGCGGAAGTGAAATTTGAAGGCGATACTCTGGTAAACCGCAT TGAGCTGAAAGGCATTGACTTTAAAGAAGACGGCAATATCCTGGGCCATAAG CTGGAATACAATTTTAACAGCCACAATGTTTACATCACCGCCGATAAACAAAA AAATGGCACTAAAGCGAATTTTAAAATTCGCCACAACGTGGAGGATGGCAGC GTGCAGCTGGCTGATCACTACCAGCAAAACACTCCAATCGGTGATGGTCCTGT TCTGCTGCCAGACAATCACTATCTGAGCACGCAAAGCGTTCTGTCTAAAGATC CGAACGAGAAACGCGATCATATGGTTCTGCTGGAGTTCGTAACCGCAGCGGG CATCACGCATGGTATGGATGAACTGTAC
attB	TCGGCCGGCTTGTCGACGACGGCGGTCTCCGTCGTCAGGATCATCCGGGC
attP	TCGTGGTTTGTCTGGTCAACCACCGCGGTCTCAGTGGTGTACGGTACAAACCC
mRFP	ATGGCTTCCTCCGAAGATGTTATCAAAGAGTTCATGCGTTTCAAAGTTCGTAT GGAAGGTTCCGTTAACGGTCACGAGTTCGAAATCGAAGGTGAAGGTGAAGGT CGTCCGTACGAAGGTACCCAGACCGCTAAACTGAAAGTTACCAAAGGTGGTC CGCTGCCGTTCGCTTGGGACATCCTGTCCCCGCAGTTCCAGTACGGTTCCAAA GCTTACGTTAAACACCCGGCTGACATCCCGGACTACCTGAAACTGTCCTTCCC GGAAGGTTTCAAATGGGAACGTGTTATGAACTTCGAGGACGGTGGTGTTGTTA CCGTTACCCAGGACTCCTCCCTGCAAGACGGTGAGTTCATCTACAAAGTTAAA CTGCGTGGTACCAACTTCCCGTCCGACGGTCCGGTTATGCAGAAAAAAACCAT

	GGGTTGGGAAGCTTCCACCGAACGTATGTACCCGGAAGATGGTGCTCTGAAA GGTGAAATCAAAATGCGTCTGAAACTGAAAGACGGTGGTCACTACGACGCTG AAGTTAAAACCACCTACATGGCTAAAAAACCGGTTCAGCTGCCGGGTGCTTA CAAAACCGACATCAAACTGGACATCACCTCCCACAACGAGGACTACACCATC GTTGAACAGTACGAACGTGCTGAAGGTCGTCACTCCACCGGTGCTTAA

Table 3. qPCR primer/probe sequences

Name	Sequence details	
P1	Primer 1 (3', Outside, reverse)	AGTTTCGTGGTTTGTCTGGT
P2	Primer 2 (Inside, unflipped device, 3' forward)	GGCGGATAAAGTTGCAGGA
P3	Primer 3 (Inside, flipped device, 3' forward)	AGCAGATCCGGAACATAATGG
Prb1	Probe 1 (unflipped device, 3' reverse)	CGGTCTCAGTGGTGTACGGTACAAA
Prb2	Probe 2 (flipped device, 3' reverse)	CGGTCTCCGTCGTCAGGATCAT



Supplemental Figure S1. Comparisons of the specific growth rates between hosts. (a) The maximum specific growth rates of each host expressing the pBBR1MCS2-based event logging device at all IPTG concentrations tested. **(b)** Maximum specific growth rates of each wild-type host; absent of expression vector and device.

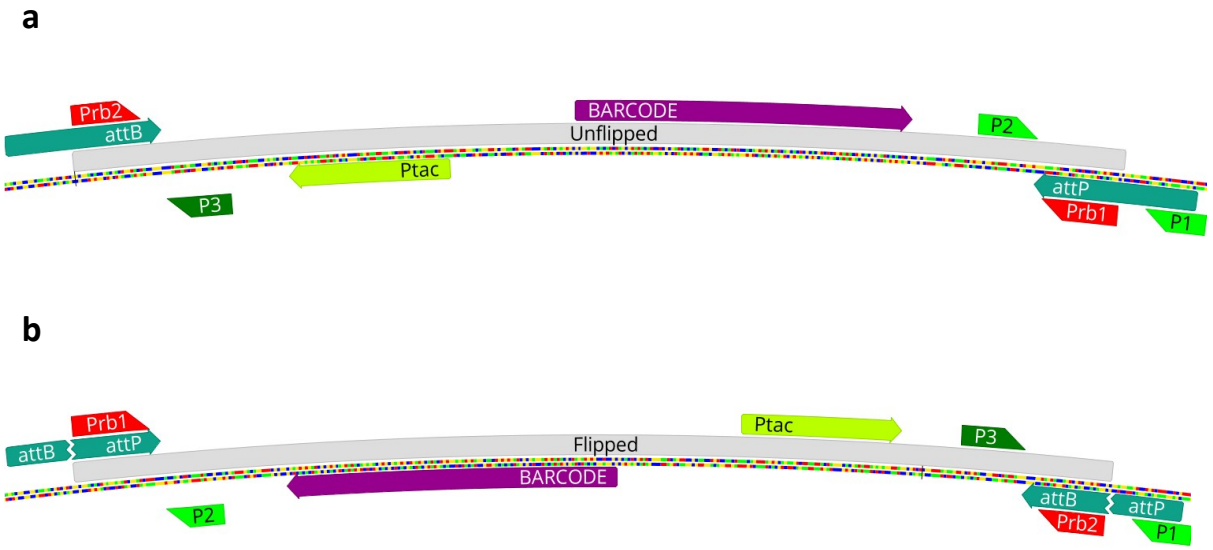


Figure S2. Device maps with primer, probe locations - (a) unflipped, (b) flipped. In the unflipped orientation primers P1 and P2 form a pair and amplifies, while P3 remains unpaired and causes very little amplification. When the device is flipped, positions of P2 and P3 are reversed, and P3 now forms a pair with P1 and amplifies.