MARINE ROBOTICS

Revealing ocean-scale biochemical structure with a deep-diving vertical profiling autonomous vehicle

John A. Breier[1]*, Michael V. Jakuba[2], Mak A. Saito[2], Gregory J. Dick[3,4,5], Sharon L. Grim[3,†], Eric W. Chan[1], Matthew R. McIlvin[2], Dawn M. Moran[2], Brianna A. Alanis[1], Andrew E. Allen[6,7], Chris L. Dupont[6], Rod Johnson[8]

Vast and diverse microbial communities exist within the ocean. To better understand the global influence of these microorganisms on Earth's climate, we developed a robot capable of sampling dissolved and particulate seawater biochemistry across ocean basins while still capturing the fine-scale biogeochemical processes therein. Carbon and other nutrients are acquired and released by marine microorganisms as they build and break down organic matter. The scale of the ocean makes these processes globally relevant and, at the same time, challenging to fully characterize. Microbial community composition and ocean biochemistry vary across multiple physical scales up to that of the ocean basins. Other autonomous underwater vehicles are optimized for moving continuously and, primarily, horizontally through the ocean. In contrast, *Clio*, the robot that we describe, is designed to efficiently and precisely move vertically through the ocean, drift laterally in a Lagrangian manner to better observe water masses, and integrate with research vessel operations to map large horizontal scales to a depth of 6000 meters. We present results that show how *Clio* conducts high-resolution sensor surveys and sample return missions, including a mapping of 1144 kilometers of the Sargasso Sea to a depth of 1000 meters. We further show how the samples obtain filtered biomass from seawater that enable genomic and proteomic measurements not possible through in situ sensing. These results demonstrate a robotic oceanography approach for global-scale surveys of ocean biochemistry.

INTRODUCTION

On Earth, the cycling and availability of carbon, nitrogen, phosphorus, and oxygen, as well as other essential nutrients and cofactors, are strongly influenced by processes that occur within the ocean over basin scales (*1*–*3*). Planktonic marine microorganisms, which collectively account for close to half of planetary primary production, drive these processes (*4*). Redfield (*5*) described the reciprocal, and chemical, relationships between ocean biology and the environment many years ago, suggesting that this relationship was the result of plankton communities building their cellular biochemical constituents. Geochemists and ecologists have incorporated these relationships into large-scale conceptual models of nutrient flows, generally referred to as biogeochemical cycles. However, it has always been challenging to identify the essential processes to model and validate their representation and predictions. While our understanding of the processes connecting planktonic microbiology to ocean chemistry has grown substantially over the years (*6*–*8*), our appreciation of the complexities in these relationships has grown, too (*9*–*11*). However, the recent inclusion of genomic (*12*, *13*), transcriptomic (*14*), proteomic (*15*), and metabolomic (*16*) measurements in ocean studies is creating great potential to understand the specific biochemical mechanisms coupling ocean life and environmental chemistry. Currently, there is keen interest in better understanding these natural processes (*17*) and urgency to do so because anthropogenic processes are altering the physical and chemical environment of the ocean, with the potential to reduce the ocean's ability to buffer large-scale climatic shifts (*18*).

To improve our understanding of Earth's biogeochemical cycles, we developed a robotic vehicle to facilitate water column ocean mapping of microbial and biochemical data (Fig. 1). This robot, called *Clio*, collects dissolved and particulate samples and measurements from the water column for microbial community proteomics, genomics, transcriptomics, organic and inorganic chemistry, and related physical data. These data types include biochemicals, sparsely mapped at present, and supporting parameters that must be measured synoptically. Sensors can measure only a subset of the necessary data; thus, sample return for shore-based analyses is essential. *Clio* collects samples and sensor data in the form of (i) high-resolution ocean depth profiles, (ii) ocean basin scale transects of the same, and (iii) targeted data collection from biochemical hotspots hidden within the ocean. Although we can make some biochemical and chemical measurements with small sample volumes (milliliters to liters), we require large volumes for other measurements (e.g., proteomics, organic carbon, and trace metals). Therefore, *Clio* is capable of processing several hundred liters of seawater per sample to return material for subsequent high-throughput laboratory analyses. *Clio* is designed for diving to depths of 6000 m and can, therefore, access most of the ocean. *Clio* drifts freely in the lateral plane to better observe ocean water masses. *Clio* operates autonomously without a tethered connection to a surface vessel; this frees the robot from the restricted motion that a tether imposes, and it frees the surface vessel to carry out synoptic science activities (fig. S1). Before *Clio*, an ensemble of sequential techniques, involving the repeated lowering and recovery of cabled equipment from research vessels, was required to collect similar full ocean depth datasets.

[1]School of Earth, Environmental, and Marine Sciences, The University of Texas Rio Grande Valley, Edinburg, TX 78539, USA. [2]Woods Hole Oceanographic Institution, Woods Hole, MA 02543, USA. [3]Department of Earth and Environmental Sciences, University of Michigan, Ann Arbor, MI 48109, USA. [4]Center for Computational Medicine and Bioinformatics, University of Michigan, Ann Arbor, MI 48109, USA. [5]Department of Ecology and Evolutionary Biology, University of Michigan, Ann Arbor, MI 48109, USA. [6]Microbial and Environmental Genomics, J. Craig Venter Institute, San Diego, CA 92121, USA. [7]Integrative Oceanography Division, Scripps Institution of Oceanography, University of California, San Diego, La Jolla, CA 92093, USA. [8]Bermuda Institute of Ocean Sciences, St. George's, GE 01, Bermuda.
*Corresponding author. Email: john.breier@utrgv.edu
†Present address: Space Science and Astrobiology Division, NASA Ames Research Center, Moffett Field, CA 94035, USA.

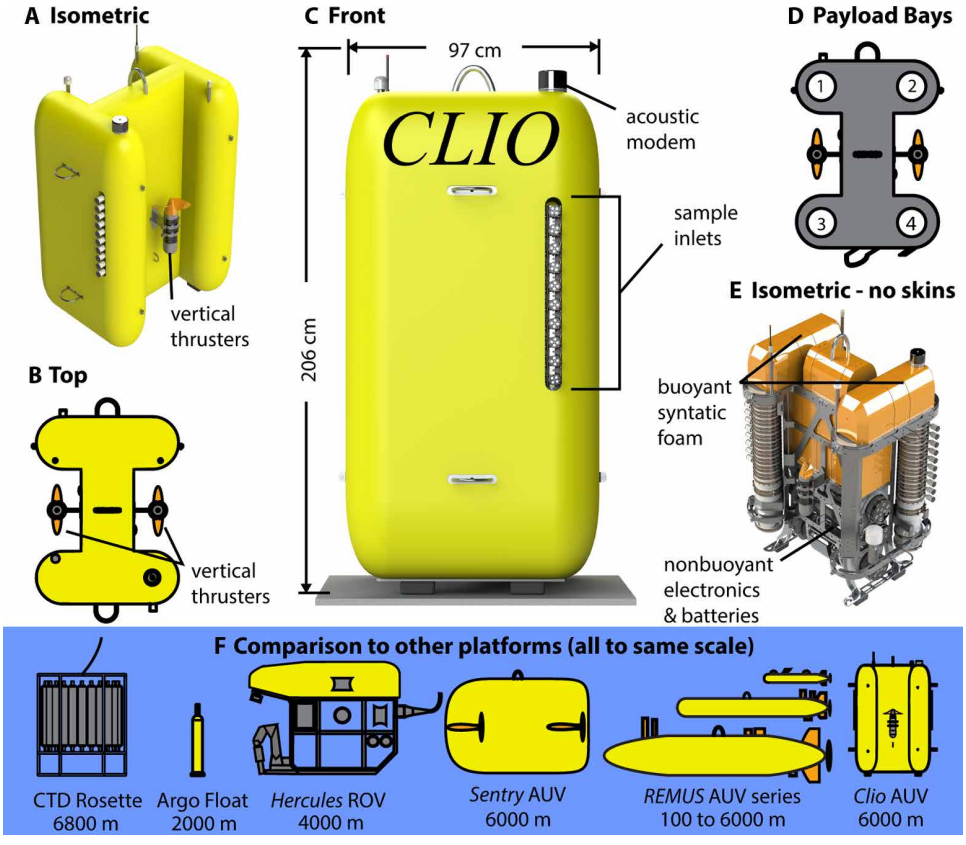

Fig. 1. *Clio* biochemical AUV. *Clio* vertically profiles the ocean water column to 6000 m and drifts horizontally in a Lagrangian manner. The (**A**) H-shape hull protects the (**B**) thrusters and creates payload bays for sensors and samplers, the inlets of which extend through the (**C**) vehicle sides. The (**D**) payload bays are arranged at the corners of the H-shape. The (**E**) floatation is near the vehicle's top, and the heaviest components are near the bottom. *Clio* is (**F**) comparable in size to other deep-diving vehicles and sampling platforms. However, its design and purpose are unique and directed at facilitating the dual goals of global ocean mapping and providing samples for biochemistry.

Existing vehicle designs are not well suited to diving deep, systematically crossing an ocean in short order, and processing large volumes of water for sample return (Fig. 1F). We addressed this challenge by developing a new robot for the most unfulfilled and currently limiting aspect of these objectives while making use of existing ocean surface vessels where they were most capable: quickly transiting between stations. Specifically, we optimized *Clio* to execute single vertical ocean data collection dives as effectively as possible because Langragian motion is better for sampling. Thus, the *Clio* design foregoes any lateral control. During deployment, *Clio* is launched from a surface vessel and conducts an autonomous sample and data collection mission over a predetermined depth range. *Clio* is recovered by the surface vessel at the end of each dive, typically lasting 7 to 14 hours, and transported to the next dive location where this process is repeated (Fig. 2 and fig. S1). In this way, *Clio* can acquire data and samples within weeks that span ocean basins.

Global-scale ocean studies

In 1934, Redfield (*19*) hypothesized that the ratios of inorganic nutrients in seawater were in a reciprocal self-regulating relationship with marine microorganisms based on the analysis of global datasets. This was not obvious and only exists in the oceans because phytoplankton dominate marine biomass and oceans are constantly mixing, but this only became apparent through the examination of global datasets. Subsequent global-scale studies have also transformed our understanding of ocean processes. The Geochemical Ocean Sections Study (GEOSECS) improved our understanding of ocean circulation (*20*). The Tropical Ocean Global Atmosphere (TOGA) program allowed us to predict better El Niño and Southern Oscillation events (*21*). The World Ocean Circulation Experiment (WOCE) and the Joint Global Ocean Flux Study (JGOFS) informed our initial understanding of ocean circulations' influence on Earth's climate and the ocean carbon cycle, respectively (*22*, *23*). Current global ocean studies, including GEOTRACES (*24*), build on these previous efforts; more details are in the Supplementary Materials.

An important part of these studies has been the systematic collection of data for the creation of ocean sectional maps of the water column. The field methods for creating ocean sectional maps have generally been consistent: (i) collect water and particulate samples from multiple depths in a vertical profile from the ocean surface to the deepest depth, (ii) repeat this process at multiple stations spread along a transect crossing an ocean, and (iii) complete multiple transects over as many research cruises to create a network of sectional maps that cover Earth. To date, ocean sectional studies rely on wire-based water column sampling technology, including water sampling rosettes and wire-deployed pumps (fig. S1B). The time to collect samples this way is determined by winch wire speed (typically 0.5 to 0.7 m s^{-1}), sample depths, pumping time, processing time, and the number of sampling deployments, which must be sequential to avoid entangling wires. It can take 1 to 2 days per station, sometimes more, for modern sectional studies such as GEOTRACES to acquire the necessary samples for all measurements (more details are in the "Sectional cruise methodology" section in the Supplementary Materials). This results in sectional cruises that are typically 30 to 65 days long, a duration that can be shortened or made more productive with robotic tools like *Clio*.

The biochemistry of the oceans, including microbial community genomics, transcriptomics, proteomics, and metabolomics, has yet to be widely measured. Global genomics surveys to date, such as the Global Ocean Sampling program (*25*) and the Tara Expedition (*26*), have focused almost entirely on the sunlit euphotic zone. The biochemistry of the rest of the ocean is largely unmapped. This fact and the results of previous ocean basin studies are inspiring efforts to launch new global microbiology and biochemistry studies (*27*), such as the proposed Biogeoscapes effort.

Clio is designed to enable global studies of this kind by speeding the collection of the necessary data. *Clio* collects a full suite of samples

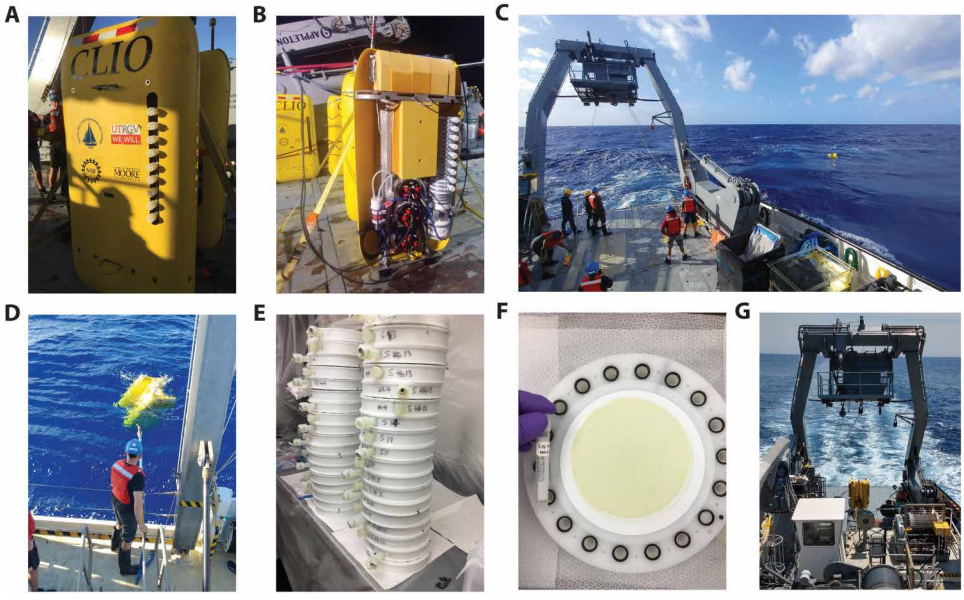

Fig. 2. Ocean profiling with *Clio*. (**A**) *Clio* is designed to collect biochemically relevant data and samples from the ocean water column. (**B**) *Clio* onboard R/V *Atlantic Explorer*, with its front and back skins removed to access the payload bays for dive preparation. (**C**) *Clio* waiting for recovery after a successful dive. (**D**) *Clio* being hooked for recovery. (**E**) One dive's worth of 18 samples returned to a shipboard clean room for processing and subsampling. (**F**) A particulate sample filtered by *Clio* at 70-m depth from 47.7 liters of oligotrophic seawater. (**G**) *Clio* being transported by R/V *Atlantic Explorer* to the next dive.

synoptically during a single dive in a fraction of the time required by current methods. *Clio* also enables autonomous sample collection concurrent with existing wire-based techniques, or other activities, which frees the research vessel and investigators to conduct more research.

Robotic oceanography for biogeochemical studies

Robotic oceanography techniques have not been used for global ocean sectional studies, but they have been used to study ocean biogeochemical processes (Fig. 1F). Profiling floats, and more recently autonomous underwater gliders, are used to collect sensor data to support biogeochemical studies, and profiling floats were part of the WOCE study (*28*). Remotely operated vehicles (ROVs) and autonomous underwater vehicles (AUVs) have been used to study biological and chemical processes in niche environments throughout the ocean.

Profiling floats and gliders are used for large-scale observations of specific core ocean parameters (*29*–*31*). They are designed for long-term deployments lasting months to years. They control their depth by changing their buoyancy using variable ballast. They typically measure salinity, temperature, dissolved oxygen, chlorophyll, and optical proxies of particle concentration but are also used for other measurements (*32*, *33*). Argo floats are one of the most common profiling float designs; more than 4000 Argo floats have been deployed throughout the ocean. They are typically deployed once for a lifetime duration of 4 or more years. During that time, they drift in a Lagrangian manner. They only control their depth to a limited number of set points, and most designs are limited to the upper 2000 m of the ocean. Profiling gliders have wings, control pitch and roll in addition to buoyancy and can move horizontally as well as vertically (*34*–*36*). These platforms typically have small payload and battery capacities that cannot accommodate large sample-processing systems. Both platforms are optimized for travel over great distances but over long time scales that would (i) limit the sample collection rate even if payload constraints were not limiting and (ii) raise greater concern regarding sample preservation. Floats and gliders are not currently well suited to the type of sample return necessary for biochemical sectional studies.

ROVs have been used as ocean-observing platforms since the 1980s (*37*–*39*). Among ocean vehicles, they have the highest payload capacities and are the most reconfigurable. The precision navigation capabilities of ROVs have enabled targeted sampling in and around seafloor seeps and hydrothermal biological "hotspots" (*40*–*43*). At these sites, the concentration of microbial life and chemical nutrients can be high; thus, processing large samples is unnecessary in many cases. However, collecting material emitted from these systems up into the water column as it mixes and dilutes with seawater does require processing large samples; the sampling systems developed for *Clio* inherits design principles from instruments developed to return samples from deep-sea hydrothermal plumes (*43*, *44*). Nevertheless, ROVs require cables, cables require winches, and cables and winches typically constrain the supporting research vessel to one activity at a time. For ocean sectional studies, ROVs provide little benefit over existing cabled sampling technologies.

AUVs have been used as ocean observing platforms since the 1990s (*45*–*47*). Optimized for longer-distance surveys and more task-specific observations than ROVs, AUVs trade smaller payload capacities for lower drag and increased range. Fully autonomous, they can operate independently from surface ships. Most AUV designs have been multipurpose. They are used widely for mapping operations, particularly at the seafloor (*47*) but also under ice (*48*). AUVs have been adapted to the study of surface plankton processes (*49*, *50*) and equipped with dedicated sample return systems to support these studies (*51*, *52*). In the coastal surface ocean, phytoplankton concentrations can be high; as a result, processing large samples is unnecessary. Therefore, AUV sample return systems designed for these surface ocean studies typically collect and process water volumes on the order of 1 to 2 liters per sample. A few AUV applications have made use of deep-sea sampling system designs and are capable of collecting larger volumes for a broader range of biochemical analyses (*53*, *54*).

Moreover, most AUV designs are based on the familiar torpedo shape and optimized for long-range lateral motion. For example, the Tethys Long Range AUVs have a range of at least 1800 km (*55*), and the long-range Autosub version is designed for a range of 6000 km (*56*). These AUVs and others like them can also dive deep, but they are not well suited to repeatedly profiling to full ocean depth nor doing so in a single deployment.

Clio is a reimagination of AUV design applied to vertical profiling. It builds upon sample return instrumentation and techniques developed for the deep sea, where biomass concentrations are low, and applies them to the full-water column. It is designed and optimized for a single purpose and a mode of science operations orthogonal to previous AUV applications. Instead of collecting core datasets at global scales or intensively observing regional processes, we designed *Clio* to complete an intensive study of global-scale biochemistry.

Experimental setting

We tested *Clio* over several years, six cruises, and 26 dives. Initial sea trials were in 2017 near the shelf break of the eastern Atlantic coast of North America, just north of the Gulf Stream and the Sargasso Sea. During 2018 and 2019, we deployed *Clio* repeatedly, on a semiquarterly basis, at the Bermuda Atlantic Time-series Study (BATS) station (31.66667°N, −64.16667°W) within the Sargasso Sea. The Sargasso Sea is an oligotrophic region of the North Atlantic Ocean where Redfield (*19*) made some of his initial observations in 1933.

Contributions

Clio is uniquely designed to facilitate global-scale studies of ocean biochemistry, move vertically through the water column with high precision, and specifically return samples from large swaths of the ocean ranging in depths from the surface to 6000 m. *Clio's* range encompasses full ocean depths everywhere except for hadal trenches. *Clio* is capable of flexible, precise vertical motion that few other ocean robots can perform, and none to our knowledge over this depth range. We demonstrate that *Clio* can move through the water column at a controlled rate as slow as desired or up to 0.8 m s^{-1} and settle at desired depths to within 50 cm in less than a minute. It can hold a vertical station within a depth window of less than 5 cm for multiple hours while sampling and sensing. *Clio* is isolated from the influence of ship heave, unlike for tethered instruments. Unlike other AUVs, *Clio* can hold station vertically while drifting in a Lagrangian manner laterally. This allows high-resolution, vertical profiling using both sensors and samplers that are ideal for targeted observations of fine chemical or biological gradients that exist within the ocean water column. Importantly for its mission, *Clio* can acquire and return a broader range of sample types, from larger water volumes, than other existing ocean robots. From samples collected in the North Atlantic Ocean and the Sargasso Sea, we demonstrated that *Clio* can be used to make metagenomic and proteomic measurements and collect sensor data and samples for more routine nutrient measurements. In June 2019, *Clio* completed its first sectional cruise in a series of nine dives spanning a 1144-km transect between Bermuda and Woods Hole, Massachusetts, USA, through the Sargasso Sea and the Gulf Stream. This transect focused on the upper 1000 m of the water column and also included *Clio*'s deepest dive to date, to 4100 m, demonstrating *Clio*'s ability to conduct ocean-crossing sectional studies.

RESULTS

In this section, we describe our AUV design for biochemical ocean mapping. We also describe the design of this robot's sample return systems, which it uses to manipulate seawater to extract samples for return. We then report on the results of field trials where the robot was used to conduct high-precision sampling within the water column.

Designing a biochemical AUV for global ocean mapping

Clio's unique H-shape design is a synthesis of vehicle control strategy, deck-side operational considerations, and data collection criteria (Fig. 1, "Vehicle design" section in Materials and Methods). We designed this robot to efficiently conduct single vertical profiles to 6000 m, stopping and loitering at many arbitrary depths to collect data and samples for return. More specifically, we designed *Clio* to autonomously collect the observations made at sectional cruise stations (e.g., GEOTRACES) but in a fraction of the time and surpassing their spatial resolution. These operations have to be repeated many times, so the robot has to be logistically simple to stage, deploy, recover, and ship worldwide.

Clio's streamlined hull is symmetric about the horizontal plane such that drag is approximately equal during ascent and descent. Stability is achieved, for this otherwise unstable hull form, by using a large separation between the vehicle's centers of mass and buoyancy. This configuration is stable up to the speed at which the destabilizing hydrodynamic moment exceeds the hydrostatic stabilizing moment. Scale model tow tank tests validated this design before manufacturing, and the full-scale field tests confirmed it (*57*). Operation above the critical speed does not endanger the vehicle but results in a limit cycle consisting of large (20° to 30°) oscillations in pitch and roll, decreasing transit efficiency.

Clio is positively buoyant at all depths. Positive buoyancy mitigates the risk of vehicle loss in the event of a control system or actuator failure. Vertical thrusters control its depth. Variable-density ballast control was also considered, but *Clio* must move rapidly between, and settle precisely at, target depths. This is easier to achieve with active thrusting over this depth range. *Clio* must thrust downward to hold depth while sampling, and it spends most of its time during a deployment holding station. It must expend energy to do so, but the energy expended for propulsion is comparable to the power it expends pumping seawater for sample collection. Unlike other AUVs designed for deep-sea applications, *Clio* does not use expendable ballast to speed descent or for emergency ascents; the metals typically used for these weights represent cleanliness concerns, and their use also complicates vehicle operations and logistics. *Clio* uses ambient pressure air bladders to increase positive buoyancy at the surface where it is most valuable: to loft antennas above the sea surface. Their volume decreases rapidly with depth so that their impact on propulsion power consumption is negligible below approximately 100-m depth. To minimize the increase in positive buoyancy with depth as seawater density increases, we favored the selection of materials with lower bulk moduli to more closely match the average compressibility of seawater (*58*).

Thrusters enable high-performance depth control for minimal complexity. *Clio* uses conventional proportional, integral, and derivative position control to command the desired thruster winding current and a Paroscientific Series 8000 Digiquartz pressure sensor running at 6 Hz to provide depth feedback. DC brushless motor controllers run an internal control loop to regulate winding current. The desired thrust is saturated as a means of controlling speed between stations; however, the result depends on ballast condition. Closed-loop speed control will be implemented in the future.

Clio and its sample return systems are sized and shaped in large part to use the same filter membranes commonly used for comparable oceanographic studies using existing wire-based sampling systems. The orientation of the sampler containers, such that filter membranes are parallel to the horizontal, is made to match existing

systems. Samples collected in this way produce a radially uniform pattern across the filter surfaces, which is essential to achieve because the filters are subdivided on the basis of this assumption; this is part of the methodology used with existing wire-based systems, and we match this with *Clio*. *Clio* is also constructed of materials selected to maximize sample integrity and cleanliness. In particular, ferrous alloys were all but eliminated. Iron is a trace and limiting nutrient in seawater. Even minute concentrations of anthropogenic iron introduced from ferrous alloys on the vehicle could influence the trace metal measurements or the biochemical measurements made with this system. In place of ferrous alloys, aluminum, titanium, and nickel alloys were used. For similar reasons, the use of oil-filled electrical junction boxes, standard on some ROVs and AUVs, was minimized, and where possible liquid fluorocarbons were used in their place. When the mineral oils commonly used for these applications inevitably leak out of their reservoirs, they coat the outer surfaces of vehicle components with a residue that can remain over many dives and become entrained into the sampling systems, potentially biasing them for some organic analyses. The fluorocarbon substitutes are more volatile than mineral oils, and residues quickly evaporate, minimizing this concern (*59*). The location of the thrusters and the shape of the vehicle are designed to minimize the influence of propeller wash on sensor data or the collected samples. The thrusters are located within vertical channels on the sides of the vehicles. The sensors are arranged at the bottom of the vehicle, which permits the collection of undisturbed sensor profile data during descent. The intakes for the sample return system are colocated with the sensors on sides orthogonal to the thrusters.

Manipulating seawater

Clio's primary task is to extract material from seawater and return it for analyses that cannot be measured by in situ sensors. Therefore, *Clio* has four payload bays (each 25 cm by 25 cm by 115 cm), each capable of containing independent, custom, high-volume in situ water filtration sampling systems and whole-water and filtrate subsampling systems (Fig. 3A). Each sampler can collect up to 19 samples depending on media size. For operations to date, we outfitted *Clio* with sampling systems in two payload bays and configured them to collect nine sample sets each consisting of (i) particulate material filtered onto 142-mm-diameter, 0.2-μm–pore size media, (ii) 200 ml of filtrate, and (iii) 200 ml of unfiltered water ("Biochemical methods" section in Supplementary Materials). Particulate samples can be preserved in situ by preservative administration. Depending on biomass concentration, this configuration typically filters 40 to 300 liters of water per sample and 1500 to 3500 liters of water total per dive; even so, *Clio* has substantial unused payload for future instrumentation (table S1).

The sample-processing and return systems are the most essential and unique tools that the robot has for interacting with the environment (see Materials and Methods). They are effectively this robot's manipulators. Cleanliness was a design concern throughout the vehicle, but especially so in the sample return systems. Apart from electronics and motor parts contained within the liquid fluorocarbon–filled stepper housing, nearly all other parts are plastic: high-density polyethylene, acetal copolymer, polycarbonate, polytetrafluoroethylene, and stereolithography resin (Durable, Formlabs). Titanium and nickel alloys were used selectively where strength was essential. Each sample container unit has an independent intake to preclude cross-contamination. The custom sample container units were designed to interconnect between the sampling valve and all other individual sample units in the set without the need for tubing (Fig. 3B), which is time consuming to interconnect and difficult to clean (*42*). The sample containers themselves are designed for critical cleaning and to facilitate sample extraction in shipboard fabricated cleanrooms (Figs. 2, E and F, and 3C).

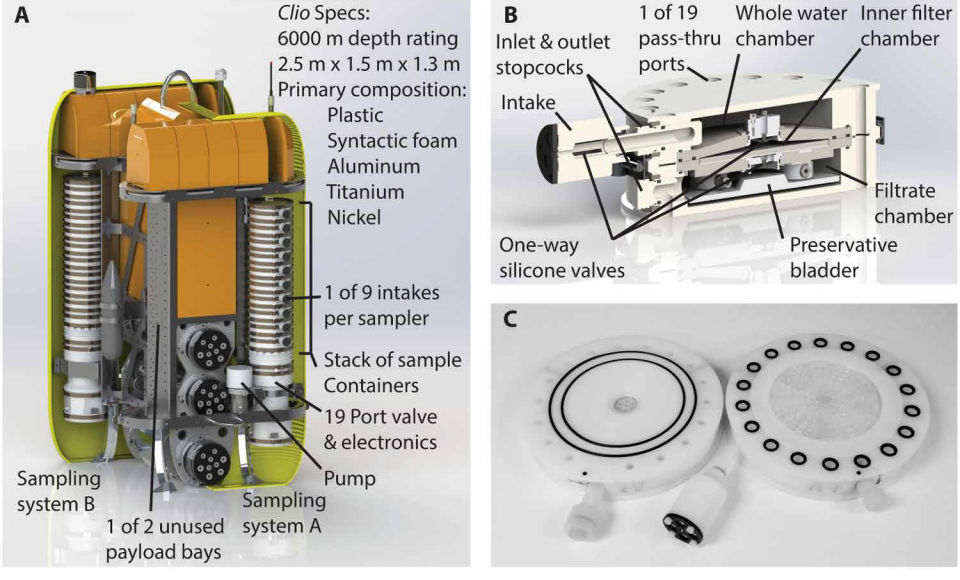

Fig. 3. *Clio* sampling system. (**A**) *Clio* with most skins removed to reveal three of four payload bays (one empty). In this configuration, two bays contain identical large-volume sample collection systems. Additional payload can be mounted in unused bays or external to the skins. To maximize sample return when using larger filters (i.e., 142-mm diameter), we designed (**B**) custom sample containers that combine the collection of 200 ml of whole water, 200 ml of filtrate, and particulate material collected by filtration into a single compact unit. These units have one-way silicone valves that allow injection of preservative into the filter chamber immediately after collection to preserve RNA, for example, until (**C**) shipboard sample extraction.

Experimental setup

Several dive types were conducted during the field trials described here. During initial sea trials, a series of five dives were executed to increasing depths, the last of which was to 2000 m in a water depth of 4100 m. During the dives at BATS station, we tested a number of vehicle configurations, diving to depths of 2000 m. The most frequent dive profiles to date were 1000-m dives completed at BATS and during the Bermuda–to–Woods Hole transect. Near the end of the Bermuda–to–Woods Hole transect, *Clio* completed a dive to 4100 m. The configuration of the sample return systems for these deployments is described in the "Biochemical methods" section in the Supplementary Materials.

AUV performance, endurance, and power efficiency

The typical 1000-m AUV profile took 16 hours to complete (e.g., Fig. 4). While diving, *Clio* was subject to Lagrangian horizontal transport dependent on the

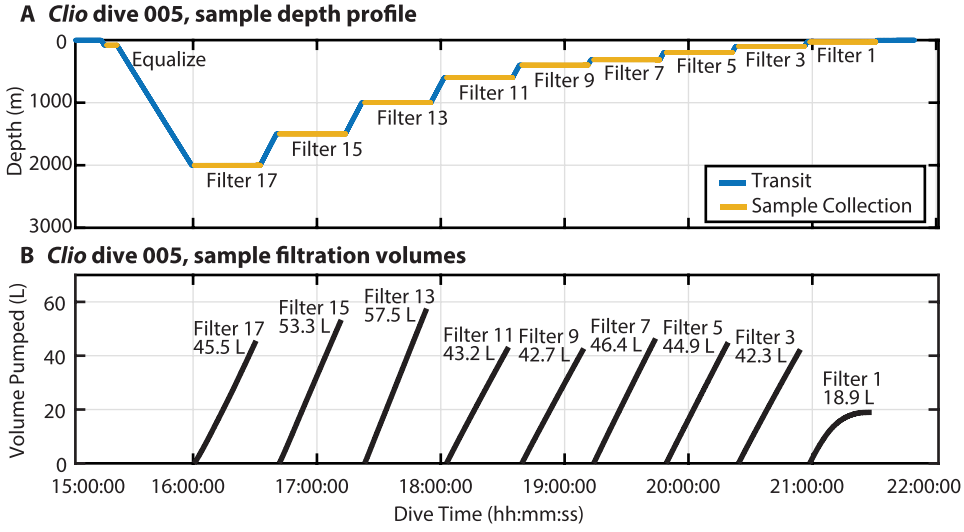

Fig. 4. Typical *Clio* dive. (**A**) *Clio* dives involve holding a shallow depth while air is purged, driving down to the deepest sampling depth, and holding at each sampling depth during the return to the surface. (**B**) At each sample depth, seawater is filtered for a predetermined maximum amount of time. Samples collected in the surface ocean (e.g., filter 1) collect biomass more rapidly because of the higher particulate concentration, which is reflected in the nonlinearity of the sample flow curve; these samples require less filtration. During this dive, *Clio* was deployed with one payload filled with nine filter units, but more often, two payload bays of filter units are deployed for parallel filtration or increased depth resolution; see table S1. Parallel filtration with independent payload bays facilitates the collection of samples for different types of measurements. In this case, the filter numbers are odd numbers between 1 and 17 and reflect the port position of the valve.

prevailing vertical currents. This is intentional; Lagrangian motion results in better samples that are more representative of the water mass being studied. As a result of Lagrangian transport, the AUV typically surfaced 2 to 15 km from the launch site; in the most extreme case, during a profile at the edge of the Gulf Stream Current, the vehicle surfaced 34 km from the launch site. The supporting research vessel typically transits at 15 km hour^{-1} and can typically reach the AUV in <1 hour; recovering the AUV back on to deck was completed in <1 hour depending on sea state. This operational tempo is in keeping with our design objectives. Meeting these objectives required careful attention to understanding and optimizing the vehicle's ballasting and power demands.

Energy expended during dives consists of (i) baseline power consumption including profiling sensors and sampling pumps and (ii) propulsion power spent to overcome buoyancy while sampling and to overcome drag while transiting. As with all vehicles, this represents a trade-off, and every vehicle has an optimal speed at which range is maximized (*60*). Slow transits reduce the energy needed to overcome drag but increase the energy needed to sustain baseload and overcome buoyancy. The energy consumed overcoming buoyancy during the descent comes close to, but does not entirely cancel, the energy negated by buoyancy on the ascent, because of asymmetry in the propellers. Fast transits reduce the energy needed to sustain baseload and overcome buoyancy but increase the energy needed to overcome drag. *Clio's* mission calls for relatively rapid transits where the power to overcome drag losses dominates. For most missions, the vehicle expends most of its battery energy while holding depth and sampling. Vehicle buoyancy represents a purely parasitic load while sampling. Therefore, the accurate prediction of propulsion power while holding depth—which varies as a function of pressure, temperature, and in situ seawater density—is necessary to maximize endurance while ensuring vehicle safety. The 26 dives to date have been used to test our understanding of the vehicle's drag, buoyancy, and coefficients of thermal expansion and bulk modulus; in these dives, thrust, power, and ballasting conditions were measured using the approach described in the "Engineering materials and methods" section in the Supplementary Materials.

Transit propulsion power depends on buoyancy and speed. Figure 5 illustrates the propulsion power consumed while transiting as a function of transit speed for various ballasts. Conservative ballasting (vehicle very buoyant) increases descent power but decreases ascent power. Descent power typically exceeds ascent power, a consequence of the additional power required to overcome buoyancy when descending, but at sufficiently high speed, the opposite holds, a consequence of asymmetry in the propellers designed for high efficiency while thrusting down to hold depth. Below the critical speed, power approximately scales as the velocity cubed. Above the critical speed of about 0.75 m s^{-1}, the power required increases markedly; however, the precise speed at which this transition occurs varies, perhaps as a result of trim. The critical speed is also somewhat different for ascent versus descent because of the relative positions of the center of mass and aerodynamic center change. We have observed stable transit at speeds as high as 0.9 m s^{-1}. The round-trip transit energy consumption, including baseload expressed as the percentage of battery capacity consumed per kilometer descent and ascent, varies between 3 and 5% near the critical speed across the practical range of ballasts. For an aggressively ballasted vehicle, this figure holds throughout the water column despite the increase in buoyancy with depth.

Clio's buoyancy changes with depth as the density of seawater, and all of the materials comprising the vehicle respond to differences in temperature and pressure. *Clio's* deepest dive, to 4100 m, allowed us to test our understanding of the vehicle's net coefficient of thermal expansion and net bulk modulus. Analysis of calibrated thruster data from *Clio*'s dive to 4100 m indicates a best-fit coefficient of thermal expansion of 3.65×10^{-5} m m^{-1} K^{-1} and a net bulk modulus of 3.3 GPa. This is 1.5 times less compressible than typical seawater (2.3 GPa) and favorable compared with most deep-diving AUVs that are typically two to five times less compressible than seawater (*61*). For a vehicle of *Clio's* size, the typical range corresponds to 90 to 150 N in added buoyancy at 6000 m, i.e., a notable increase in parasitic propulsion load. The value for *Clio* is 50 N.

Figure 5 illustrates the power consumed while holding depth to collect samples. Background vehicle processes and sensors consume 1.2% of battery capacity hour^{-1}. Each additional pump used while sampling consumes another 0.4% of battery capacity hour^{-1}. The energy consumed by the thrusters varies with depth and depends on the surface ballast condition. For conservative ballasting, i.e., 100 N

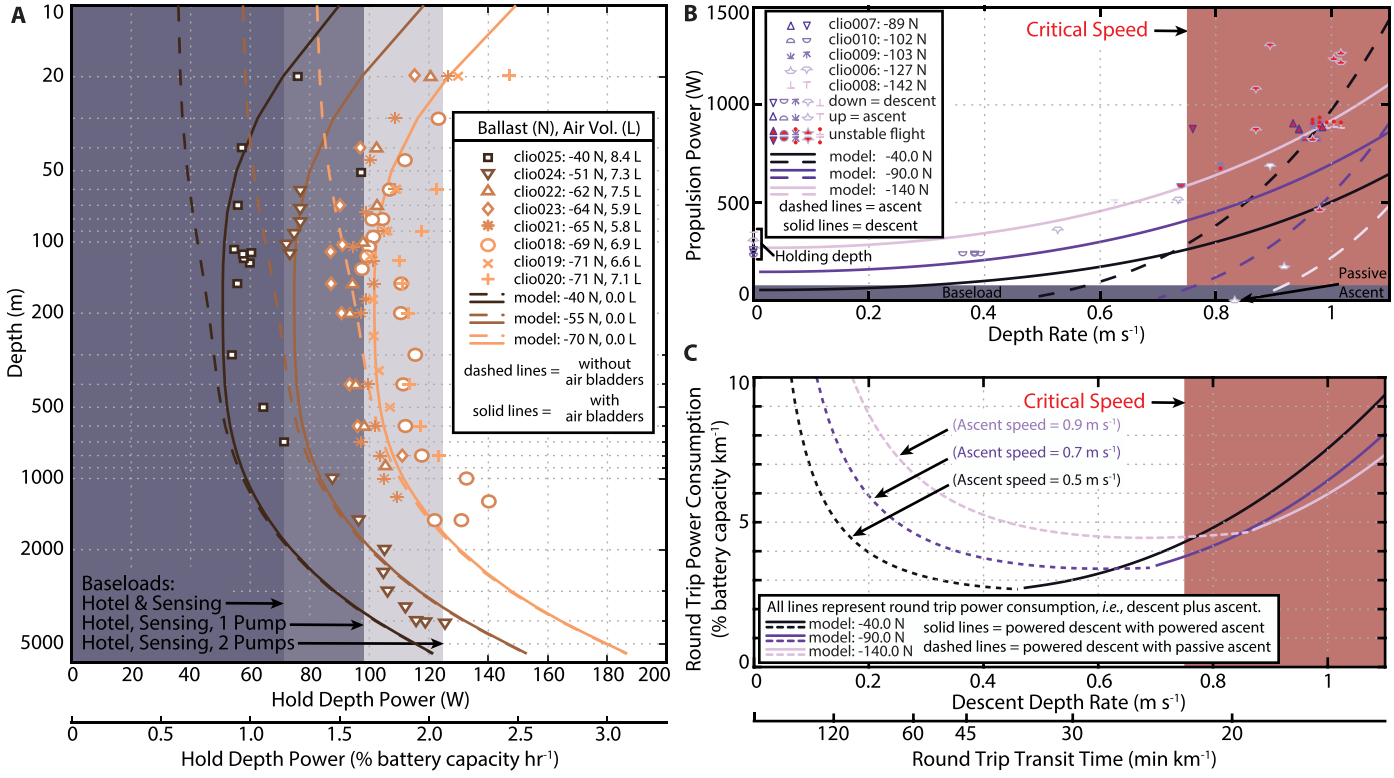

Fig. 5. Power consumption. Power consumption, while holding depth (**A**) consists of baseload, power to run pumps, and propulsion power to overcome vehicle buoyancy (including air bladders). Vehicle buoyancy varies with depth. Solid lines are predictions for propulsion power consumption, assuming a global mean temperature and density profile (*56*), whereas markers represent data for in situ profiles measured by *Clio*. Power consumption while transiting (**B**) depends on speed and vehicle buoyancy, plus baseload. Above the critical speed at which *Clio* transitions to unstable flight, effective drag rises markedly, resulting in a concomitant increase in propulsion power. The model predictions apply for stable flight only (sometimes observed above the critical speed). (**C**) Propulsion energy consumed to execute a midwater round trip for a 1-km profile. The dashed portions of the model curves indicate power that would be expended during passive ascent to slow the vehicle and reflect energy consumed by baseload only during a passive ascent at the terminal speed indicated. The model curves in (A) and (C) allow accurate prediction of total energy consumption for mission planning.

buoyant or 1.4% of vehicle displacement, energy consumption is 1.5 to 3% of battery capacity hour^{-1}. For a more aggressive ballasting state, i.e., 50 N buoyant or 0.7% of vehicle displacement, energy consumption is reduced to 1 to 2% of battery capacity hour^{-1}. Propulsion energy consumption is lowest at intermediate depths of 100 to 1000 m where the air bladder volume has been compressed and vehicle buoyancy has yet to be strongly influenced by pressure or increasing background density. Figure 5 shows that with aggressive ballasting, the energy expended by propulsion while holding depth is commensurate with the energy required for sampling throughout *Clio's* 6000-m depth range, thus validating a key assumption during the design process that chose thrusters as the sole means of depth control.

Baseline cleanliness and genomic resolution

Baseline cleanliness was assessed using genomic analyses of samples returned from a *Clio* profile. Complementary DNA (cDNA) sequences were scanned for putative contaminant taxa identified by Salter *et al.* (*62*) and the interpretation of Sheik *et al.* (*63*) for low-biomass sequenced samples from the Census of Deep Life. Nodes that match in taxonomy to these were flagged as putative contaminants (fig. S2 and table S2). In this profile, we assessed background cDNA sequence contamination to be between 0.2 and 4.5%, and 1% on average.

Genomic analyses also show the microbial community structure and genomic resolution achieved by *Clio* (Fig. 6 and table S2). Figure 6 shows the results from nine samples collected by *Clio* from depths of 30 to 2000 m and analyzed for microbial DNA and RNA measured as cDNA ("Biochemical methods" section in the Supplementary Materials). Up to 38,280 DNA sequences and 24,493 cDNA sequences were obtained from filtered volumes of 18 to 57 liters, using a one-sixteenth split portion of the filtered biomass. In two cases, insufficient biomass or extraction methods, or both, resulted in low cDNA yields (table S2, samples D5 and D17); in all other cases, filter biomass was sufficient to generate sequencing libraries with good resolution that were representative of the water column community. These results demonstrate core functionality and baseline necessary volumes and illustrate, to some degree, the amount of data returned from these biochemical measurements and contained within just one profile.

The genomic data in Fig. 6 (A and B) are categorized at the phylum level into 27 groups, but this is just a fraction of the data. A deeper characterization is possible within each of these levels. Figure 6 (C and D) shows the classification of the Proteobacteria phylum into a further 37 clades. Moreover, the cDNA is itself representative of only a fraction of the RNA transcriptome that can be measured for these groups.

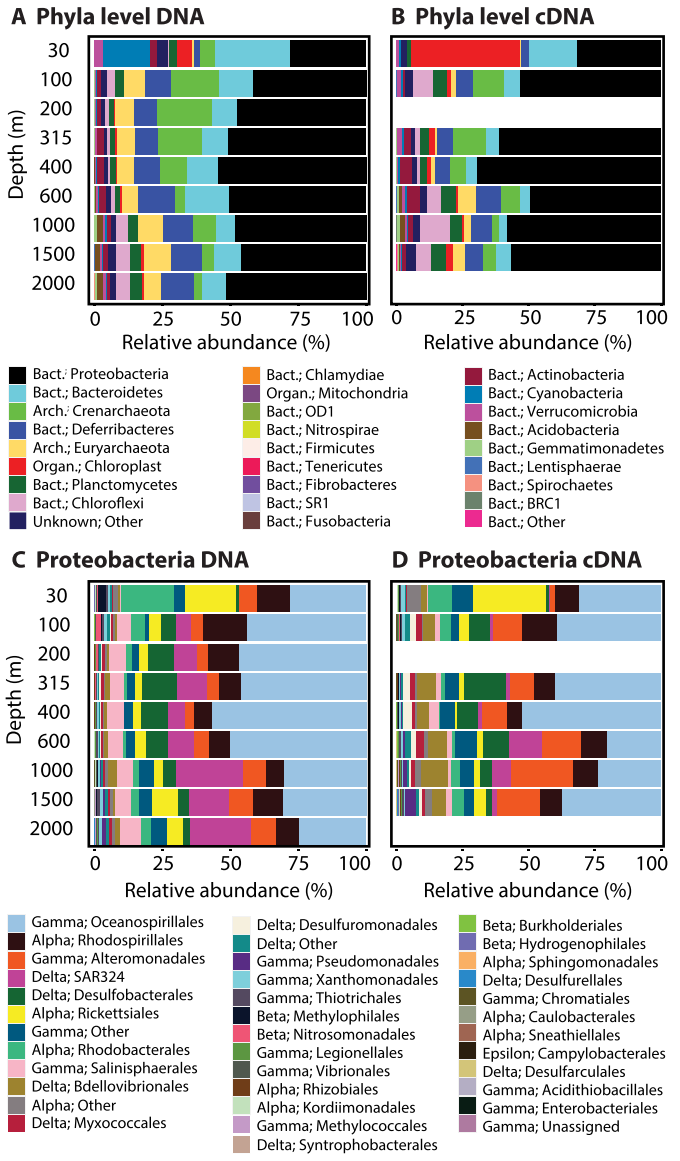

Fig. 6. *Clio* DNA and cDNA samples. Relative abundances of phyla in *Clio* (**A**) DNA and (**B**) cDNA samples and relative abundances specifically within Proteobacteria (**C**) DNA and (**D**) cDNA. Samples are organized by depth on the *y* axis and relative abundance on the *x* axis. Two cDNA samples, AR20-5-D5 at 200 m and AR20-5-D17 at 2000 m, do not show on the graph because of either or both insufficient biomass and extraction methods resulting in low cDNA yields.

Nor does this include the proteomic and metabolomic datasets that can be acquired from the same samples.

High-resolution proteomic measurements

The proteomic analysis of samples returned from a *Clio* profile collected at BATS station shows the biochemical measurements that can be made by this method (Figs. 7 and 8 and table S3). Proteomic analysis can only be performed on returned samples, as is true of metabolomic analysis as well. Figure 7 (A and B) shows the results from 10 samples of microbial biomass collected by *Clio* on 14 April 2018, at BATS station, and analyzed using metaproteomic methods ("Biochemical methods" section in the Supplementary Materials) (tables S3 and S4). From these samples, 1577 proteins were identified using high-resolution liquid chromatography–mass spectrometry from filtered volumes of 56 to 135 liters per depth (similar to Fig. 4), using one 50% split portion of the filtered biomass (*15*, *64*). The vertical distributions of these proteins follow the niche depth range for the various microbial groups, as illustrated by a subset of proteins taxonomically attributed to *Prochlorococcus, Synechococcus*, and *Pelagibacter* (Fig. 7B).

Clio can also conduct higher-resolution sampling than existing wire-based techniques. To demonstrate this, on 24 October 2018, at BATS station, *Clio* collected five samples across a chlorophyll maximum at 5-m increments while holding depth to a precision of <5 cm (Fig. 7, C and D, and tables S5 to S8). In comparison, wire-deployed in situ pumps, which were collecting samples synoptically, varied in depth by more than 15 m because of high winds and waves. Metaproteomic analysis of these targeted, high-resolution *Clio* samples revealed distinct depth differences within the 5-m sampling increments, with numerous proteins at the top (Fig. 7E) of the chlorophyll feature being significantly more abundant than those at the center and vice versa, based on Fisher's exact test ($P < 0.05$) (Fig. 7F). These results are consistent with prior observations, including for several *Prochlorococcus* proteins: Urea and phosphate transporters were more abundant above the maximum where nitrogen is scarcest (Fig. 7, E and G), and the iron-free electron transport protein flavodoxin was most abundant at the chlorophyll maximum where iron scarcity is known to occur (Fig. 7, F and G). Nutrient stress biomarkers such as these are valued for their ability to diagnose the environmental stresses faced by each major species present and how their biochemical responses change in time and space (*15*). Also, different isoforms of key biochemical proteins, such as glyceraldehyde-3-phosphate dehydrogenase, were observed with depth corresponding to high- and low-light *Prochlorococcus* ecotypes. These results demonstrate how *Clio*'s unique depth keeping capabilities can lead to new observations that are otherwise very difficult to achieve.

Sectional transect

During June 2019, in a series of nine dives spanning the 1144-km transect between Bermuda and Woods Hole, Massachusetts, USA, we used *Clio* to collect (i) 158 sample sets for proteins, metals, and nutrients (table S9) and (ii) complementary sensor data (table S10). *Clio*'s optical beam transmission data indicate how low suspended material concentration was in this region (fig. S3). Consequently, *Clio* had to filter 20,878 liters of seawater to collect these samples. The biochemical data from this cruise are being analyzed and will be reported elsewhere. Figure 8A shows the level 3, monthly averaged, chlorophyll a estimates from the MODIS-Aqua spectroradiometer for June 2019 (see the "Remote-sensing methods" section in the Supplementary Materials), and the *Clio* transect stations. Figure 8 also shows the results for chlorophyll a, measured by *Clio*'s fluorometer, and dissolved inorganic nitrogen measured from samples collected by *Clio*. A comparison of the remotely sensed MODIS-Aqua chlorophyll a to the in situ *Clio* observations shows the presence of a deep chlorophyll maximum (DCM) layer at depths between 80 and 150 m throughout this region (Fig. 8B). This feature indicates a region of active primary productivity that cannot be observed by remote-sensing techniques and must be observed within the water column.

During this transect, we also implemented an adaptive sampling algorithm to better target discrete features like the DCM. During

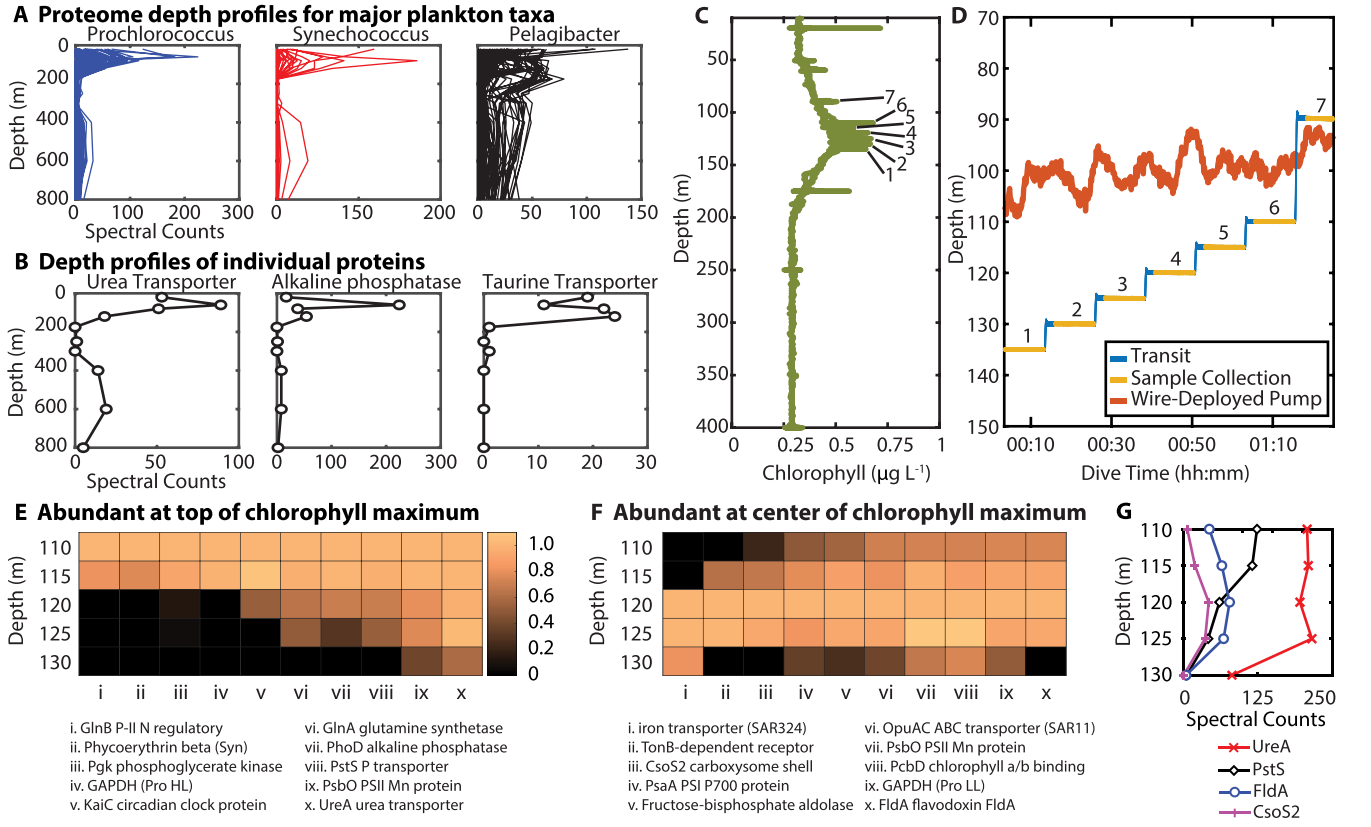

Fig. 7. *Clio* high-resolution protein sampling. (**A**) Protein profiles categorized by major microbial taxa and (**B**) selected nutrient response proteins from *Prochlorococcus* (urea transporter and alkaline phosphatase) and *Pelagibacter* (taurine transporter) from dives Clio-007/008 from BATS station. (**C** to **G**) Higher-resolution sampling of the DCM on dive Clio-014. (C) *Clio*'s chlorophyll profile; sample locations numbered. (D) *Clio* depth versus time, with the same locations labeled, shows holding station to <1 m, whereas depth versus time for a concurrent wire-deployed pump shows the limitation that ship heave imposes on depth precision for cabled instruments. Significant differences (Fisher's exact test $P < 0.05$) were observed in (E) 10 selected proteins more abundant at the top (110 m) or (F) center (120 m) of the chlorophyll maximum (ratio of total exclusive spectral counts, normalized and log transformed; see tables S5 to S8). All proteins are *Prochlorococcus* unless labeled otherwise (Syn is *Synechococcus*, Pro is *Prochlorococcus*; HL and LL refer to high- and low-light ecotypes). (**G**) Protein profiles of phosphate transporter PstS, urea transporter UreA, flavodoxin FldA (iron-free substitute for ferredoxin), and CsoS2 carboxysome shell protein demonstrate observation of biologically meaningful differences in P, N, Fe, and C metabolisms with 5-m resolution.

Clio dives 024 and 025, the depths of a subset of stations were left unspecified with instructions to sample the DCM. *Clio* autonomously determined the depth of the DCM on the basis of its chlorophyll fluorometer during descent. After completing several preprogrammed stations at deeper depths, *Clio* ascended to distribute the remaining stations at depths across the DCM. This adaptive sampling approach is similar to that taken by Zhang *et al.* (*65*) to sample thin phytoplankton layers using a conventional AUV for lateral surveys. This capability, coupled with high-resolution sampling, will enable even finer sampling resolution of discrete features like the DCM in the future.

The sectional cruise was completed in 14 days. If the same set of samples were collected without *Clio*, we estimate that the cruise would have taken twice as long. Moreover, this transect is a relatively short distance for ocean sectional studies; the time savings would typically be greater. During this section, *Clio* also completed a profile to a depth of 4100 m. These data, and *Clio*'s reliable performance, demonstrate that this robot can complete the sectional cruises that we designed it to perform.

Biochemical significance of trials

During the BATS and Sargasso Sea deployments, we used *Clio* to better understand the vertical structure of plankton communities and their relationship to available nutrients. The thousands of functional proteins observed represent a tremendous amount of information, but the characteristics of the ecosystem can be discerned from a few example proteins. Key proteins increase in response to extreme nutrient scarcity for nitrogen, phosphorus, and iron; for example, vertical distributions of a urea transporter, phosphate transporter, and flavodoxin from the cyanobacterium *Prochlorococcus* reflect adaptations to scarce N, P, and Fe (Fig. 7G). Similarly, proteins can be indicative of key metabolic processes such as photosynthetic carbon fixation (PsbO and CsoS2; Fig. 7, F and G) from *Prochlorococcus* and the use of the organic molecule taurine as a substrate for respiration in the bacterium *Pelagibacter* (Fig. 7B). Together, these signals can provide a powerful assessment of the metabolic capabilities of each major microbial taxon present in the community, indicating what is controlling their growth and their contribution to the marine carbon cycle at each point in space and time.

DISCUSSION

These results demonstrate that *Clio* can be used to observe ocean biochemistry across large sections of the ocean. We showed that the

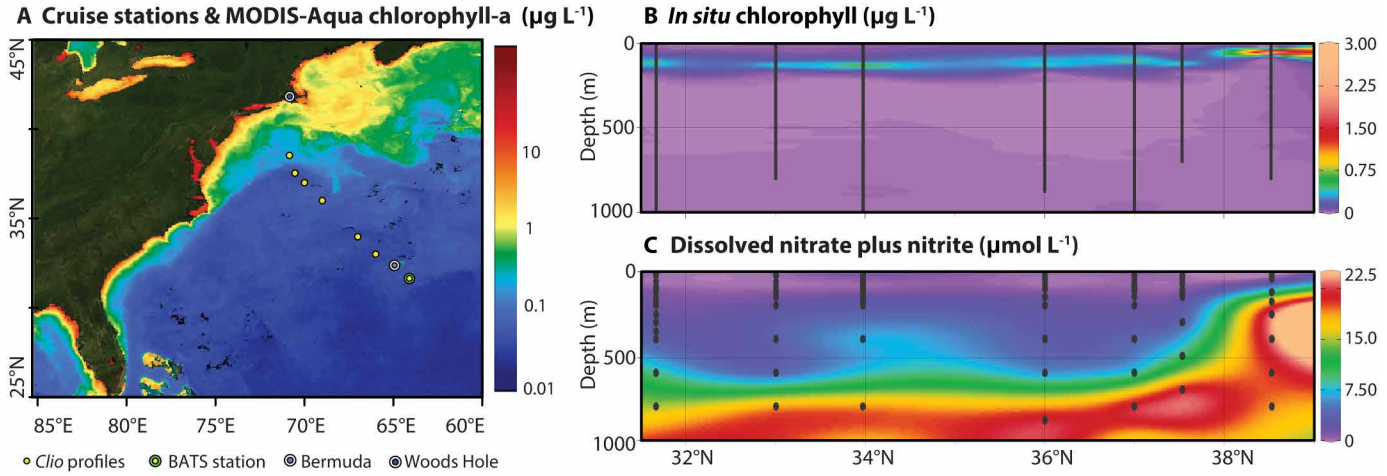

Fig. 8. Ocean biogeochemical structure revealed. (**A**) *Clio* sectional cruise stations superimposed over the level 3, monthly averaged chlorophyll a estimates from the MODIS-Aqua spectroradiometer for June 2019. (**B**) In situ chlorophyll a measured by *Clio*'s fluorometer: Gray lines are overlapping dots representing point data that are spatially interpolated, and (**C**) dissolved nitrate plus nitrite measured from samples collected by *Clio*.

vehicle can efficiently transit the water column while stopping to loiter precisely at depths as needed. We showed how the vehicle, its sample-processing systems, and the methods for using them can collect samples preserving genomic information with low background contamination levels. We also showed that *Clio* returns samples for metaproteomic analysis that can only be realized from the processing of large volumes of water, much larger than can be collected by other current AUVs. We demonstrated how *Clio* maps chemical and physical properties across sections of the ocean using (i) sensor data, e.g., chlorophyll, salinity, temperature, and optical backscatter, and (ii) data produced from sample analysis. We showed, as an example, how the sectional map of chlorophyll a, measured in situ by *Clio's* sensor, reveals the DCM environment hidden from remote satellite sensors. We also demonstrated, using nitrate plus nitrite data, how the analysis of returned samples collected by this method can reveal the distribution of biochemically relevant ocean parameters that are difficult to measure with existing sensors. The genomic and proteomic measurements illustrate the richness of these datasets; however, the data that we showed are only a limited fraction of what can be measured in these analyses. Furthermore, the data that we showed represent only a partial set of the important biochemical compounds that could be measured with this method. Therefore, the datasets that *Clio* enables are in a class of "big data" that is challenging to represent. In the simplest case, if a distinct sectional map were created for every biochemical compound so measured, the result would be many 10,000 s of mapping products. The oceanographic community is now only at the beginning of the process of studying this wide range of compounds and understanding their roles in the cycling of nutrients between the environment and living systems. Our goal is that *Clio*, and robots like it, will contribute to a global-scale biochemical mapping of the ocean (*66*), allowing researchers to answer fundamental questions about the limiting constraints on ocean life and the role of ocean life in regulating Earth's chemical environment. Moreover, our goal is that *Clio* will accelerate this research so that we can be better informed about how Earth's ocean and climate are changing and what we can do to mitigate the negative consequences of these changes.

Limitations and future steps

A limitation of *Clio* is that it is not suited to under-ice applications. However, *Clio*'s sample return systems can be used on other ocean vehicles, such as the hybrid ROV *NUI* (*67*), suited to under-ice operations. Also, we intentionally designed this version of *Clio* with a maximum depth of 6000 m because most of the ocean is shallower than this. However, a deeper diving hadal version of *Clio* is possible. Notably, *Clio* still has substantial unused scientific payload capacity and underused computational, data storage, and battery capacity. Thus, there is strong potential to add new sensors and in situ analysis systems, develop new autonomous behaviors that can better navigate the water column biogeochemical environment, and even develop the capacity to carry out experimental treatments on marine microorganisms, or seawater chemistry, under ambient conditions.

MATERIALS AND METHODS

Vehicle design

Clio is similar in size to water sampling rosettes (Fig. 1F), which are common to oceanographic research vessels, and can be deployed, recovered, and shipped with similar equipment. On deck, the vehicle stands vertically on integral landing skids and can be launched and recovered from this position. The vehicle has an aluminum structural chassis that supports all components including the (i) two vertical thrusters, (ii) three pressure housings for electronics and batteries below, (ii) syntactic foam flotation above, (iv) four sample return payload bays around the perimeter of the midsection, and (v) sensors near the bottom (Fig. 1). The vehicle's free-flooded internal volume, including the sample return payload bays, is enclosed in high-density polyethylene shells. These shells are removable, which provides operators easy access to the sample return systems and the power and communication cables (Fig. 2B). At the top of the vehicle is an acoustic modem that allows operators to acoustically range on the vehicle and receive regular vehicle status updates. Also, on top are three independently powered depth-activated beacons, a GPS/Iridium unit to provide an over-the-horizon location, a strobe for line-of-sight location, and a VHF (very high frequency)

radio beacon, which alert operators to the vehicle's position when it surfaces. The computer systems, batteries, and thrusters leverage designs are inherited from previous deep-sea AUVs: *ABE* (*68*), *Sentry* (*69*), *Nereus* (*70*), and *NUI* (*67*). *Clio* carries a Paroscientific pressure sensor, a Seabird 49 Fastcat CTD sensor, and typically also a WET Labs 25-cm pathlength transmissometer, a WET Labs ECO FL chlorophyll a fluorometer, and an Aanderaa dissolved oxygen optode. Components are pressure rated to 6000 m or greater.

Sampling system design

The *Clio* sampling systems inherit design features and principles from previous deep-sea ROV and AUV sampling systems (*43*, *44*, *53*), but their overall design, and many features, have been uniquely developed for this open-ocean application. Specifically, they use custom high-flow rate valves to speed sample collection and minimize power consumption while pumping; the diameter of the fluidic path within the valves and the rest of the fluidic system is never smaller than 9.5 mm. Each valve has 19 ports and is actuated by a direct-drive NEMA 42 stepper motor with a bipolar torque of 40 N·m. The stepper motors are equipped with absolute encoders to ensure that their position is known at all times. The sample containers are first followed by the control valve, then a deep-sea positive displacement pump (8 liters min^{-1} pump head, Mclane Research Laboratories) that pulls seawater through the system, and last a digital flow meter (SPX model, Seametrics) that records sample flow rate and volume. The flow rate through the system depends on the filter media used and amount of suspended particulate material in the sample, but typically, the average sampling flow rate achieved is 1 to 2 liters min^{-1}. The control and motor power electronics for the samplers are pressure-tolerant designs located within the fluorocarbon-filled stepper motor housings. The custom sample units that we developed to maximize sample capacity and cleanliness are described in the "Engineering materials and methods" section in the Supplementary Materials.

SUPPLEMENTARY MATERIALS

robotics.sciencemag.org/cgi/content/full/5/48/eabc7104/DC1
Supplementary Text: Background, materials, and methods.
Fig. S1. *Clio* robotic approach to ocean biochemical profiling.
Fig. S2. Abundances of putative contaminant sequences in *Clio* DNA and cDNA samples.
Fig. S3. Ocean environmental structure revealed using *Clio*.
Table S1. *Clio* sampling configuration used in Sargasso Sea study.
Table S2. North Atlantic shelf break DNA and RNA.
Table S3. BATS station proteins.
Table S4. Spectral counts and annotations of proteins from *Clio* dive 007 and 008 BATS station. (Microsoft Excel format).
Table S5. High-resolution chlorophyll maximum sampling profile.
Table S6. Total exclusive spectral counts of selected proteins abundant at the top of the chlorophyll maximum from Clio-014 high-resolution sampling from 110 to 130 m.
Table S7. Total exclusive spectral counts of selected proteins abundant at the center of the chlorophyll maximum from Clio-014 high-resolution sampling from 110 to 130 m.
Table S8. Spectral counts and annotations of 595 identified proteins from Clio-014 high-resolution sampling 110 to 130 m at BATS station.
Table S9. Sargasso Sea dissolved inorganic nitrogen.
Table S10. Temperature, salinity, chlorophyll a, turbidity, and optical beam transmission.
References (*71*–*95*)

REFERENCES AND NOTES

1. K. R. Arrigo, Marine microorganisms and global nutrient cycles. *Nature* **437**, 349–355 (2005).
2. M. J. Follows, S. Dutkiewicz, S. Grant, S. W. Chisholm, Emergent biogeography of microbial communities in a model ocean. *Science* **315**, 1843–1846 (2007).
3. G. J. Herndl, T. Reinthaler, Microbial control of the dark end of the biological pump. *Nat. Geosci.* **6**, 718–724 (2013).
4. D. A. Bryant, The beauty in small things revealed. *Proc. Natl. Acad. Sci. U.S.A.* **100**, 9647–9649 (2003).
5. A. C. Redfield, The biological control of chemical factors in the environment. *Am. Sci.* **46**, 205–221 (1958).
6. J. B. Waterbury, S. W. Watson, R. R. Guillard, L. E. Brand, Widespread occurrence of a unicellular, marine, planktonic, cyanobacterium. *Nature* **277**, 293–294 (1979).
7. S. J. Giovannoni, M. S. Rappé, K. L. Vergin, N. L. Adair, 16S rRNA genes reveal stratified open ocean bacterioplankton populations related to the green non-sulfur bacteria. *Proc. Natl. Acad. Sci. U.S.A.* **93**, 7979–7984 (1996).
8. A. K. Hawley, H. M. Brewer, A. D. Norbeck, L. Paša-Tolić, S. J. Hallam, Metaproteomics reveals differential modes of metabolic coupling among ubiquitous oxygen minimum zone microbes. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 11395–11400 (2014).
9. M. V. Brown, M. Ostrowski, J. J. Grzymski, F. M. Lauro, A trait based perspective on the biogeography of common and abundant marine bacterioplankton clades. *Mar. Genomics* **15**, 17–28 (2014).
10. Z. I. Johnson, E. R. Zinser, A. Coe, N. P. McNulty, E. M. S. Woodward, S. W. Chisholm, Niche partitioning among prochlorococcus ecotypes along ocean-scale environmental gradients. *Science* **311**, 1737–1740 (2006).
11. G. Rocap *et al.*, Genome divergence in two Prochlorococcus ecotypes reflects oceanic niche differentiation. *Nature* **424**, 1042–1047 (2003).
12. S. Ganesh, D. J. Parris, E. F. DeLong, F. J. Stewart, Metagenomic analysis of size-fractionated picoplankton in a marine oxygen minimum zone. *ISME J.* **8**, 187–211 (2014).
13. N. Kashtan, S. E. Roggensack, J. W. Berta-Thompson, M. Grinberg, R. Stepanauskas, S. W. Chisholm, Fundamental differences in diversity and genomic population structure between Atlantic and Pacific *Prochlorococcus*. *ISME J.* **11**, 1997–2011 (2017).
14. K. Anantharaman, J. A. Breier, C. S. Sheik, G. J. Dick, Evidence for hydrogen oxidation and metabolic plasticity in widespread deep-sea sulfur-oxidizing bacteria. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 330–335 (2013).
15. M. A. Saito, A. Dorsk, A. F. Post, M. R. McIlvin, M. S. Rappé, G. R. DiTullio, D. M. Moran, Needles in the blue sea: Sub–species specificity in targeted protein biomarker analyses within the vast oceanic microbial metaproteome. *Proteomics* **15**, 3521–3531 (2015).
16. C. L. Fiore, K. Longnecker, M. C. Kido Soule, E. B. Kujawinski, Release of ecologically relevant metabolites by the cyanobacterium *Synechococcus elongatus* CCMP 1631. *Environ. Microbiol.* **17**, 3949–3963 (2015).
17. V. Marx, When microbiologists plunge into the ocean. *Nat. Methods* , 133–136 (2020).
18. S. C. Doney, The growing human footprint on coastal and open-ocean biogeochemistry. *Science* **328**, 1512–1516 (2010).
19. A. C. Redfield, On the proportions of organic derivatives in sea water and their relation to the composition of plankton, in *James Johnstone Memorial Volume*, R. J. Daniel, Ed. (Liverpool Univ. Press, 1934), pp. 176–192.
20. W. S. Broecker, Thermohaline circulation, the Achilles heel of our climate system: Will man-made CO_2 upset the current balance? *Science* **278**, 1582–1588 (1997).
21. M. J. McPhaden, A. J. Busalacchi, R. Cheney, J.-R. Donguy, K. S. Gage, D. Halpern, M. Ji, P. Julian, G. Meyers, G. T. Mitchum, P. P. Niiler, J. Picaut, R. W. Reynolds, N. Smith, K. Takeuchi, The Tropical Ocean-Global Atmosphere observing system: A decade of progress. *J. Geophys. Res. Oceans.* **103**, 14169–14240 (1998).
22. A. Ganachaud, C. Wunsch, Improved estimates of global ocean circulation, heat transport and mixing from hydrographic data. *Nature* **408**, 453–457 (2000).
23. M. J. R. Fasham, B. M. Baliño, M. C. Bowles, A new vision of ocean biogeochemistry after a decade of the Joint Global Ocean Flux Study (JGOFS). *Ambio* **30**, 4–30 (2001).
24. The GEOTRACES Group, The GEOTRACES Intermediate Data Product 2014. *Mar. Chem.* **177**, 1–8 (2015).
25. D. B. Rusch, A. L. Halpern, G. Sutton, K. B. Heidelberg, S. Williamson, S. Yooseph, D. Wu, J. A. Eisen, J. M. Hoffman, K. Remington, K. Beeson, B. Tran, H. Smith, H. Baden-Tillson, C. Stewart, J. Thorpe, J. Freeman, C. Andrews-Pfannkoch, J. E. Venter, K. Li, S. Kravitz, J. F. Heidelberg, T. Utterback, Y.-H. Rogers, L. I. Falcón, V. Souza, G. Bonilla-Rosso, L. E. Eguiarte, D. M. Karl, S. Sathyendranath, T. Platt, E. Bermingham, V. Gallardo, G. Tamayo-Castillo, M. R. Ferrari, R. L. Strausberg, K. Nealson, R. Friedman, M. Frazier, J. C. Venter, The Sorcerer II global ocean sampling expedition: Northwest Atlantic through eastern tropical Pacific. *PLOS Biol.* **5**, e77 (2007).
26. P. Bork, C. Bowler, C. de Vargas, G. Gorsky, E. Karsenti, P. Wincker, *Tara* Oceans studies plankton at planetary scale. *Science* **348**, 873 (2015).
27. A. E. Santoro, Crystal ball: The microbial map of the ocean. *Environ. Microbiol. Rep.* **11**, 35–37 (2019).
28. A. Ganachaud, C. Wunsch, Oceanic nutrient and oxygen transports and bounds on export production during the World Ocean Circulation Experiment. *Global Biogeochem. Cycles* **16**, 1057 (2002).
29. M. A. Spall, Circulation in the Canary Basin: A model/data analysis. *J. Geophys. Res. Oceans* **95**, 9611–9628 (1990).

30. L. Wu, Z. Jing, S. Riser, M. Visbeck, Seasonal and spatial variations of Southern Ocean diapycnal mixing from Argo profiling floats. *Nat. Geosci.* **4**, 363–366 (2011).
31. Y. Ohno, T. Kobayashi, N. Iwasaka, T. Suga, The mixed layer depth in the North Pacific as detected by the Argo floats. *Geophys. Res. Lett.* **31**, L11306 (2004).
32. N. L. Williams, L. W. Juranek, R. A. Feely, K. S. Johnson, J. L. Sarmiento, L. D. Talley, A. G. Dickson, A. R. Gray, R. Wanninkhof, J. L. Russell, S. C. Riser, Y. Takeshita, Calculating surface ocean pCO_2 from biogeochemical Argo floats equipped with pH: An uncertainty analysis. *Global Biogeochem. Cycles* **31**, 591–604 (2017).
33. J. K. B. Bishop, T. J. Wood, Year-round observations of carbon biomass and flux variability in the Southern Ocean. *Global Biogeochem. Cycles* **23**, GB2019 (2009).
34. M. J. Perry, B. S. Sackmann, C. C. Eriksen, C. M. Lee, Seaglider observations of blooms and subsurface chlorophyll maxima off the Washington coast. *Limnol. Oceanogr.* **53**, 2169–2179 (2008).
35. D. C. Webb, P. J. Simonetti, C. P. Jones, SLOCUM: An underwater glider propelled by environmental energy. *IEEE J. Ocean. Eng.* **26**, 447–452 (2001).
36. S. Glenn, C. Jones, M. Twardowski, L. Bowers, J. Kerfoot, J. Kohut, D. Webb, O. Schofield, Glider observations of sediment resuspension in a Middle Atlantic Bight fall transition storm. *Limnol. Oceanogr.* **53**, 2180–2196 (2008).
37. D. Yoerger, J. Newman, J.-J. Slotine, Supervisory control system for the JASON ROV. *IEEE J. Ocean. Eng.* **11**, 392–400 (1986).
38. D. M. Karl, A. M. Brittain, B. D. Tilbrook, Hydrothermal and microbial processes at Loihi Seamount, a mid-plate hot-spot volcano. *Deep Sea Res. Part A Oceanogr. Res. Pap.* **36**, 1655–1673 (1989).
39. B. I. Bergström, J. Larsson, J.-O. Pettersson, Use of a remotely operated vehicle (ROV) to study marine phenomena: I. Pandalid shrimp densities. *Mar. Ecol. Prog. Ser.* **37**, 97–101 (1987).
40. J. P. Cowen, D. A. Copson, J. Jolly, C.-C. Hsieh, H.-T. Lin, B. T. Glazer, C. G. Wheat, Advanced instrument system for real-time and time-series microbial geochemical sampling of the deep (basaltic) crustal biosphere. *Deep Sea Res. Part I Oceanogr. Res. Pap.* **61**, 43–56 (2012).
41. C. D. Taylor, K. W. Doherty, S. J. Molyneaux, A. T. Morrison III, J. D. Billings, I. B. Engstrom, D. W. Pfitsch, S. Honjo, Autonomous Microbial Sampler (AMS), a device for the uncontaminated collection of multiple microbial samples from submarine hydrothermal vents and other aquatic environments. *Deep Sea Res. Part I Oceanogr. Res. Pap.* **53**, 894–916 (2006).
42. D. A. Butterfield, K. K. Roe, M. D. Lilley, J. A. Huber, J. A. Baross, R. W. Embley, G. J. Massoth, Mixing, reaction and microbial activity in the sub-seafloor revealed by temporal and spatial variation in diffuse flow vents at Axial Volcano. *Geophys. Monogr. Ser.* **144**, 269–289 (2004).
43. J. A. Breier, C. G. Rauch, K. McCartney, B. M. Toner, S. C. Fakra, S. N. White, C. R. German, A suspended-particle rosette multi-sampler for discrete biogeochemical sampling in low-particle-density waters. *Deep Sea Res. Part I Oceanogr. Pap.* **56**, 1579–1589 (2009).
44. J. A. Breier, C. S. Sheik, D. Gomez-Ibanez, R. T. Sayre-McCord, R. Sanger, C. Rauch, M. Coleman, S. A. Bennett, B. R. Cron, M. Li, C. R. German, B. M. Toner, G. J. Dick, A large volume particulate and water multi-sampler with *in situ* preservation for microbial and biogeochemical studies. *Deep Sea Res. Part I Oceanogr. Res. Pap.* **94**, 195–206 (2014).
45. B. Allen, R. Stokey, T. Austin, N. Forrester, R. Goldsborough, M. Purcell, C. von Alt, REMUS: A Small, Low Cost AUV; System Description, Field Trials and Performance Results, in *OCEANS '97 MTS/IEEE Conference Proceedings* (IEEE, 1997), vol. 2, pp. 994–1000.
46. H. Schmidt, J. G. Bellingham, M. Johnson, D. Herold, D. M. Farmer, R. Pawlowicz, Real-time frontal mapping with AUVs in a coastal environment, in *OCEANS 96 MTS/IEEE Conference Proceedings. The Coastal Ocean–Prospects for the 21st Century* (IEEE, 1996), vol. 3, pp. 1094–1098.
47. M. A. Tivey, H. P. Johnson, A. Bradley, D. Yoerger, Thickness of a submarine lava flow determined from near-bottom magnetic field mapping by autonomous underwater vehicle. *Geophys. Res. Lett.* **25**, 805–808 (1998).
48. K. W. Nicholls, E. P. Abrahamsen, K. J. Heywood, K. Stansfield, S. Østerhus, High-latitude oceanography using the autosub autonomous underwater vehicle. *Limnol. Oceanogr.* **53**, 2309–2320 (2008).
49. T. O. Fossum, G. M. Fragoso, E. J. Davies, J. E. Ullgren, R. Mendes, G. Johnsen, I. Ellingsen, J. Eidsvik, M. Ludvigsen, K. Rajan, Toward adaptive robotic sampling of phytoplankton in the coastal ocean. *Sci. Robot.* **4**, eaav3041 (2019).
50. K. M. Yamahara, C. M. Preston, J. Birch, K. Walz, R. Marin III, S. Jensen, D. Pargett, B. Roman, W. Ussler III, Y. Zhang, J. Ryan, B. Hobson, B. Kieft, B. Raanan, K. D. Goodwin, F. P. Chavez, C. Scholin, *In-situ* autonomous acquisition and preservation of marine environmental DNA using an autonomous underwater vehicle. *Front. Mar. Sci.* **6**, 373 (2019).
51. L. E. Bird, A. Sherman, J. Ryan, Development of an active, large volume, discrete seawater sampler for autonomous underwater vehicles, in *OCEANS '07 MTS/IEEE Conference Proceedings* (IEEE, 2007), pp. 1–5.
52. Y. Zhang, B. Kieft, B. W. Hobson, J. P. Ryan, B. Barone, C. M. Preston, B. Roman, B.-Y. Raanan, R. Marin III, T. C. O'Reilly, C. A. Rueda, D. Pargett, K. M. Yamahara, S. Poulos, A. Romano, G. Foreman, H. Ramm, S. T. Wilson, E. F. De Long, D. M. Karl, J. M. Birch, J. G. Bellingham, C. A. Scholin, Autonomous tracking and sampling of the deep chlorophyll maximum layer in an open-ocean eddy by a long-range autonomous underwater vehicle. *IEEE J. Ocean. Eng.* **45**, 1308–1321 (2020).
53. D. L. Valentine, G. B. Fisher, O. Pizarro, C. L. Kaiser, D. Yoerger, J. A. Breier, J. Tarn, Autonomous marine robotic technology reveals an expansive benthic bacterial community relevant to regional nitrogen biogeochemistry. *Environ. Sci. Technol.* **50**, 11057–11065 (2016).
54. A. F. Govindarajan, J. Pineda, M. Purcell, J. A. Breier, Species- and stage-specific barnacle larval distributions obtained from AUV sampling and genetic analysis in Buzzards Bay, Massachusetts, USA. *J. Exp. Mar. Biol. Ecol.* **472**, 158–165 (2015).
55. B. W. Hobson, J. G. Bellingham, B. Kieft, R. McEwen, M. Godin, Y. Zhang, Tethys-class long range AUVs - extending the endurance of propeller-driven cruising AUVs from days to weeks, in *2012 IEEE/OES Autonomous Underwater Vehicles (AUV)* (IEEE, 2012), pp. 1–8.
56. D. T. Roper, A. B. Phillips, C. A. Harris, G. Salavasidis, M. Pebody, R. Templeton, S. V. S. Amma, M. Smart, S. M. Phail, Autosub long range 1500: An ultra-endurance AUV with 6000 Km range, in *OCEANS 2017 - Aberdeen* (IEEE, 2017), pp. 1–5.
57. M. V. Jakuba, J. A. Breier, D. Gómez-Ibáñez, K. Tradd, M. A. Saito, Clio: An autonomous vertical sampling vehicle for global ocean biogeochemical mapping, in *2018 IEEE/OES Autonomous Underwater Vehicles Workshop (AUV)* (IEEE, 2018), pp. 1–8.
58. S. Levitus, *Climatological Atlas of the World Ocean* (NOAA Professional Paper 13, U.S. Department of Commerce, National Oceanic and Atmospheric Administration, 1982).
59. J. Bell, J. Betts, E. Boyle, MITESS: A moored in situ trace element serial sampler for deep-sea moorings. *Deep Sea Res. Part I Oceanogr. Res. Pap.* **49**, 2103–2118 (2002).
60. A. M. Bradley, M. D. Feezor, H. Singh, F. Yates Sorrell, Power systems for autonomous underwater vehicles. *IEEE J. Ocean. Eng.* **26**, 526–538 (2001).
61. C. C. Eriksen, Systems and methods for compensating for compressibility and thermal expansion coefficient mismatch in buoyancy controlled underwater vehicles. U.S. Patent 8,726,827 (2014).
62. S. J. Salter, M. J. Cox, E. M. Turek, S. T. Calus, W. O. Cookson, M. F. Moffatt, P. Turner, J. Parkhill, N. J. Loman, A. W. Walker, Reagent and laboratory contamination can critically impact sequence-based microbiome analyses. *BMC Biol.* **12**, 87 (2014).
63. C. S. Sheik, B. K. Reese, K. I. Twing, J. B. Sylvan, S. L. Grim, M. O. Schrenk, M. L. Sogin, F. S. Colwell, Identification and removal of contaminant sequences from ribosomal gene databases: Lessons from the Census of Deep Life. *Front. Microbiol.* **9**, 840 (2018).
64. M. A. Saito, E. M. Bertrand, M. E. Duffy, D. A. Gaylord, N. A. Held, W. J. Hervey IV, R. L. Hettich, P. D. Jagtap, M. G. Janech, D. B. Kinkade, D. H. Leary, M. R. McIlvin, E. K. Moore, R. M. Morris, B. A. Neely, B. L. Nunn, J. K. Saunders, A. I. Shepherd, N. I. Symmonds, D. A. Walsh, Progress and challenges in ocean metaproteomics and proposed best practices for data sharing. *J. Proteome Res.* **18**, 1461–1476 (2019).
65. Y. Zhang, R. S. McEwen, J. P. Ryan, J. G. Bellingham, Design and tests of an adaptive triggering method for capturing peak samples in a thin phytoplankton layer by an autonomous underwater vehicle. *IEEE J. Ocean. Eng.* **35**, 785–796 (2010).
66. M. A. Saito, C. Breier, M. Jakuba, M. McIlvin, D. Moran, Envisioning a chemical metaproteomics capability for biochemical research and diagnosis of global ocean microbiomes, in *The Chemistry of Microbiomes* (National Academies Press, 2017), chap. 5, pp. 29–36.
67. M. V. Jakuba, C. R. German, A. D. Bowen, L. L. Whitcomb, K. Hand, A. Branch, S. Chien, C. M. Farland, Teleoperation and robotics under ice: Implications for planetary exploration, in *2018 IEEE Aerospace Conference* (IEEE, 2018), pp. 1–14.
68. D. R. Yoerger, A. M. Bradley, B. B. Walden, H. Singh, R. Bachmayer, Surveying a subsea lava flow using the Autonomous Benthic Explorer (ABE). *Int. J. Syst. Sci.* **29**, 1031–1044 (1998).
69. C. L. Kaiser, D. R. Yoerger, J. C. Kinsey, S. Kelley, A. Billings, J. Fujii, S. Suman, M. Jakuba, Z. Berkowitz, C. R. German, A. D. Bowen, The design and 200 day per year operation of the autonomous underwater vehicle Sentry, in *2016 IEEE/OES Autonomous Underwater Vehicles (AUV)* (IEEE, 2016), pp. 251–260.
70. L. L. Whitcomb, M. V. Jakuba, J. C. Kinsey, S. C. Martin, S. E. Webster, J. C. Howland, C. L. Taylor, D. Gomez-Ibanez, D. R. Yoerger, Navigation and control of the Nereus hybrid underwater vehicle for global ocean science to 10,903 m depth: Preliminary results, in *2010 IEEE International Conference on Robotics and Automation* (IEEE, 2010), pp. 594–600.
71. H. Craig, K. K. Turekian, The GEOSECS Program: 1973–1976. *Earth Planet. Sci. Lett.* **32**, 217–219 (1976).
72. K. A. Fanning, Nutrient provinces in the sea: Concentration ratios, reaction rate ratios, and ideal covariation. *J. Geophys. Res. Oceans* **97**, 5693–5712 (1992).
73. L. D. Talley, R. A. Feely, B. M. Sloyan, R. Wanninkhof, M. O. Baringer, J. L. Bullister, C. A. Carlson, S. C. Doney, R. A. Fine, E. Firing, N. Gruber, D. A. Hansell, M. Ishii, G. C. Johnson, K. Katsumata, R. M. Key, M. Kramp, C. Langdon, A. M. Macdonald, J. T. Mathis, E. L. McDonagh, S. Mecking, F. J. Millero, C. W. Mordy, T. Nakano, C. L. Sabine, W. M. Smethie, J. H. Swift, T. Tanhua, A. M. Thurnherr, M. J. Warner, J.-Z. Zhang, Changes in ocean heat, carbon content, and ventilation: A review of the first decade of GO-SHIP global repeat hydrography. *Ann. Rev. Mar. Sci.* **8**, 185–215 (2016).

74. W. J. Jenkins, M. Hatta, J. N. Fitzsimmons, R. Schlitzer, N. T. Lanning, A. Shiller, N. R. Buckley, C. R. German, D. E. Lott III, G. Weiss, L. Whitmore, K. Casciotti, P. J. Lam, G. A. Cutter, K. L. Cahill, An intermediate-depth source of hydrothermal ^{3}He and dissolved iron in the North Pacific. *Earth Planet. Sci. Lett.* **539**, 116223 (2020).
75. C. Measures, M. Hatta, J. Fitzsimmons, P. Morton, Dissolved Al in the zonal N Atlantic section of the US GEOTRACES 2010/2011 cruises and the importance of hydrothermal inputs. *Deep Sea Res. Part II Oceanogr. Res. Pap.* **116**, 176–186 (2015).
76. NASA Goddard Space Flight Center, Ocean Ecology Laboratory, Ocean Biology Processing Group, Moderate-resolution Imaging Spectroradiometer (MODIS) Aqua chlorophyll data; 2019 reprocessing (2020).
77. J. A. Breier, B. M. Toner, S. C. Fakra, M. A. Marcus, S. N. White, A. M. Thurnherr, C. R. German, Sulfur, sulfides, oxides and organic matter aggregated in submarine hydrothermal plumes at 9°50′N East Pacific Rise. *Geochim. Cosmochim. Acta* **88**, 216–236 (2012).
78. A. McCarthy, E. Chiang, M. L. Schmidt, V. J. Denef, RNA preservation agents and nucleic acid extraction method bias perceived bacterial community composition. *PLOS ONE* **10**, e0121659 (2015).
79. J. G. Caporaso, C. L. Lauber, W. A. Walters, D. Berg-Lyons, J. Huntley, N. Fierer, S. M. Owens, J. Betley, L. Fraser, M. Bauer, N. Gormley, J. A. Gilbert, G. Smith, R. Knight, Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. *ISME J.* **6**, 1621–1624 (2012).
80. J. J. Kozich, S. L. Westcott, N. T. Baxter, S. K. Highlander, P. D. Schloss, Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the MiSeq Illumina sequencing platform. *Appl. Environ. Microbiol.* **79**, 5112–5120 (2013).
81. A. M. Seekatz, C. M. Theriot, C. T. Molloy, K. L. Wozniak, I. L. Bergin, V. B. Young, Fecal microbiota transplantation eliminates *Clostridium difficile* in a murine model of relapsing disease. *Infect. Immun.* **83**, 3838–3846 (2015).
82. A. M. Eren, L. Maignien, W. J. Sul, L. G. Murphy, S. L. Grim, H. G. Morrison, M. L. Sogin, Oligotyping: Differentiating between closely related microbial taxa using 16S rRNA gene data. *Methods Ecol. Evol.* **4**, 1111–1119 (2013).
83. A. M. Eren, H. G. Morrison, P. J. Lescault, J. Reveillaud, J. H. Vineis, M. L. Sogin, Minimum entropy decomposition: Unsupervised oligotyping for sensitive partitioning of high-throughput marker gene sequences. *ISME J.* **9**, 968–979 (2014).
84. S. M. Huse, L. Dethlefsen, J. A. Huber, D. M. Welch, D. A. Relman, M. L. Sogin, Exploring microbial diversity and taxonomy using SSU rRNA hypervariable tag sequencing. *PLOS Genet.* **4**, e1000255 (2008).
85. P. D. Schloss, S. L. Westcott, T. Ryabin, J. R. Hall, M. Hartmann, E. B. Hollister, R. A. Lesniewski, B. B. Oakley, D. H. Parks, C. J. Robinson, J. W. Sahl, B. Stres, G. G. Thallinger, D. J. Van Horn, C. F. Weber, Introducing mothur: Open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Appl. Environ. Microbiol.* **75**, 7537–7541 (2009).
86. C. S. Hughes, S. Foehr, D. A. Garfield, E. E. Furlong, L. M. Steinmetz, J. Krijgsveld, Ultrasensitive proteome analysis using paramagnetic bead technology. *Mol. Syst. Biol.* **10**, 757 (2014).
87. A. Bankevich, P. A. Pevzner, TruSPAdes: Barcode assembly of TruSeq synthetic long reads. *Nat. Methods* **13**, 248–250 (2016).
88. A. Bankevich, S. Nurk, D. Antipov, A. A. Gurevich, M. Dvorkin, A. S. Kulikov, V. M. Lesin, S. I. Nikolenko, S. Pham, A. D. Prjibelski, A. V. Pyshkin, A. V. Sirotkin, N. Vyahhi, G. Tesler, M. A. Alekseyev, P. A. Pevzner, SPAdes: A new genome assembly algorithm and its applications to single-cell sequencing. *J. Comput. Biol.* **19**, 455–477 (2012).
89. D. Antipov, A. Korobeynikov, J. S. McLean, P. A. Pevzner, hybridSPAdes: An algorithm for hybrid assembly of short and long reads. *Bioinformatics* **32**, 1009–1015 (2016).
90. A. D. Prjibelski, I. Vasilinetc, A. Bankevich, A. Gurevich, T. Krivosheeva, S. Nurk, S. Pham, A. Korobeynikov, A. Lapidus, P. A. Pevzner, ExSPAnder: A universal repeat resolver for DNA fragment assembly. *Bioinformatics* **30**, i293–i301 (2014).
91. M. Rho, H. Tang, Y. Ye, FragGeneScan: Predicting genes in short and error-prone reads. *Nucleic Acids Res.* **38**, e191 (2010).
92. C. L. Dupont, J. P. McCrow, R. Valas, A. Moustafa, N. Walworth, U. Goodenough, R. Roth, S. L. Hogle, J. Bai, Z. I. Johnson, E. Mann, B. Palenik, K. A. Barbeau, J. C. Venter, A. E. Allen, Genomes and gene expression across light and productivity gradients in eastern subtropical Pacific microbial communities. *ISME J.* **9**, 1076–1092 (2015).
93. L. I. Gordon, J. C. Jennings Jr., A. A. Ross, J. M. Krest, A suggested protocol for continuous flow automated analysis of seawater nutrients (phosphate, nitrate, nitrite and silicic acid) in the WOCE Hydrographic Program and the Joint Global Ocean Fluxes Study, in *Methods Manual WHPO 91-1* (WOCE Hydrographic Program Office, 1993).
94. F. A. J. Armstrong, C. R. Stearns, J. D. H. Strickland, The measurement of upwelling and subsequent biological process by means of the Technicon Autoanalyzer® and associated equipment. *Deep Sea Res.* **14**, 381–389 (1967).
95. K. P. Koltermann, V. V. Gouretski, K. Jancke, *Hydrographic Atlas of the World Ocean Circulation Experiment (WOCE) Volume 3: Atlantic Ocean*, M. Sparrow, P. Chapman, J. Gould, Eds. (International WOCE Project Office, 2011).

Acknowledgments: We thank the crews of R/V *Neil Armstrong* and R/V *Atlantic Explorer* for assistance and support with the ocean deployments. We thank the NASA Goddard Space Flight Center, Ocean Ecology Laboratory, Ocean Biology Processing Group for the Moderate-resolution Imaging Spectroradiometer (MODIS) Aqua Chlorophyll Data, NASA OB.DAAC, Greenbelt, MD, USA, https://oceancolor.gsfc.nasa.gov/l3/ (accessed on 1 May 2020), and the ESA and developers of the SNAP tool used to illustrate these data. We thank B. Searle, J. Haskin, R. Chmiel, and M. Kellog for assistance at sea. **Funding:** The development of *Clio* was supported by U.S. National Science Foundation (NSF) grants OCE-1333212 to J.A.B., M.V.J., and M.A.S. and OCE-1334727 to G.J.D. The Sargasso Sea study was supported by U.S. NSF grants OCE-1658030 to M.A.S. and M.V.J., OCE-1657885 to R.J., and OCE-1658067 to J.A.B. U.S. NSF grants OCE-1924554 to M.A.S. and M.V.J., and OCE-1924508 to J.A.B. also supported this work. The Gordon and Betty Moore Foundation grant GBMF3782 provided support to M.A.S. U.S. NSF grants OCE-1756884 and OCE-1136477 and The Gordon and Betty Moore Foundation grant GBMF3828 provided support to A.E.A. The National Oceanic and Atmospheric Administration (NOAA) Educational Partnership Program with Minority-Serving Institutions Cooperative Agreement Award #NA16SEC4810009 provided support to J.A.B. and B.A.A.; the contents of this article are solely the responsibility of the authors and do not necessarily represent the official views of the U.S. Department of Commerce, NOAA. **Author contributions:** J.A.B. designed the sampling systems. M.V.J. designed the vehicle. The overall vehicle concept and use case scenarios were envisioned by J.A.B., M.A.S., M.V.J., and G.J.D. The proteomic methodology and associated metagenomic libraries were developed by M.A.S., M.R.M., D.M.M., A.E.A., and C.L.D., and they performed the analyses. G.J.D. and S.L.G. developed the DNA/RNA methodology and performed the analyses. R.J., M.V.J., M.A.S., M.R.M., S.L.G., E.W.C., B.A.A., and J.A.B. developed the field oceanographic methods and performed the fieldwork. J.A.B. wrote the manuscript with contributions from all authors. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** The hydrographic, nutrient, and protein data for this study have been deposited in Biological and Chemical Oceanography Data Management Office (www.bco-dmo.org) (project 765945), and raw mass spectrometry and DNA sequence datasets have been deposited at the EBI PRIDE (project PXD018067) and NCBI repositories (project PRJNA607796), respectively. All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials.

Submitted 18 May 2020
Accepted 21 October 2020
Published 25 November 2020
10.1126/scirobotics.abc7104

Citation: J. A. Breier, M. V. Jakuba, M. A. Saito, G. J. Dick, S. L. Grim, E. W. Chan, M. R. McIlvin, D. M. Moran, B. A. Alanis, A. E. Allen, C. L. Dupont, R. Johnson, Revealing ocean-scale biochemical structure with a deep-diving vertical profiling autonomous vehicle. *Sci. Robot.* **5**, eabc7104 (2020).

Revealing ocean-scale biochemical structure with a deep-diving vertical profiling autonomous vehicle

John A. Breier, Michael V. Jakuba, Mak A. Saito, Gregory J. Dick, Sharon L. Grim, Eric W. Chan, Matthew R. McIlvin, Dawn M. Moran, Brianna A. Alanis, Andrew E. Allen, Chris L. Dupont and Rod Johnson

Sci. Robotics **5**, eabc7104.
DOI: 10.1126/scirobotics.abc7104