MOLECULAR BIOLOGY

Barcoded microbial system for high-resolution object provenance

Jason Qian[1,2,3]*, Zhi-xiang Lu[1,2]*, Christopher P. Mancuso[4]*, Han-Ying Jhuang[1]*, Rocío del Carmen Barajas-Ornelas[5]*, Sarah A. Boswell[1,2]*, Fernando H. Ramírez-Guadiana[5], Victoria Jones[1,6]†, Akhila Sonti[4], Kole Sedlack[4]‡, Lior Artzi[5], Giyoung Jung[7], Mohammad Arammash[1], Mary E. Pettit[1], Michael Melfi[1], Lorena Lyon[1], Siân V. Owen[6], Michael Baym[2,6], Ahmad S. Khalil[4,8], Pamela A. Silver[1,8], David Z. Rudner[5], Michael Springer[1,2]§

Determining where an object has been is a fundamental challenge for human health, commerce, and food safety. Location-specific microbes in principle offer a cheap and sensitive way to determine object provenance. We created a synthetic, scalable microbial spore system that identifies object provenance in under 1 hour at meter-scale resolution and near single-spore sensitivity and can be safely introduced into and recovered from the environment. This system solves the key challenges in object provenance: persistence in the environment, scalability, rapid and facile decoding, and biocontainment. Our system is compatible with SHERLOCK, a Cas13a RNA-guided nucleic acid detection assay, facilitating its implementation in a wide range of applications.

Globalization of supply chains has substantially complicated the process of determining the origins of agricultural products and manufactured goods. Determining the origins of these objects can be critical, for example, in cases of foodborne illness, but current labeling technologies are prohibitively labor intensive and easy to subvert (*1*). Tools that label persons or objects passing through a location of interest could also be useful to law enforcement as a complement to fingerprinting and video surveillance (*2*). Microbial communities offer a potential alternative to standard labeling approaches. Any object gradually adopts the naturally occurring microbes present in its environment (*3*, *4*), so it has been suggested that the microbial composition of an object could be used to determine its provenance (*5*). Challenges with this approach include variability of resident microbial community abundance over time; similarities of microbial composition between different locations; and the requirement for extensive, expensive, and time-consuming mapping of natural environments.

To circumvent these challenges, we propose the deliberate introduction and use of synthetic, nonviable microbial spores harboring barcodes that uniquely identify locations of interest (e.g., food production areas). These synthetic spores would offer a sensitive, inexpensive, and safe way to map object provenance provided that several important criteria are met, including: (i) the microbes must be compatible with growth at industrial scale; (ii) the synthetic spores must be biocontained and not viable in the wild to prevent adverse ecological effects; (iii) the synthetic spores must persist in the environment and reliably label objects that pass through it; and (iv) the encoding and decoding of information about object provenance must be rapid, sensitive, and specific. Similar barcoding approaches have been explored previously to model pathogen transmission (*6*, *7*) but did not explicitly address those challenges. Here, we report the barcoded microbial spores (BMS) system, a scalable, safe, and sensitive system that uses DNA-BMS mixtures to permit the determination of object provenance (Fig. 1A).

The BMS system leverages the natural ability of spores to persist for long periods in the

[1]Department of Systems Biology, Harvard Medical School, Boston, MA 02115, USA. [2]Laboratory of Systems Pharmacology, Harvard Medical School, Boston, MA 02115, USA. [3]Biological and Biomedical Sciences Program, Harvard Medical School, Boston, MA 02115, USA. [4]Department of Biomedical Engineering and Biological Design Center, Boston University, Boston, MA 02215, USA. [5]Department of Microbiology and Immunobiology, Harvard Medical School, Boston, MA 02115, USA. [6]Department of Biomedical Informatics, Harvard Medical School, Boston, MA 02115, USA. [7]Synthetic Biology Center, Massachusetts Institute of Technology, Cambridge, MA 02139, USA. [8]Wyss Institute for Biologically Inspired Engineering, Harvard University, Boston, MA 02115, USA.
*These authors contributed equally to this work.
†Present address: Quanterix, Billerica, MA 01821, USA.
‡Present address: The University of Illinois College of Medicine, Rockford, IL 61107, USA.
§Corresponding author. Email: michael_springer@hms.harvard.edu

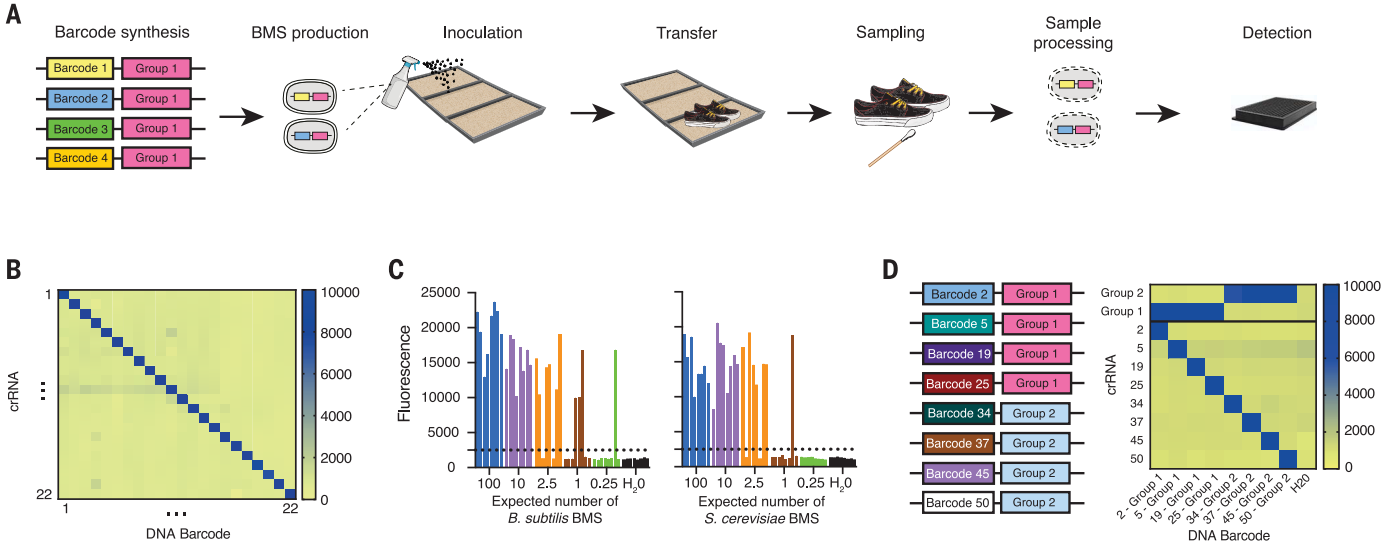


Fig. 1. BMS can be specifically and sensitively detected. (**A**) Schematic of the BMS application and detection pipeline. (**B**) Heatmap of endpoint fluorescence values from in vitro SHERLOCK reactions of all combinations of 22 barcodes and 22 crRNAs assessing specificity of each barcode-crRNA pair. (**C**) Detection limit of *B. subtilis* and *S. cerevisiae* BMS by SHERLOCK (each of the eight biological replicates for each spore concentration are shown). Spore numbers are calculated on a per-reaction basis. (**D**) Heatmap of endpoint fluorescence values from in vitro SHERLOCK reactions testing the specificity of four barcodes for group 1 crRNA and four barcodes for group 2 crRNA as detected by either specific or group crRNA.

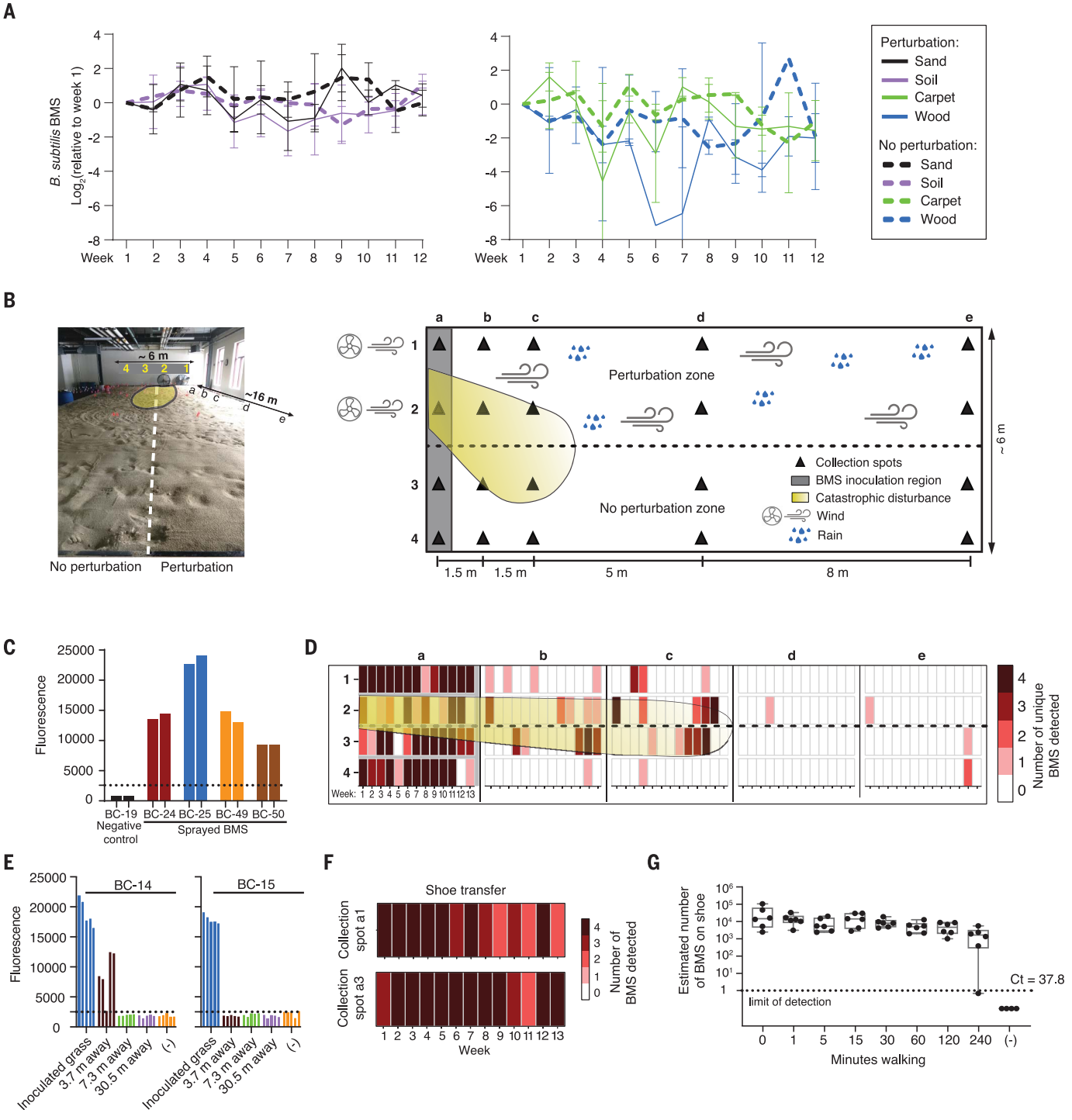
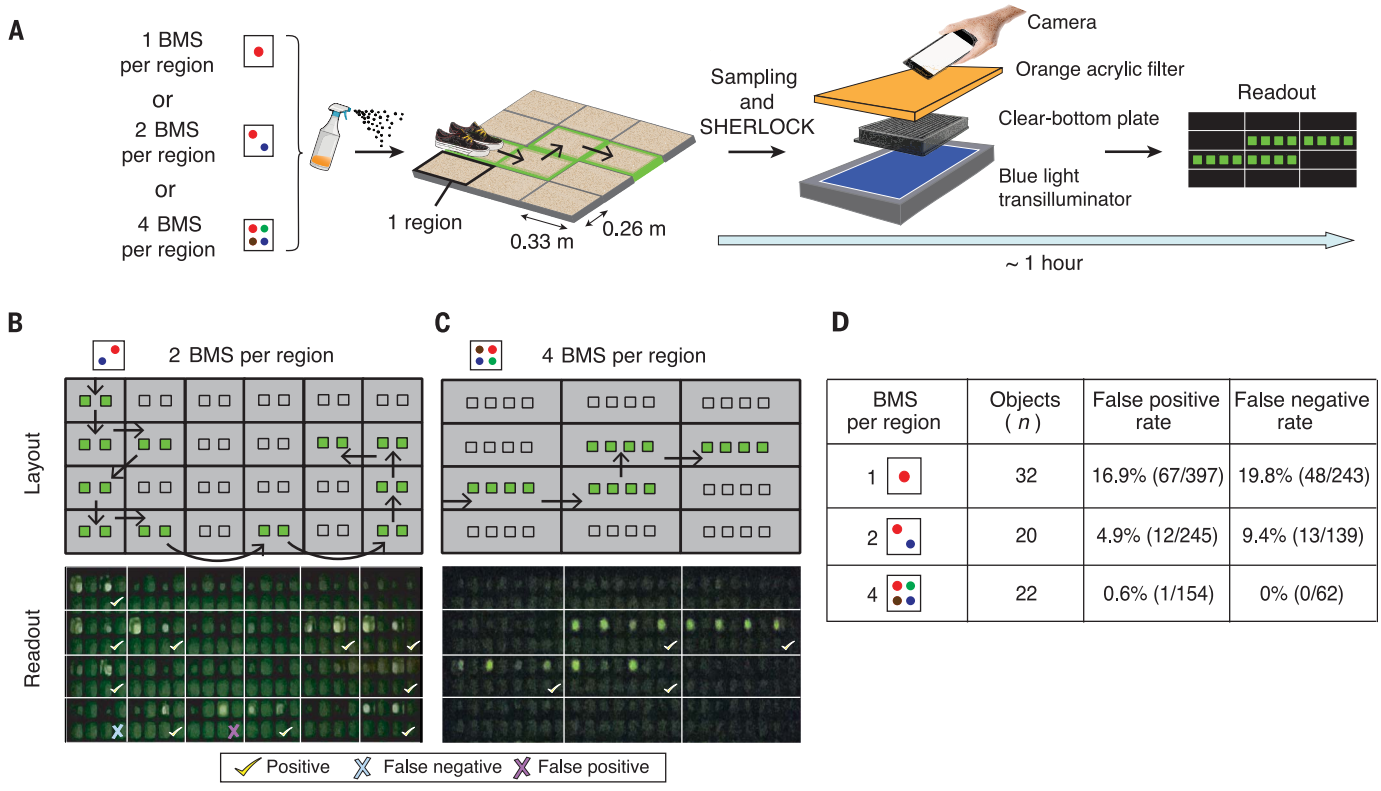


Fig. 2. Persistence, transferability, and maintenance of BMS. (**A**) BMS persisted on sand, soil, carpet, and wood over 3 months on ~1-m^2 test surfaces in incubators. The *y*-axis shows the *B. subtilis* BMS number relative to week 1 levels. Bars represent standard deviations. Perturbations were simulated wind, rain, vacuuming, or sweeping. (**B**) Photograph and schematic of a large-scale (~100-m^2) sandpit. *B. subtilis* BC-24 and BC-25 BMS and *S. cerevisiae* BC-49 and BC-50 BMS were inoculated in the shaded gray region in the diagram. The yellow shadow represents the area over which the top 5 cm of sand from a 1.5-m^2 area of the inoculation region was redistributed after a large fan fell over. (**C**) SHERLOCK signal from the four BMS from the inoculated region "a" in the large-scale experiment. Dashed line is the threshold for positive calls. BC-19 is the negative control. (**D**) BMS persist at collection point "a" (within the inoculated region) and do not spread to collection points "d" or "e." Heatmap depicts the number of BMS (out of four) detected by SHERLOCK at each collection point over 13 weeks. (**E**) BMS persist on grass in an outdoor environment for at least 5 months. The grass region was inoculated with *B. subtilis* BC-14 and 15 BMS. Samples from actual BMS-inoculated region 3.7, 7.3, and 30.5 m away from the inoculated region were tested by SHERLOCK using crRNA 14 and 15 (each of the five biological replicates for each grass region are shown). (**F**) BMS were transferred onto shoes by walking in the inoculated region "a" in the sandpit and were detected by SHERLOCK. (**G**) Abundance of BC-25 BMS on shoes after up to 240 min of walking on noninoculated outdoor areas. The *y*-axis shows the BMS count based on the qPCR standard curve.



D

BMS per region	Objects (*n*)	False positive rate	False negative rate
1	32	16.9% (67/397)	19.8% (48/243)
2	20	4.9% (12/245)	9.4% (13/139)
4	22	0.6% (1/154)	0% (0/62)

Fig. 3. Determining object provenance using BMS and field-deployable detection. (**A**) Schematic of experiment design and field-deployable method to determine previous locations of an object. Each region was inoculated with one, two, or four nonredundant BMS. Green outlines indicate the object path through a subset of the regions. SHERLOCK reactions were imaged using a mobile phone camera to photograph the reaction plate through a filter and under portable blue light illumination. (**B** and **C**) (Top) Path of an object overlaid on a reaction plate. (Bottom) Photograph of SHERLOCK reaction plate overlaid with correct or incorrect calls. The call for each region is denoted by color (yellow check: true positive, blue cross: false negative, purple cross: false positive). (**D**) Statistics for SHERLOCK provenance predictions of objects traversing regions inoculated with one, two, or four BMS per region. The false-positive and false-negative rates for one and two BMS per region were based on the criteria of one or more positive calls; the rates for four specific BMS per region were based on the criteria of two or more positive calls.

environment without growth (*8*). We designed nonredundant DNA barcodes and integrated them into the genomes of *Bacillus subtilis* and *Saccharomyces cerevisiae* spores, creating a set of BMS that can be used combinatorially to provide a nearly infinite set of identification codes. The BMS can be manufactured at scale using standard cloning and culturing techniques, inoculated to surfaces by spraying, and transferred to objects that come into contact with the inoculated surface. To identify barcodes, BMS sampled from objects can be lysed and decoded with a range of methods including SHERLOCK, a recombinase polymerase amplification (RPA) method, coupled with a Cas13a-based nucleic acid detection assay (*9*), quantitative polymerase chain reaction (qPCR), and sequencing (Fig. 1A and fig. S1A).

The BMS are designed not to affect the native environment into which they are applied. First, we used auxotrophic strains that require amino acid supplementation for growth. Second, we made the cells germination deficient. For *B. subtilis* spores, we deleted the genes encoding the germinant receptors and the genes that encode the cell wall lytic enzymes required to degrade the specialized spore cell wall. Incubation of $>10^{12}$ spores generated from this mutant strain showed they were unable to form colonies or grow in rich medium and remained stable and nongerminating at room temperature for >3 months (fig. S2, A and C). For *S. cerevisiae*, we boiled spores for 30 min to heat-kill vegetative cells and spores before application. Incubation of $>10^{8}$ boiled spores on rich medium yielded no colonies (fig. S2, B and D to F). All antibiotic resistance cassettes used to generate the BMS were removed by site-specific recombination to prevent horizontal gene transfer of resistance genes to other organisms in the environment. Finally, the inserted barcode does not encode any gene and should not confer any fitness advantage if horizontally transferred. We tracked the microbiome of soil samples and found that inoculation with BMS had insignificant effects on the microbiome compared with natural changes over time or in response to watering (fig. S3). We also note that *B. subtilis* and *S. cerevisiae* are both commonly found in environmental and food samples.

Multiple BMS can be applied and then decoded simultaneously. We designed a series of tandem DNA barcodes, each with a Hamming distance of >5, allowing $>10^{9}$ possible individual barcodes. To test the specificity of our barcode design in a field-deployable system, we constructed 22 barcodes and their matching CRISPR RNAs (crRNAs) and assayed all permutations in vitro using SHERLOCK. All 22 crRNAs clearly distinguished the correct barcode target (Fig. 1B). To scale the system, we devised a rapid and facile method to screen a large number of barcodes and crRNAs in parallel to eliminate those with cross-reactivity or background. We validated and performed pooled *n-1 barcode* RPA reactions in vitro with corresponding crRNA and water RPA controls testing 94 crRNA-barcode pairs, eliminating 17 for high background and seven

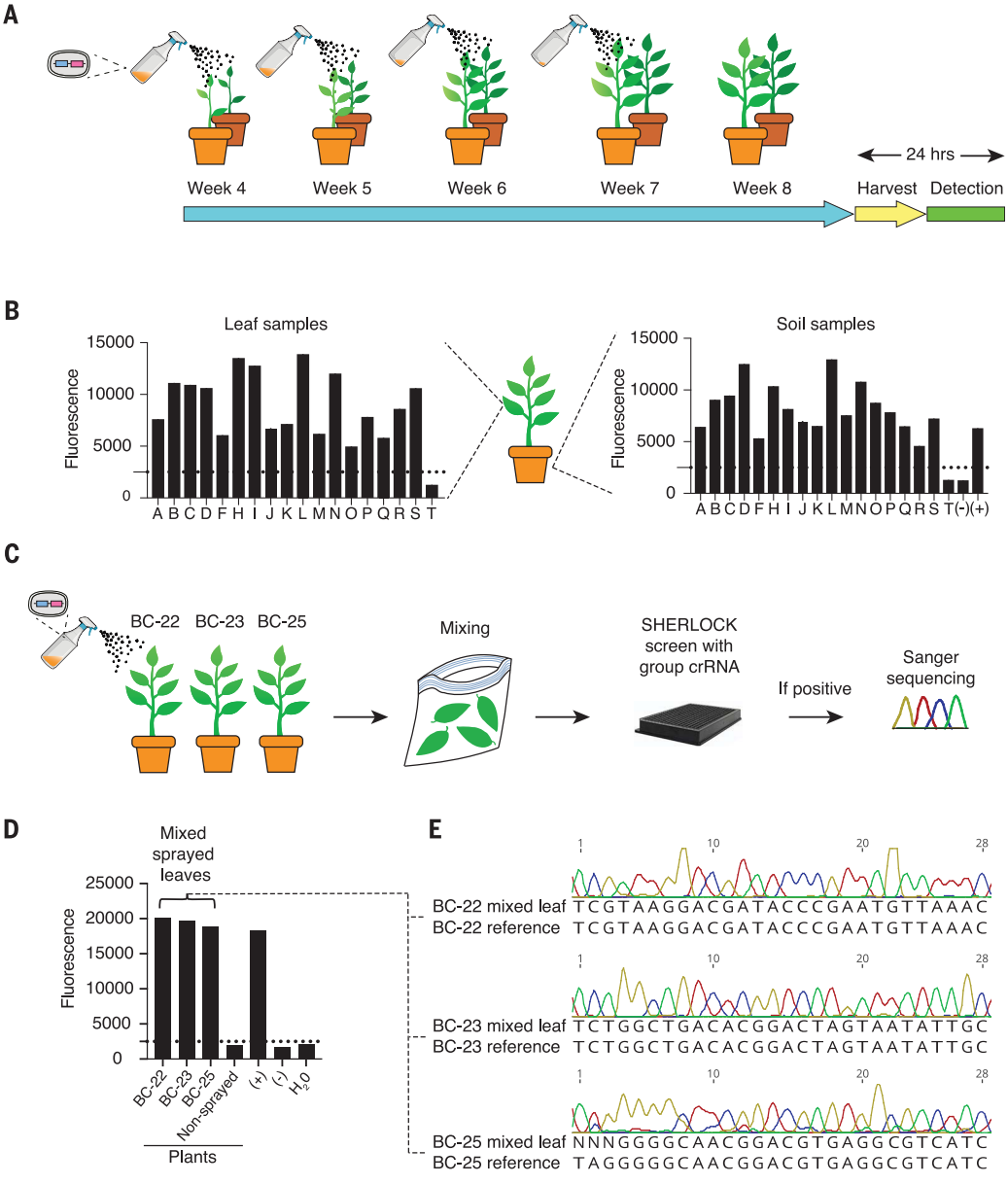


Fig. 4. Determining the provenance of produce using BMS. (**A**) Eighteen plants were inoculated with distinct *B. subtilis* BMS and inoculated once per week (four times total). (**B**) Detection of BMS on plants and soil after harvesting by SHERLOCK with a group crRNA. The *y*-axis shows endpoint fluorescence values. Plants A to S were sprayed with BMS; plant T was not sprayed. (−), negative control without DNA template; (+), positive control from DNA. Dashed line is the threshold for positive calls. (**C**) Leaves from plants inoculated with different BMS were mixed together, SHERLOCK was used to confirm the presence of the BMS, and then Sanger sequencing was used to identify the origin of each leaf. (**D**) Leaves were screened for the presence of BMS by SHERLOCK using a group crRNA. The *y*-axis shows endpoint fluorescence values. (+), group 2 positive DNA; (−), group 1 DNA; (H_2O), water control. (**E**) Sanger sequencing identified the plant of origin of the mixed leaves.

for cross-reactivity (fig. S1, B and C). To test sensitivity and specificity in vivo, we integrated 57 barcodes into *B. subtilis* and 11 into *S. cerevisiae*. We developed an efficient spore lysis protocol using heat and sodium hydroxide (fig. S4), which allowed us to achieve near single-spore resolution for detection using SHERLOCK (Fig. 1C). In vivo and in vitro specificity screenings of crRNA-barcode pairs gave similar results (fig. S1, C and D). In addition, our barcodes are tandemly designed with a specific sequence and a shared group sequence (fig. S1A) to aid in high-throughput detection settings where only a subset of samples contains the BMS of interest. The group sequence is compatible with field-deployable detections and can be used to determine whether a BMS of interest is present before using a second assay to specifically identify the BMS (Fig. 1D). This two-step process solves the throughput limitations of field-deployable detection and lowers the costs of sequencing.

The BMS system is robust and can function on different surfaces in simulated real-world environments. First, in ~1-m^2-scale experiments (fig. S5A and table S6), we used qPCR to detect and quantify BMS directly from surface samples or surface swabs. We found that BMS persisted on sand, soil, carpet, and wood surfaces for at least 3 months with little to no loss over time (Fig. 2A and fig. S5, B and C). Notably, multiple tested perturbations (e.g., simulated wind, rain, vacuuming, or sweeping; see fig. S5A) did not reduce our ability to detect BMS from the surface. Second, we constructed an ~100-m^2 indoor sandpit (Fig. 2B and fig. S6), inoculated one region with BMS (Fig. 2C), and were able to readily detect the BMS for 3 months using SHERLOCK (Fig. 2D and fig. S7). Perturbations did not cause appreciable spreading to noninoculated areas (Fig. 2D and fig. S7); even a catastrophic disturbance in which a fan fell into the sand, only spread the BMS several meters (fig. S6). In an outdoor environment, BMS inoculated on grass was still detectable after 5 months of exposure to natural weather, with minimal spreading outside of the inoculated region (Fig. 2E). This is consistent with low levels of re-aerosolization reported for other spore-forming Bacilli (*10*).

The BMS can be transferred onto objects that pass through test environments. In ~1-m^2-scale testing, BMS could be transferred onto

rubber or wooden objects simply by placing them on the BMS-inoculated surface for several seconds, yielding up to ~100 spores/μl of reaction input (fig. S5D). At ~100-m^2 scale, BMS were reliably transferred onto shoes worn in the inoculated sandpit (Fig. 2F and fig. S8). Furthermore, the BMS transferred onto shoes could still be detected even after walking on noninoculated surfaces for several hours, although BMS counts decreased by twofold with 2 hours of walking as quantified by qPCR (Fig. 2G and fig. S9). We were unable to detect spores on noninoculated surfaces after walking on them with shoes that had traveled through BMS-inoculated regions (fig. S10). We conclude that the BMS can persist in the environment without significant spreading, are transferable onto objects that pass through the environment, are retained on these objects, and can be sensitively and specifically detected using SHERLOCK.

The BMS system can be used to label specific locations of interest to determine whether a person or object has passed through them. We divided different surfaces into grids; inoculated each grid region with one, two, or four nonredundant BMS (Fig. 3A); and traversed them with different test objects (e.g., shoes). To mimic in-field deployment, we used a portable light source, an acrylic filter, and a mobile phone camera to image the SHERLOCK readout (Fig. 3A and fig. S11A) and to determine object provenance (Fig. 3, B and C, and figs. S11 to S13). Provenance could be determined in the field within ~1 hour from sample collection. To evaluate the sensitivity and specificity of our system for determining an object's provenance, we considered different criteria for classification with varying numbers of BMS per region. Object provenance could be determined with a 0.6% (1/154) false-positive rate and a 0% (0/62) false-negative rate if regions were inoculated with four BMS based on the criteria of two or more positive BMS calls. Inoculating with only one or two BMS per region still permitted object provenance to be determined, albeit at the higher error rates (Fig. 3D and fig. S14). Provenance could be determined on all four surface types tested (sand, soil, carpet, and wood) (fig. S13); further validation will be needed to determine error rates in real-world environments. This experiment demonstrates that the BMS can be used to determine object provenance at meter-scale resolution, which would be extremely difficult to achieve using natural microbiome signatures (*11*).

The BMS system offers a flexible and comprehensive approach to determining food provenance. Foodborne illness is a global health issue with an estimated 48 million cases each year in the United States alone (*12*). There is an urgent need for rapid methods for identifying the source of food contamination; current approaches often take weeks and are costly because of the complex modern market chain (*13*). Plants inoculated with *B. subtilis* BMS allowed us to map laboratory-grown leafy plants back to the specific pot in which they were grown (Fig. 4, A and B, and fig. S15). BMS were inoculated four times, beginning 1 week after the first set of leaves appeared, to match recommended inoculation protocols for *Bacillus thuringiensis* (*Bt*) spores, a U.S. Food and Drug Administration–approved biocide that is widely used in agriculture (*14*). One week after the final BMS inoculation, a leaf and a soil sample from each pot were harvested and tested using SHERLOCK. All of the samples were positively detected except for the two plants that had received variant group barcode sequences, demonstrating the specificity of detection (Fig. 4B and fig. S15, B and C). Using Sanger sequencing, we then identified the pot in which each plant was grown for all 18 BMS (fig. S15D). The process from DNA extraction to sequence identification took <24 hours. This time frame could likely be shortened to hours with massively parallelized hybridization-based detection (*15*).

Cross-association of BMS-inoculated plants does not compromise the determination of provenance. To simulate cross-association that could occur during food processing, we mixed leaves from plants that were inoculated with a specific BMS per plant. Unlike the other surfaces that we inoculated, BMS-inoculated plants did not transfer as easily to objects that came into contact with the plants (fig. S5 versus fig. S16, A and B). Although there was detectable transfer between leaves, the amount of transfer still allowed Sanger sequencing to cleanly determine the origin of each leaf (Fig. 4, C to E, and fig. S16, C and D).

Bt could be used to determine food provenance. We used *Bt* spores applied during farming as a surrogate to test whether BMS would persist through conditions of a real-world food supply chain. For plants of known *Bt* inoculation status, we correctly identified all *Bt*-positive and -negative plants (38 total plants) (fig. S17). Furthermore, we detected *Bt* on 10 of 24 store-bought produce items of a priori unknown *Bt* status (figs. S17B and S18). Unexpectedly, BMS and *Bt* spores remained detectable even after washing, boiling, frying, and microwaving (fig. S19), highlighting the potential to determine provenance from cooked foods. These results show the potential for using the BMS system to determine produce provenance.

Our work shows how rationally engineered microbial spores can be manufactured in a high-throughput manner to provide a new solution to the object provenance problem. We have shown that BMS (i) persist in the environment; (ii) do not spread out of the inoculation area; (iii) transfer from soil, sand, wood, and carpet to contacting objects; and (iv) permit sensitive and rapid readout using laboratory and field-deployable methods. The ability to rapidly label objects and determine their provenance in real-world environments has a broad range of applications across agriculture, commerce, and forensics (*2*). Preliminary data suggest that BMS could work across various environments, although extensive validation in a wider range of real-world conditions is needed. Future iterations of our BMS system could be engineered for limited propagation and actively contained for use in highly trafficked areas. This system could also provide time-resolved information about location history, making it useful for an even wider range of applications.

REFERENCES AND NOTES

1. P. M. (Nel) Wognum, H. Bremmers, J. H. Trienekens, J. G. A. J. van der Vorst, J. M. Bloemhof, *Adv. Eng. Inform.* **25**, 65–76 (2011).
2. J. Gooch, B. Daniel, V. Abbate, N. Frascione, *Trends Analyt. Chem.* **83**, 49–54 (2016).
3. S. Lax *et al*., *Science* **345**, 1048–1052 (2014).
4. C. Jiang *et al*., *Cell* **175**, 277–291.e31 (2018).
5. S. Lax *et al*., *Microbiome* **3**, 21 (2015).
6. P. Buckley *et al*., *Appl. Environ. Microbiol.* **78**, 8272–8280 (2012).
7. P. A. Emanuel *et al*., *Appl. Environ. Microbiol.* **78**, 8281–8288 (2012).
8. N. Ulrich *et al*., *PLOS ONE* **13**, e0208425 (2018).
9. J. S. Gootenberg *et al*., *Science* **356**, 438–442 (2017).
10. A. H. Bishop, C. M. O'Sullivan, A. Lane, M. C. Butler Ellis, W. J. Sellors, *Lett. Appl. Microbiol.* **64**, 364–369 (2017).
11. R. I. Adams, A. C. Bateman, H. M. Bik, J. F. Meadow, *Microbiome* **3**, 49 (2015).
12. E. Scallan, P. M. Griffin, F. J. Angulo, R. V. Tauxe, R. M. Hoekstra, *Emerg. Infect. Dis.* **17**, 16–22 (2011).
13. C. W. Hedberg *et al*., *Emerg. Infect. Dis.* **14**, 311–313 (2008).
14. G. Sanahuja, R. Banakar, R. M. Twyman, T. Capell, P. Christou, *Plant Biotechnol. J.* **9**, 283–300 (2011).
15. M. Sarwat, M. M. Yamdagni, *Crit. Rev. Biotechnol.* **36**, 191–203 (2016).
16. B.-M. Koo *et al*., *Cell Syst.* **4**, 291–305.e7 (2017).
17. A. J. Meeske *et al*., *Proc. Natl. Acad. Sci. U.S.A.* **112**, 6437–6442 (2015).
18. P. Schaeffer, J. Millet, J. P. Aubert, *Proc. Natl. Acad. Sci. U.S.A.* **54**, 704–711 (1965).
19. H. C. Lim *et al*., *PLOS Genet.* **15**, e1008284 (2019).
20. A. A. Horwitz *et al*., *Cell Syst.* **1**, 88–96 (2015).
21. D. M. Gohl *et al*., *Nat. Biotechnol.* **34**, 942–949 (2016).
22. E. Bolyen *et al*., *Nat. Biotechnol.* **37**, 852–857 (2019).
23. B. J. Callahan *et al*., *Nat. Methods* **13**, 581–583 (2016).
24. Q. Chang, Y. Luan, F. Sun, *BMC Bioinformatics* **12**, 118 (2011).
25. L. A. Klobutcher, K. Ragkousi, P. Setlow, *Proc. Natl. Acad. Sci. U.S.A.* **103**, 165–170 (2006).
26. P. S. V. Devi, T. Ravinder, C. Jaidev, *J. Invertebr. Pathol.* **88**, 163–168 (2005).

ACKNOWLEDGMENTS

We thank participating laboratory members for useful feedback, O. Mazor of the HMS Research Instrumentation Core Facility for technical consultations and instrument design and fabrication, and K. Rhea and P. Buckley for BMS-labeled grass samples. **Funding:** This work was supported by DARPA BRICS grant no. HR001117S0029. J.Q. is supported by an NSF GRFP. **Author contributions:** J.Q., Z.L., C.P.M., H.Y.J., R.C.B.-O., and S.A.B. conceived the study and performed most key experiments and data analysis (supervised by M.S.). J.Q., Z.L., C.P.M., H.Y.J., R.C.B.-O., S.A.B., and F.H.R.-G. wrote the paper (with M.S. and D.Z.R.). V.J., A.S., K.S., L.A., G.J., M.A., M.E.P., M.M., L.L., and S.V.O. assisted with technical

experiments. S.A.B. purified Cas13 proteins. R.C.B.-O. and F.H.R.-G. designed and constructed *B. subtilis* strains and purified spores (supervised by D.Z.R.). D.Z.R., P.A.S., A.S.K., and M.B. provided critical insights. All authors reviewed the manuscript. **Competing interests:** U.S. provisional patent application no. US 62/958,512 has been filed on behalf of M.S. and the President and Fellows of Harvard College. **Data and materials availability:** All data are available in the article or the supplementary materials.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/368/6495/1135/suppl/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S19
Tables S1 to S6

View/request a protocol for this paper from *Bio-protocol*.

13 December 2019; accepted 31 March 2020
10.1126/science.aba5584

Barcoded microbial system for high-resolution object provenance

Jason Qian, Zhi-xiang Lu, Christopher P. Mancuso, Han-Ying Jhuang, Rocío del Carmen Barajas-Ornelas, Sarah A. Boswell, Fernando H. Ramírez-Guadiana, Victoria Jones, Akhila Sonti, Kole Sedlack, Lior Artzi, Giyoung Jung, Mohammad Arammash, Mary E. Pettit, Michael Melfi, Lorena Lyon, Siân V. Owen, Michael Baym, Ahmad S. Khalil, Pamela A. Silver, David Z. Rudner and Michael Springer

Science **368** (6495), 1135-1140.
DOI: 10.1126/science.aba5584

DNA barcodes in small packages

Under adverse environmental conditions, some microorganisms form spores that provide robust protection for genetic material. Qian *et al.* developed a system in which DNA barcodes are encapsulated inside nongerminating microbial spores and can be dispersed on objects or in the environment (see the Perspective by Nivala). These barcoded spores provide a durable, specific marker that can be read out quickly with simple equipment. When applied to soil, the spores can be transferred to and from objects around them, enabling tracking at meter-scale resolution. On plant leaves, the spores are not readily transferred, and the authors demonstrate a potential use for tracking agricultural products.

Science, this issue p. 1135; see also p. 1058