

Open camera or QR reader and  
scan code to access this article  
and other resources online.



# Testing Rover Science Protocols to Identify Possible Biosignatures on Mars: Achieving Sampling Goals Under a Highly Constrained Time Line

R. Aileen Yingst,<sup>1</sup> Julie K. Bartley,<sup>2</sup> Thomas J. Chidsey, Jr.,<sup>3</sup> Barbara A. Cohen,<sup>4</sup> Natalie Curran,<sup>4</sup> Brian M. Hynek,<sup>5</sup> Linda C. Kah,<sup>6</sup> Michelle E. Minitti,<sup>7</sup> Michael D. Vanden Berg,<sup>3</sup> Rebecca M.E. Williams,<sup>1</sup> John Gemperline,<sup>5</sup> Michael Lotto,<sup>5</sup> Sarah Black,<sup>8</sup> Bruce C. Bartley,<sup>9</sup> and Taylor Pearson<sup>2</sup>

## Abstract

At a Mars analog site in Utah, we tested two science operation methods for data acquisition and decision-making protocols: a scenario where the tactical day is replanned, but major adjustments may still be made before plan delivery; and a scenario in which the sol path must largely be planned before a given tactical planning day and very few adjustments to the plan may be made. The goal was to provide field-tested insight into operations planning for rover missions where science operations must facilitate the efficient choice of sampling locations at a site relevant to searching for habitability and biosignatures. Results of the test indicate that replanning sol paths did not result in a sol cost savings nor did it improve science return or optimal biologically relevant sample collection. In addition because facies variations in an environment can be subtle and evident only at scales below orbital resolution, acquiring systematic observations is crucial. We also noted that while spectral data provided insight into the chemical components as a whole at this site, they did not provide a guide to targets for which the traverse should be altered. Finally, strategic science planning must include a special effort to account for terrain. Key Words: Rover—Science protocols—Mars analog. Astrobiology 22, 1310–1329.

## 1. Introduction

Planetary exploration relies primarily on remote observation, and an increasing amount of data are being provided by *in situ* landers and mobile scientific platforms. One of the most challenging aspects of rover-based science operations is the need for science teams to balance maximizing scientific return and minimizing expenditure of resources, such as time, energy, and data volume. The most frequently used strategy of maximizing science return is one

that rarely retreads previously covered ground, which can mean leaving a site before certain observations can be conducted or limiting the number or type of activities that can be achieved during any given planning period. The trade-offs involved in resource conservation and data collection can therefore lead to an either real or perceived limitation of the scientific understanding of a field area.

The nature of assessing the success and efficiency of operational choices, however, is difficult, and understanding the most resource-effective ways to conduct remote

<sup>1</sup>Planetary Science Institute, Tucson, Arizona, USA.

<sup>2</sup>Department of Geology, Gustavus Adolphus College, St. Peter, Minnesota, USA.

<sup>3</sup>Utah Geological Survey, Salt Lake City, Utah, USA.

<sup>4</sup>Goddard Space Flight Center, Greenbelt, Maryland, USA.

<sup>5</sup>Department of Geological Sciences, University of Colorado Boulder, Boulder, Colorado, USA.

<sup>6</sup>Department of Earth and Planetary Sciences, University of Tennessee-Knoxville, Knoxville, Tennessee, USA.

<sup>7</sup>Framework, Silver Spring, Maryland, USA.

<sup>8</sup>USGS, Flagstaff, Arizona, USA.

<sup>9</sup>Concordia College, Moorhead, Minnesota, USA.

fieldwork for maximum science benefit represents a challenging, yet critical, dimension of planetary exploration.

Terrestrial rover analog tests provide one mechanism to address the effectiveness of science operations for remote research (*e.g.*, Greeley *et al.*, 1994; Whittaker *et al.*, 1997; Arvidson *et al.*, 2000; Stoker *et al.*, 2001, 2002; Lee, 2007; Lee *et al.*, 2017; Fong *et al.*, 2010; Eppler *et al.*, 2013; Graham *et al.*, 2015; Lim *et al.*, 2018; Osinski *et al.*, 2019). These tests, however, often rely heavily on the successful operation of complex technical equipment. In such cases, equipment error or malfunction often negatively impacts scientific decision-making processes and thus impedes the study's ability to isolate science operations as a variable distinct from equipment function (Eppler *et al.*, 2013; Cohen *et al.*, 2015; Graham *et al.*, 2015; Yingst *et al.*, 2015).

By contrast, the GeoHeuristic Operational Strategies Testing (GHOST) program isolates the science operations' decision-making process by using commercial, off-the-shelf technology, minimizing variables related to rover and equipment function (Yingst *et al.*, 2018, 2020). Previous GHOST field tests have focused on investigating the relative efficiency of executing all science activities on a single traverse, compared with using a reconnaissance pass—a method of traverse previously termed a “walkabout,” in which the remote instruments are used to assess the site, and then the site is traversed again, with the more resource-intensive instruments used on a subset of high-priority targets—before detailed contact science on a subset of high-value targets (Yingst *et al.*, 2016, 2018, 2020).

Building from the results of previous experiments, we provide in the present study results of a test designed specifically to explore the effect of these two science operation strategies, given an overall scenario in which the number of planning periods and the ability to adjust the plan of observations are both highly constrained.

## 2. Methods

### 2.1. Understanding rover operations in the context of GHOST field tests

Conducting field geology with a remote, roving science platform requires the use of instrument-acquired data to make interpretations comparable with those made by a traditional field geologist. In such a scenario, available science data directly inform decisions about subsequent observations and activities to plan (*i.e.*, what observations to make), not the method of acquiring those data (*i.e.*, what tools to use). To test the science protocols without having malfunctioning equipment, communication failures, or other such adverse condition derail results, GHOST field tests use a “roverless roving” approach, where generalized commercial, off-the-shelf instruments provide visual, compositional, and geochemical data similar to flight-ready instruments (Yingst *et al.*, 2014). In such a scenario, humans stand in for the rover itself but do not allow their geologic knowledge to inform data acquisition or interpretation, thus replicating the rover's inability to yield independent analysis or suggest appropriate follow-on observations.

GHOST field tests also utilize a highly compressed planning schedule. Rover science operations with high latency (*i.e.*, an extended time between commanding and execution, as experienced in most planetary missions)

require significant planning to best utilize limited resources. For Mars, this typically means that, in a single tactical planning shift, scientists will plan activities for 1 to 4 martian days (or “sols”), then send that entire package up to the rover to be executed while subsequent sols are being planned. Limits to time, power, data volume, and other mission resources also change from sol to sol, requiring any test of science operation efficiency to estimate what science observations and enabling rover activities fit into a given planning cycle. GHOST tests use the Mars Science Laboratory (MSL) mission (Vasavada *et al.*, 2014) as a baseline, leveraging multiple years of MSL science operation to estimate the duration and power costs of commonly executed activities. This estimate also encompasses data downlink, or the data available to the teams for planning subsequent sols; for simplicity of execution, it is generously assumed that all data acquired on one sol can be downlinked for availability to the teams planning the next sol. It is important to note that the sol cost utilized in this study is based on the unique payload of the Curiosity rover, which differs from that of the earlier Mars Exploration Rover (MER) missions (Arvidson *et al.*, 2000; Squyres *et al.*, 2003; Bass *et al.*, 2005) and the Mars 2020 Perseverance rover mission (Milkovich *et al.*, 2022).

In particular, the number and type of activities that Mars 2020 (currently operating as of this writing) may fit into any given planning period may differ somewhat from what is assumed here.

Based on MSL averages, we baseline activities during a single sol as 1 h of active remote data acquisition, plus either a drive of 50–100 m or multiple contact science observations. The estimated sol cost for each activity is shown in Table 1. To calculate the total sols used for each activity, each rover action requested during the field test was recorded and converted to an estimated number of sols that would be required to execute that action on Mars. This allowed the GHOST field team to use limited field time to perform multiple planning sols.

### 2.2. Instruments used in GHOST

Commercial off-the-shelf instruments were chosen to mimic the general data format and resolution that have been common to Mars rovers (*e.g.*, Black and Hynes, 2018). These instruments include those dedicated to both landscape and hand lens scale imaging, visible to near-infrared (VNIR) spectrometry, and mineralogical analysis. Commercial instrument details are outlined in Table 2.

We used a Canon Rebel XL digital single-lens reflex camera with a macrolens to provide the range of resolutions that might be covered by rover RGB cameras, including both mast-mounted and arm-mounted imagers. The field-portable VNIR spectrometer (Malvern Panalytical) provided rapid collection of hyperspectral data, resulting in whole-rock mineralogy including that of phyllosilicate, Fe-oxides/hydroxide, sulfate, and silica phases. Field spectra were analyzed using the Environment for Visualizing Images (ENVI) software from L3/Harris Geospatial Solutions. Mineral absorptions were manually matched to known library spectra using the USGS splib07 spectral library (Kokaly *et al.*, 2017) and the Compact Reconnaissance Imaging Spectrometer for Mars (CRISM) Analysis Toolkit (CAT) spectral library available on the Planetary Data System Geosciences Node at <https://pds-geosciences.wustl.edu/missions/mro/crism.htm>

TABLE 1. ESTIMATED SOL COST FOR EACH ROVER ACTIVITY

<i>Sol cost</i>	<i>Activities</i>	<i>Instrument equivalence</i>
1	Up to three remote composition analyses Two small- or one large-context mosaic One downlook image Drive (50–100 m) or bump with PDI	ChemCam or SuperCam Mastcam or Mastcam-Z or Navcam MARDI or Navcam Mobility
1	Up to three remote composition analyses Two small or one large context mosaic One downlook image Targeted contact science	ChemCam, SuperCam Mastcam or Mastcam-Z or Navcam MARDI or Navcam MAHLI/APXS/DRT or SHERLOC/WATSON or SHERLOC/PXL
1	“Touch and Go” rapid contact science One remote composition observation or one small-context mosaic Drive (50–100 m) or bump with PDI One downlook image	MAHLI/APXS/DRT or SHERLOC/WATSON or SHERLOC/PXL ChemCam or SuperCam, Mastcam or Mastcam-Z Mobility MARDI or Navcam
2	Drive for targeting One very large-context mosaic Up to four remote composition analyses One downlook image	Mobility Mastcam or Mastcam-Z ChemCam, SuperCam MARDI or Navcam
5	Drill	Drill/cache hardware, drill documentation instruments

PDI= postdrive imaging.

Powder X-ray diffraction (XRD; Olympus Terra) provided bulk mineralogy and was also used to calculate oxide abundances, comparable with data obtained by X-ray fluorescence (XRF). The field-portable XRD could identify sulfate, Fe-oxide/hydroxide, and silicate phases. XRD analyses used in this study were able to suggest the presence of phyllosilicate phases, but specific identifications were not made available to the field teams, as this would require additional sample processing that was not available to Mars rovers. Therefore, such identification of phyllosilicate phases was not provided to the field teams. Diffraction pattern matching was completed using X Powder software. Field samples were matched to known library patterns using the “difdata” library from the American Mineralogist Crystal Structure Database (Downs and Hall-Wallace, 2003). After determining the bulk mineralogy, quantitative mineral abundances were converted to elemental oxide percentages to simulate XRF data.

### 2.3. Operational testing scenarios

In this field investigation, we tested two distinct approaches to rover-based science operations and data acquisition under a constrained timeline. The first approach was “tac-

tically responsive” and is comparable with that used by the MER and MSL missions. In this scenario, a notional plan is strategically developed a day or two in advance of a given tactical planning cycle and is based on the analysis of both orbital and *in situ* data. In a tactically responsive scenario, changes to the notional observations and activities are relatively easy to make as the most recent data are downlinked. The team can readily decide to adjust the sol path to remain in the area for analysis of additional targets or to conduct a short drive to a target deemed more suitable to strategic goals.

By comparison, a “strategically constrained” approach is designed to focus on strategic goals and is comparable with that used by the M2020 (Mars 2020) mission. In this scenario, goals for a given region of interest (ROI, or campaign) are determined strategically using primarily orbital data. In this scenario, a strategic campaign science group produces a sol-by-sol path of fixed length that includes a prioritized list of observations, potential targets for those observations, and specific goals for mission deliverables (such as samples to be acquired).

The primary goal of this scenario is to ensure that goals for both daily tactical planning (Wilson *et al.*, 2017) and the overall mission timeline can be met. Such

TABLE 2. FIELD AND ANALOGOUS MARS INSTRUMENTATION

<i>Data type</i>	<i>Mars analog instrument</i>		<i>Field instrument</i>	<i>Specifications</i>
	<i>MSL</i>	<i>Mars 2020</i>		
Imaging	Mastcam, MAHLI, Navcams	Mastcam-Z, WATSON, Navcams	DSLR camera with macrolens	~20 $\mu\text{m}$ /pixel at 10 cm from the front of the lens
VNIR	Mastcam	Mastcam-Z, SHERLOC	TerraSpec HALO field-portable imaging spectrometer	350–2500 nm with 6 nm spectral resolution
XRD/XRF	CheMin, APXS	PIXL	Olympus Terra field portable XRD	5–55° 2 $\theta$ with Cu-K $\alpha$ source

DSLR=Digital Single-lens reflex; MSL=Mars Science Laboratory; VNIR=visible to near-infrared; XRD=X-ray diffraction; XRF=X-ray fluorescence.

a strategically constrained approach allows the tactical team minimal opportunity to change the sol path, even if prechosen target regions do not meet strategic expectations. Changes to the sol path must be argued several sols in advance. Choosing to use additional sols in a deviation from the primary campaign would require subtracting sols and, thus, other observations from the future sol path.

#### 2.4. Data acquisition and personnel

The primary goal of this field study was to compare “tactically responsive” and “strategically constrained” science operation protocols in terms of the ability to extract appropriate contextual data for a sampling mission with a highly constrained timeline, where the mission science goals are focused on identifying and characterizing biosignatures and habitability. The test was designed to track deviations from the plan and the resultant differences in science outcomes and determine the number of sols used by each method.

Both operational teams started with the same notional plan of observations and activities, based on orbital-scale data, and both explored four discrete ROIs in sequence. The “tactically responsive” team could adjust the time spent at each target, trade observations, or replace a strategically planned sample. By contrast, if the “strategically constrained” team were to exceed the preplanned sol allotment, the sols spent over that allotment and the science achieved in those extra sols would be considered a sol cost that would be subtracted from later activities.

Data acquisition teams were structured similarly to previous tests (Yingst *et al.*, 2016, 2020). One individual served as the independent site expert, reconnoitering the site and providing remote data sets for the teams to use in mission planning. Each of the field teams comprised two experts, one self-identified expert in sedimentology and one self-identified expert in rover operation. Science teams approached the field site blind and conducted prefield strategic planning to define hypotheses and plan notional observations and activities through the four ROI sites using only data similar to what would be available for rover mission planning (*i.e.*, orbitally derived visible and spectral data).

The science team was then divided into a two-person Tactical (tactically responsive) team and a two-person Strategic (strategically constrained) team to command the “rovers” along the planned traverse and interpret the returned data. Each science team independently conducted remote investigation of the field site through its specific human rover avatar. At the same time, an independent Tiger team examined the site using standard terrestrial geological field methods. The Tiger team provided a baseline of science results against which the rover-enabled teams could compare their findings.

### 3. Field Site

The exercise was conducted in a site considered analogous to both the MSL (Gale crater) and Mars 2020 (Jezero crater) landing sites. The site, which is located near the Interstate 80 corridor west of Salt Lake City, Utah (Fig. 1), contains geologic evidence for a variety of ancient habitable environments, including marine environments (Paleozoic bedrock) and fluvial, lacustrine, and groundwater spring

environments (Quaternary deposition). Orbital detection of both clay and carbonate phases is consistent with arguments that suggest these lithologies represent enhanced potential for preservation of biosignatures (*e.g.*, Ehlmann *et al.*, 2008; Milliken *et al.*, 2010; Wray *et al.*, 2016; Horgan *et al.*, 2020). This GHOST test focused on four ROIs on the western flank of the Grassy Mountains (Fig. 1), all of which were identified from orbital images as locations of high interest with respect to preserving evidence of habitability or biosignatures.

Bedrock at the site consists of Permian marine limestone, dolomite, and sandstone of the Oquirrh Group deposited in an inland sea (Clark and Oviatt, 2018; Hintze and Kowallis, 2021). A combination of Cordilleran orogenesis and basin and range extension resulted in north-to-south trending crustal blocks that represent the primary mountainous regions (Clark *et al.*, 2017). The landscape was then modified within pluvial environments of the Pleistocene Lake Bonneville, which was associated with the deposition of silt, sand, and gravel in the lowlands; the formation of two distinct geomorphic shorelines (*i.e.*, the 18 kya Bonneville shoreline at an elevation of  $\sim 1550$  m, and the 15 kya Provo shoreline at  $\sim 1445$  m); and deposition of a distinctive tufa proximal to the Provo shoreline (Reheis *et al.*, 2014; Oviatt and Schroeder, 2016; Clark and Oviatt, 2018).

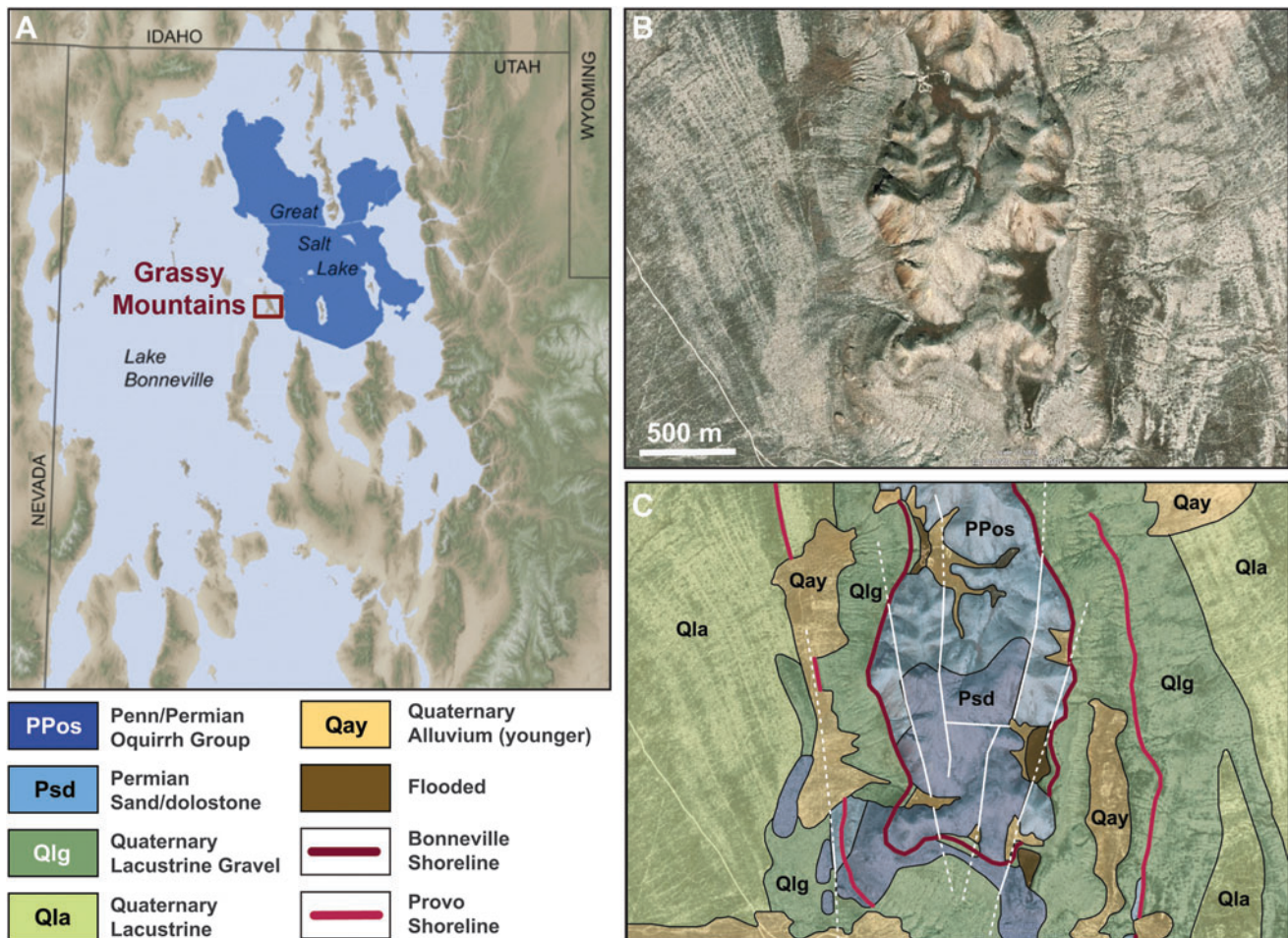
Lake Bonneville developed during the last glacial maximum, starting  $\sim 30$  kya, and covered nearly  $50,000$  km<sup>2</sup>, with maximum depths of  $\sim 300$  m (Chen and Maloof, 2017) (Fig. 1). For comparison, the putative lake at Jezero crater was  $\sim 1885$  km<sup>2</sup> and may have been as deep as 270 m (Goudge *et al.*, 2018).

## 4. Fieldwork

### 4.1. Prefieldwork site assessment

Science teams conducted a prefieldwork assessment of the  $\sim 1$  km<sup>2</sup> field site using data similar in resolution to the orbital data available for Mars. Panchromatic and visible wavelength orthoimages with a resolution of 1 m/pixel were provided by the Utah Geologic Survey and used to simulate High Resolution Imaging Science Experiment (HiRISE) images. This data set was also coarsened to 5 m/pixel to simulate Context Camera images. Simulated HiRISE stereo pairs were used to create a high-resolution (5 m/pixel) digital elevation model (DEM) and hill-shade map of the study area. An additional 70 m/pixel DEM, analogous to high-resolution stereo camera (HRSC) products, was created by degrading a 30 m/pixel US Geological Survey DEM of the region. CRISM data were simulated with Landsat 7 Thematic Mapper multispectral data. Spectral data were analyzed for relevant common materials (*e.g.*, smectite, silica, iron, carbonate) and compiled into a mineral map of the field site.

The field teams surveyed orbital data before arrival at the field site. Teams worked together to create the initial hypotheses about the field area, so that they began from a common set of observations, summarized in Figs. 2 and 3. Each team then separately planned traverses and stops for each  $100 \times 100$  m ROI (A, B, C, and D) that each believed would provide the best opportunity to characterize the geology of the region, identify possible biosignatures, and collect biologically relevant samples. The Tactical team created a generalized set of observations for each area, which



**FIG. 1.** Regional geology. (A) Regional extent of Lake Bonneville, compared with that of the current Great Salt Lake; GHOST test site locality was in the southern portions of Grassy Mountain (red rectangle). (B) Satellite image of the southern Grassy Mountain field site, courtesy of the Utah Geological Survey (<https://geology.utah.gov/apps/intgeomap/#>). (C) Geologic map of southern Grassy Mountain, showing distribution of key geologic units, location of Bonneville and Provo shorelines, and position of major N-S trending faults; adapted from Clark *et al.* (2017). GHOST = GeoHeuristic Operational Strategies Testing; N-S = north-south.

included potential stations for data and sample collection, while the Strategic team created a sol-by-sol plan for each campaign region.

#### 4.2. Data collection and analysis

Data collection in the field followed the methods described by Yingst *et al.* (2016, 2018, 2020). In the absence of engineering data, all paths were assumed to be traversable, and human rover avatars made the final judgment on the traversability of a region *in situ*. This allowance increased the efficiency of science teams and allowed ~15–20 sols of rover activities to be completed per day in the field.

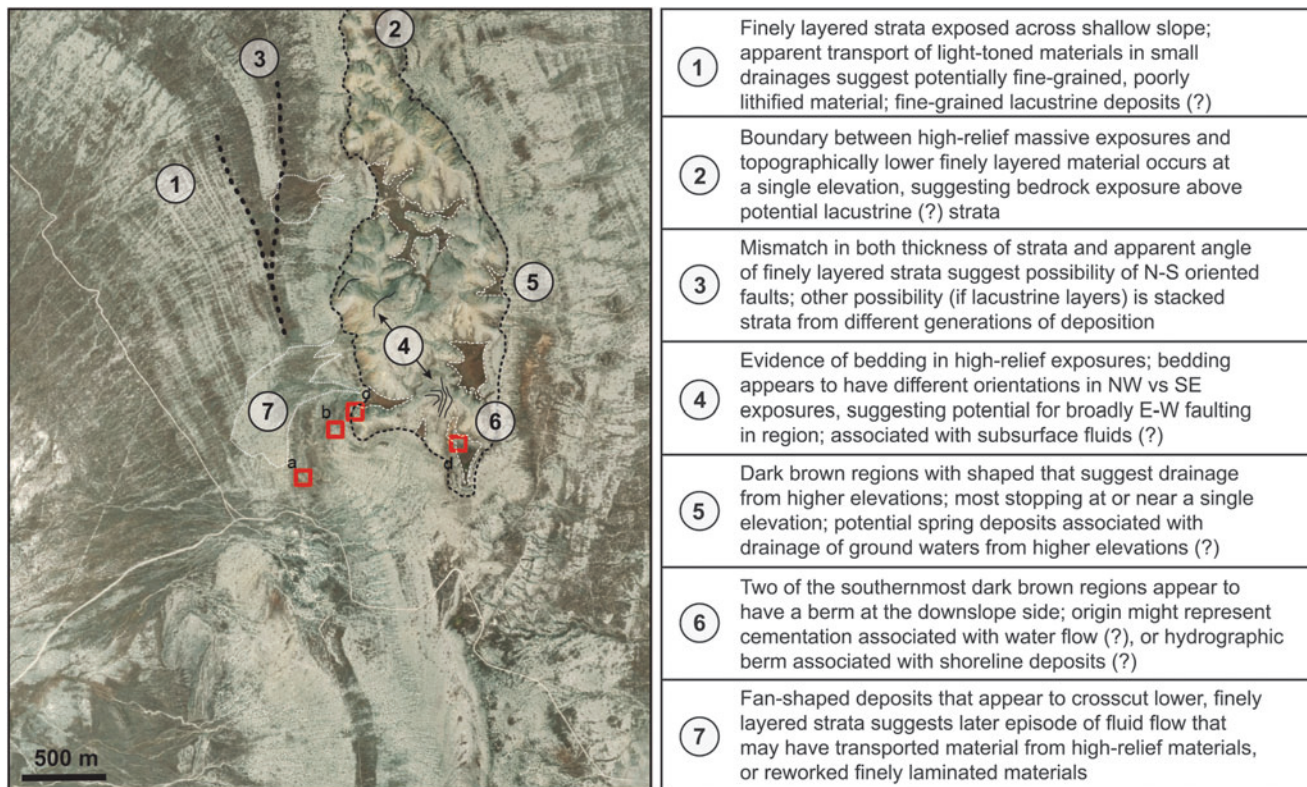
Both the Tactical and Strategic teams began analysis with an initial panoramic image of the site upon entering a new campaign region. Additional mosaics were collected after the completion of each drive. Rover teams could only drive to points that were visible in their collected images, limiting them to 100 m or less per drive sol. Upon arrival, teams could deploy VNIR and/or XRD-type instruments to collect mineralogical and chemical data. Data could then be used to

infer the composition of the rocks in the region. Additional high-resolution visible images (MAHLI/WATSON and Mastcam/Mastcam-Z) could also be collected for analysis and interpretation. Both teams were also free to choose how to traverse each site, including whether to execute a single traverse or to execute a reconnaissance pass (Yingst *et al.*, 2016, 2018, 2020), and then interrogate a subset of areas previously examined in the first pass.

However, the starting point for the “rovers,” and thus the order that the ROIs were visited, was driven entirely by the logistics of transporting equipment and personnel. This retains the integrity of the test, as experience has shown that a traverse starting point that is random with respect to the “best” geology is a real-world case; the approach to any site is usually driven by factors unrelated to science.

Concurrent with GHOST operations, the Tiger team independently investigated the site on foot and was not limited to exploring only areas that were chosen as ROIs. Geological assessment by the Tiger team provided an independent baseline interpretation of the geology, the depositional environments, and the prioritization of potential sampling localities.





**FIG. 2.** Orbital image of the southern Grassy Mountain field site, with red squares marking predetermined ROIs assigned for remote rover operations. Regions and boundaries highlighted on the image are associated with numbered, key observations agreed upon by team members before the start of remote operations. Lithologies for high-resource science (e.g., drilling or sampling) focused on materials with a high potential for preservation of biosignatures, and included fine-grained, potentially lacustrine sedimentary rock; potential hydrothermal materials associated with fractures; high-relief bedrock lithologies; and regions inferred to have experienced later fluid flow with subsequent ponding, precipitation from springs, or shoreline depositional environments. ROI, region of interest.

## 5. Data

### 5.1. Strategic team activities

The Strategic team identified high-priority locations for sampling and then planned science operations to provide the necessary contextual information for acquisition of those samples, while staying within sol counts proposed for this exercise. Using the prefield analysis (cf. Section 4.1, Figs. 2 and 3), the Strategic team determined that the site of ROI D was the highest priority for sampling, which included dark and light units within the distributary channels and the bounding ridges, each of which was determined to likely contain evidence of fluid-derived minerals and potential organic matter associated with overland flow and fluid ponding. An additional sample from exposed bedrock upstream of the channel was also planned to compare the geochemistry of bedrock with stratigraphically superimposed ponded materials.

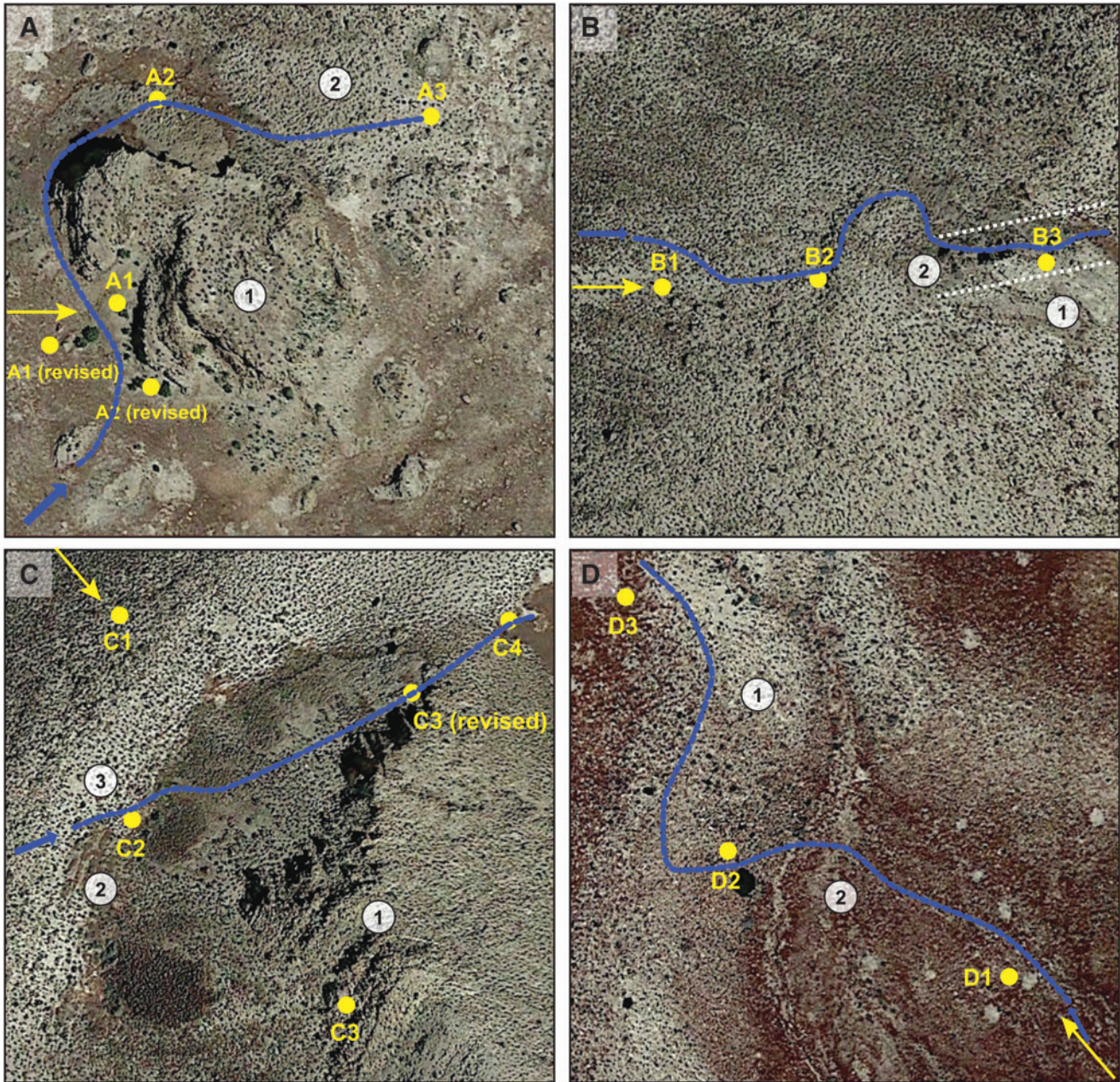
The strategic interest in the ROI D site, however, placed an operational limit on potential samples from other ROI sites. It was, therefore, determined that the remaining sites would be used primarily to build a regional context. At the other sites (ROI A–C), the team strategically planned to focus sampling on dark-toned, erosionally resistant bedrock and light-toned, fine-grained material associated with what was interpreted as regional lacustrine deposition. The team

also planned to investigate exposed bedrock and the potential cross-cutting phases to explore potential chemical alteration of bedrock and potential ground sources of fluid movement.

Although the Strategic team built a plan that was consistent with both the test scenario and the geological interpretations derived during prefield assessment, the addition of ground-based data at ROI A that revealed a likely sedimentary origin and potential biosignature retention (Fig. 4) required immediate and substantial revision of the strategic plan, from remote and contact science only to sample collection. The Strategic team determined a new sampling strategy that focused on (1) acquiring samples that would aid in the reconstruction of the number and chemistry of fluid events that affected the bedrock lithologies, while (2) retaining sample acquisition “slots” for the predetermined highest priority sampling locality late in the mission at ROI D. This new strategy led to the sampling of dolomitic karst breccia identified by remote and contact science in ROI B bedrock (Fig. 5), and carbonate-bearing soil crust and calcite vein material cutting through bedrock in ROI C (Fig. 6).

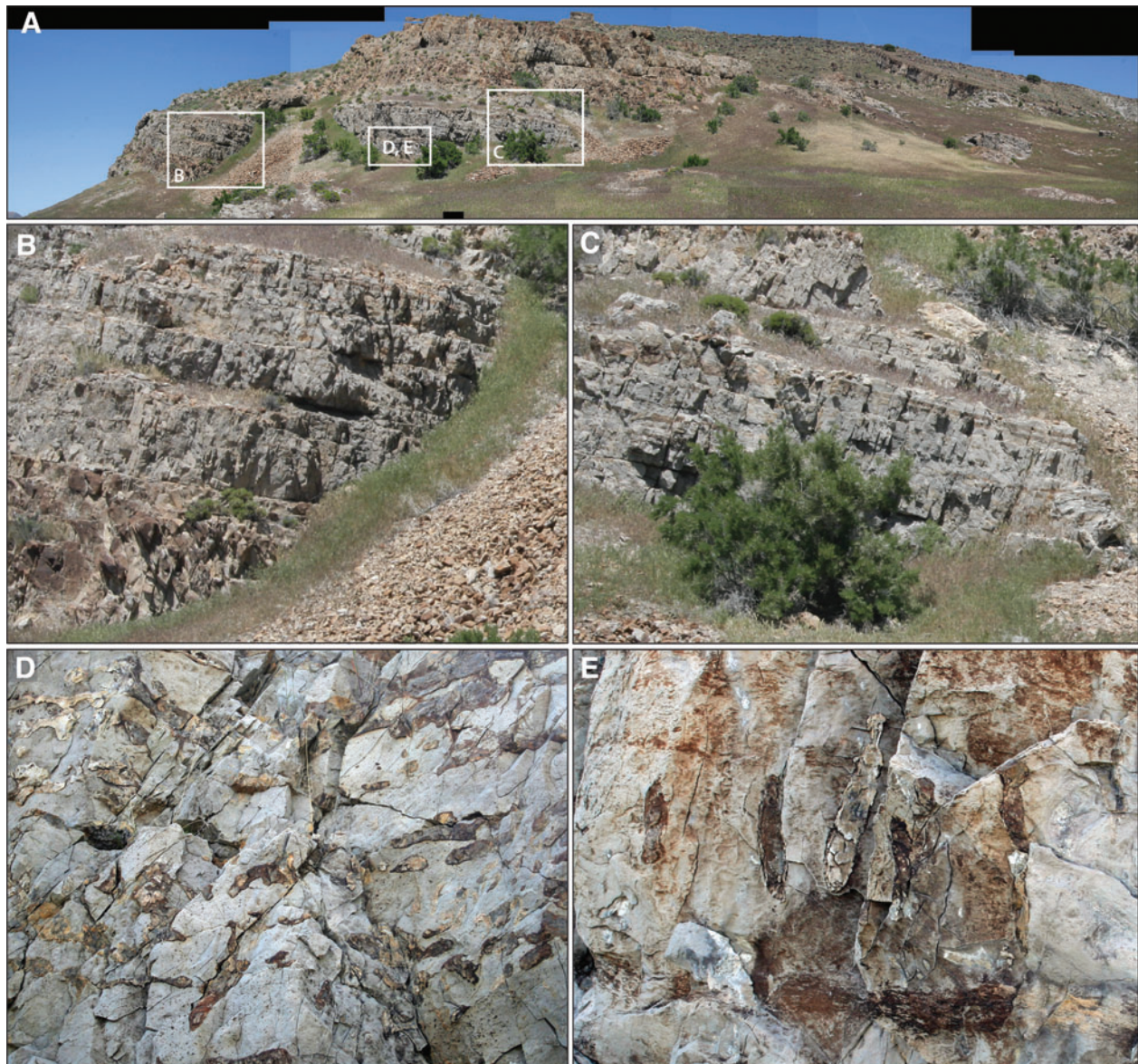
During a prefield assessment, the Strategic team identified in ROI D regional geomorphic evidence of the presence of a long-term lacustrine environment, marked by apparent shorelines that were expressed as uniform topographic ridges along the bedrock hillsides, and the potential for





**FIG. 3.** Orbital images of individual ROIs assigned for remote rover operations; images are 400 m across. **(A)** Key observations in ROI-A include exposure of inferred high-relief bedrock **(1)** with erosionally resistant, dipping strata, and low-relief light-toned materials **(2)** inferred to represent exposure of fine-grained lacustrine sediment. **(B)** Key observations in ROI-B include light-toned, flat-lying, finely layered strata **(1)** and a distinct E-W trending lineation **(2)** that defined a contact between light-toned layers and dark-toned rubbly material inferred to represent a diagenetic fracture and a potential conduit for movement of subsurface fluids. **(C)** Key observations in ROI-C include exposure of high-relief bedrock **(1)** with erosionally resistant, dipping strata, and low-relief light- to dark-toned materials **(2)** inferred to represent exposure of fine-grained lacustrine sediment. **(D)** Key observations in ROI-D include an apparent berm **(1)** texturally different from both regional high-relief bedrock and inferred finely layered lacustrine strata, and patterned ground associated with migrating channel features **(2)** inferred to represent possibly seasonal changes in groundwater flow and potential spring deposition. Yellow arrows represent direction of approach determined by the strategically constrained team, and yellow points represent predetermined localities for science operations as listed in Table 3. Blue arrows represent direction of approach determined by the tactically responsive team, and dashed blue paths represent the initial plan for traverse through each campaign area for the tactically responsive team as listed in Table 4. E-W= east-west.





**FIG. 4.** Strategic team assessment of ROI A. (A) Mosaic of region upon entering ROI A. (B, C) Characteristic alternation of gray and brown bedding inferred to represent alternation of lithologies, with hints of irregular nodular material within gray horizons. Images are  $\sim 10$  m across. (D, E) Close-up of nodules; greater resistance to erosion and measured composition suggested diagenetic chert, which is commonly associated in terrestrial environments with high potential for biosignature preservation.

short-term lacustrine environments uphill from potential shoreline ridges. Upon arrival at the locality, it became apparent that the lacustrine materials that may have been associated with topographic shorelines were not preserved as materials that could be sampled in this operational scenario. With the sol limit approaching, the Strategic team had to quickly select the three samples for this campaign region and, therefore, focused on the assessment and sampling of potential fluvial-deltaic deposits to characterize the region's fluvial history.

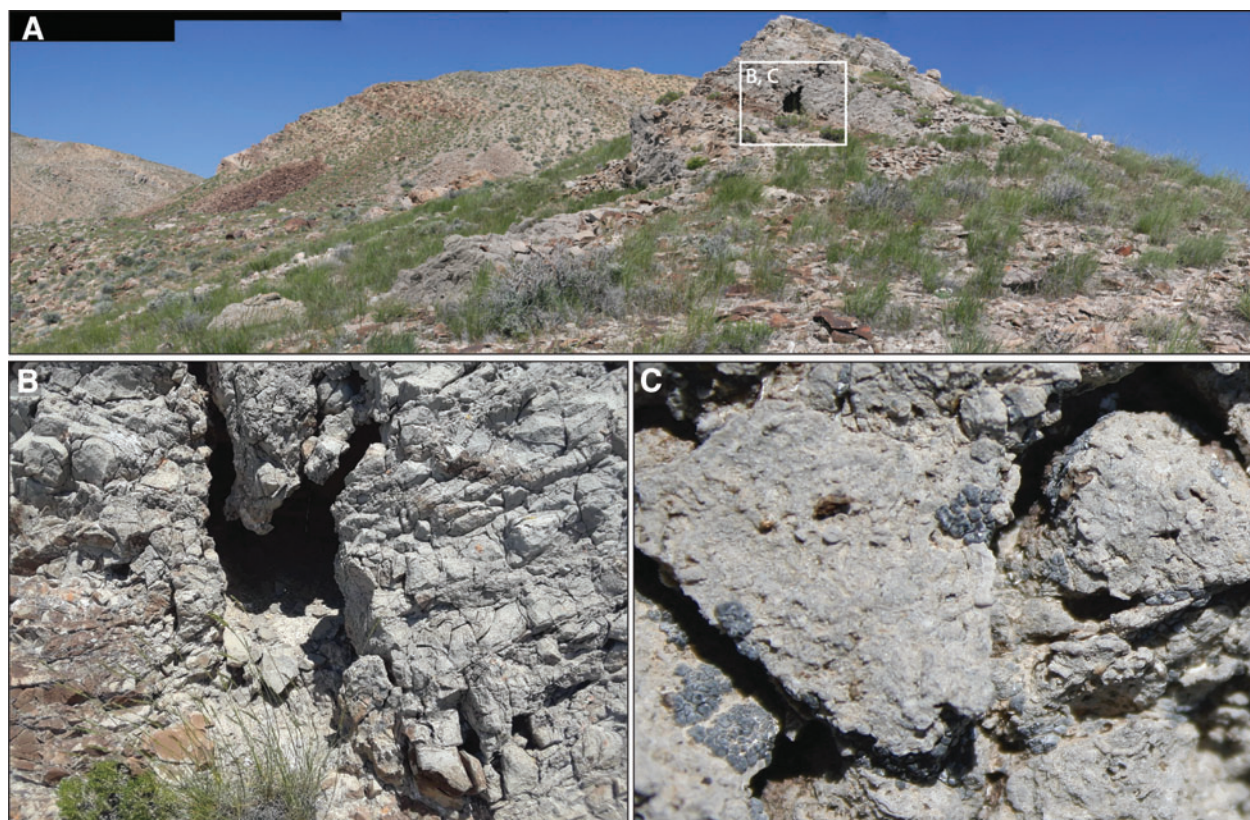
The Strategic team strategy, data and sample acquisition, and interpretations are summarized in Table 3 and described in detail in the Supplementary Material. The team used an estimated 55 sols for observations, incurring no extra sol cost for adding observations.

The Strategic team felt substantial pressure from having to reorganize from its strategic planning. Because of the small size of the four ROIs, the team chose not to use an initial reconnaissance pass to reconnoiter the site before choosing where to use more resource-intensive interrogation.

### 5.2. Tactical team activities

Although the Tactical team also began with identifying potential data collection sites and sample targets, the team did not make a firm sampling plan; instead, it planned to make sampling decisions based on data acquired during the traverse. Using the prefieldwork analysis (Section 4.1; Figs. 2 and 3), the Tactical team identified potential opportunities to sample bedrock and fluvial-lacustrine material





**FIG. 5.** Strategic team assessment of ROI B. (A) Mosaic of accessible region in ROI B. Box is an ~5-m-long axis. (B) Workspace mosaic of karst-like feature within gray dolomitic lithologies. Cave is ~1.5 m high. (C) Close-approach image of karst breccia showing discrete clasts of host limestone that have been secondarily cemented by later diagenetic fluids. Image is 10  $\mu\text{m}$ /pixel.

throughout the campaign areas. For this team, ROI A, with hypothesized exposures of layered bedrock units, and ROI D, with likely lacustrine deposits, initially appeared most promising for sample acquisition. ROIs A and B were hypothesized to contain bedrock exposures; therefore, the team chose to focus analysis on the dark-toned, erosionally resistant bedrock and, secondarily, to acquire context for the light-toned material that was inferred to represent regional fluvio-lacustrine deposition. ROI C provided likely targets for evaluating the light-toned material [area (3) of Fig. 3C].

The Tactical team, during strategic planning, envisioned acquiring 1–2 samples in ROI A; 0–1 sample in ROI B; 1–2 samples in ROI C; and 2–3 samples in ROI D. The Tactical team also chose not to utilize a reconnaissance pass strategy to acquire data, believing that the ROIs were small enough that the time expenditure would be less efficient than a single pass through the terrain.

During the prefieldwork assessment, the Tactical team identified geomorphic evidence of layered sedimentary bedrock and evidence for a younger fluvio-lacustrine environment, marked by linear features of bright material that might have been shorelines. Upon arrival at the ROI, the hypothesized shoreline features were difficult to see and lacked clear geologic units that could be sampled. Bedrock lithologies, however, were well-suited to analysis and sampling (Fig. 7).

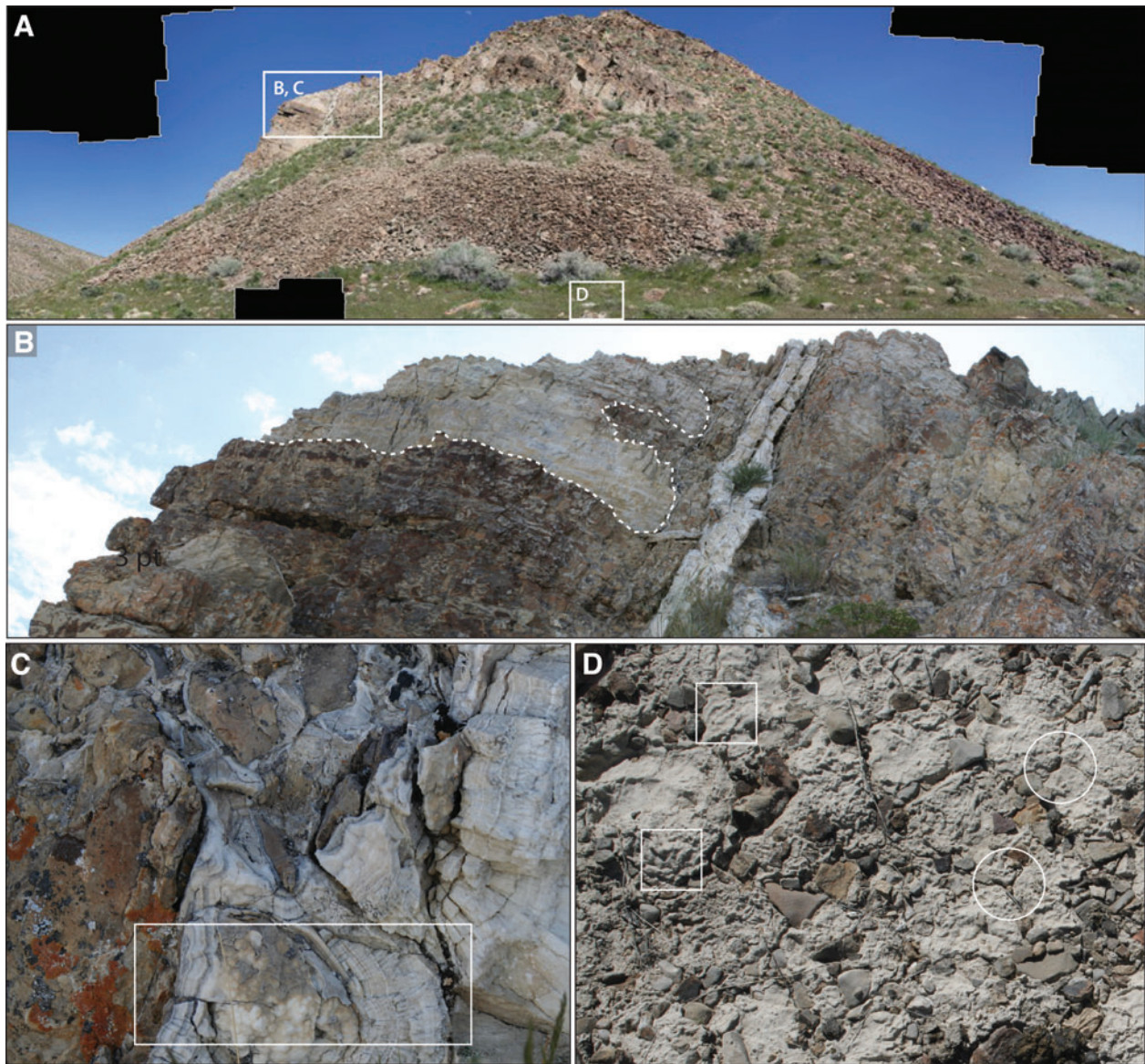
Contact science data were particularly important in choosing which areas to sample, as they provided discriminatory

clues to biosignature retention potential (*e.g.*, grain size below sand; clay compositions; Fig. 8). The team chose to focus sampling efforts on primary lithologies with a higher biosignature potential and, as a result, acquired a sample set that was representative, although not inclusive, of the full range of diagenetic events from this ROI, that is, the dolomitic mudstone bedrock of ROI A (although not the chert.); jarosite-bearing dolomite in ROI B, as representing a potentially distinct hydrologic regime and thus a second interval of habitability; and for ROI C, carbonate bedrock.

Similar to the Strategic team, the Tactical team struggled with how to interrogate ROI D. Ultimately, the Tactical team chose to sample a silty potentially clay-bearing layer on the basin floor as having the highest potential for retaining biosignatures and organic matter, and rimming material as a potential shoreline deposit.

The Tactical team interpreted the field area to contain at least three temporally separate intervals of water activity and potential habitability. The first resulted in deposition of the chert-carbonate bedrock, which was interpreted to record aqueous deposition of carbonate rock. One or more episodes of secondary dissolution and mineral formation followed, resulting in the formation of chert nodules, dissolution to form karst features, cementation of karst, and calcite vein formation. Finally, the youngest aqueous environment, best observed in ROI D, included the development of a standing body of water with shoreline deposits and a clear inlet channel. Tactical team strategy, data, sample acquisition,





**FIG. 6.** Strategic team assessment of ROI C. (A) Mosaic upon entering ROI C. Box labeled (B, C) is a 5-m-long axis. (B) Postbump mosaic of region associated with a prominent vein. Host lithologies retain similar characteristics of alternating brown and gray dolomite with chert nodules; pattern of coloration here suggests that color may reflect differential porosity and alteration by late diagenetic, subsurface fluids (dotted line; vein is ~2 m wide). (C) Close-approach image of vein edge shows discrete clasts of host dolomite that have been secondarily cemented by fracture-filling cement; multiple changes in color within vein-filling cement (rectangle) suggest potential to reconstruct evolution of fluids during precipitation of fracture fill. Image is 10  $\mu\text{m}/\text{pixel}$ . (D) Workspace image of lacustrine deposits, as inferred from orbit. Substrate shows mudcracks (circles) with near  $120^\circ$  intersections that suggest multiple wet-dry events, and wrinkle structures (squares) that suggest sediment cohesion by clay or organic constituents. Unfortunately, this material was deemed not able to be sampled. Squares are ~20 cm across.

and interpretations are summarized in Table 4 and described in detail in the Supplementary Material. The Tactical team used an estimated 56 sols to complete its observations.

## 6. Science Return and Sol Cost

The ultimate goal of this field test was to better understand how a biologically relevant context or samples may be missed when highly constrained time and sampling goals must be met.

Deviating from the preplanned sol-by-sol traverse as little as possible required the Strategic team to weigh the anticipated benefits of an unplanned stop versus the sol cost under far more constrained conditions than those set by previous missions. In this study, we first compare the number of sols each team used to acquire both data and samples. We then qualitatively compare the science results that each team derived from the acquired data by assessing the key geologic features that each team identified and interpreted. The baseline of potential key features and interpretations that



TABLE 3. STRATEGIC TEAM DATA AND INTERPRETATIONS

	<i>Initial plan</i>	<i>Sample strategy</i>	<i>Acquired plan</i>	<i>Key observations and interpretation</i>	<i>Final sample(s)</i>
ROI A	A1: 180° mosaic, clast survey. Option: bump for remote mineralogy, possible contact science. A2: 270° mosaic, clast survey. Option: bump for remote mineralogy, possible contact science. A3: workspace and starboard mosaic, remote mineralogy.	At workspace A2 the decision needs to be made to <i>either</i> sample erosionally resistant bedrock or potential fine-grained material exposed at A3.	A1: Panoramic mosaic, clast survey. A2: Mosaic of cliff face, remote and contact science. Revised A2: sample collection, additional targeted remote, and contact science for context.	(1) Finely layered, inferred lacustrine deposits not exposed in this region. (2) Erosionally resistant bedrock has dips suggesting isolated bedrock block; compelling strata. (3) Remote mineralogy and contact science suggest that bedrock consists of both detrital (siliciclastic) and chemical (carbonate) sediments; with evidence of multiple stages of early to late diagenetic fluid flow ( <i>i.e.</i> , occurrence of chert and dolomite in carbonate phases).	Sample: Dolomite-hosted chert nodule.
ROI B	B1: 180° mosaic, clast survey, untargeted remote mineralogy. B2: workspace and starboard mosaic, untargeted mineralogy. B3: 180°–270° mosaic, untargeted remote mineralogy in workspace. Option: bump for contact science.	Potential for two samples at locality B3, including light-toned layered bedrock, and darker toned materials that appear to cross-cut light-toned strata.	B1: Not traversable. B2: Panoramic mosaic, workspace imaging. Revised B3: Workspace imaging, contact science, sample collection.	(1) Light-toned layered bedrock is not exposed in this region; clast survey suggests that light-toned material consists of nonlithified clay-sized material that encompasses clasts derived from local erosionally resistant bedrock. (2) Proposed fracture was illusion based on topographic expression and shadows. (3) Workspace imaging of cave-like feature reveals extensive karst horizon ( <i>i.e.</i> , dissolution of primary lithologies), collapse breccia, and cementing agent. (4) Red staining associated with specifically grainer lithologies, possible contribution of detrital siliciclastic sand to carbonate lithology.	Sample: Dolomitic karst breccia.
ROI C	C1: 180° mosaic, clast survey, untargeted remote mineralogy, possible contact science. C2: 270° mosaic, untargeted remote mineralogy, possible contact science. C3: 180° mosaic, targeted remote mineralogy in the workspace.	Highest priority for sampling is the light-toned and overlying dark-toned strata within a bump or two of locality C2.	C1: Panoramic mosaic, clast survey, contact science. C2: Context mosaic, remote and contact science. C3: Cliff-face mosaic. Revised C3: additional targeted remote and contact science for context. C4: Context imaging (unsuccessful).	(1) Light-toned layered bedrock is not exposed in this region; clast survey suggests that light-toned material consists of nonlithified clay-sized material that encompasses clasts derived from local erosionally resistant bedrock. Fines are mudcracked and show folded surface textures consistent with organic content. (2) Dark-toned material is all float from erosional-resistant lithologies. (3) Cliffside records carbonate lithologies with more or less detrital siliciclastic input and diagenetic chert, indicating subtle variation in depositional condition. (4) Prominent vein not visible in orbital images; vein-fill shows evidence for multiple stages of postdepositional fluid flow.	Sample: Quartz and carbonate-bearing soil crust. Sample: Calcite vein.

(continued)

TABLE 3. (CONTINUED)

	Initial plan	Sample strategy	Acquired plan	Key observations and interpretation	Final sample(s)
ROI D	D1: 270° mosaic, clast survey. Option: targeted remote mineralogy of pale and darker materials, possible contact science. D2: workspace mosaic, mosaic toward channel. Option: targeted remote mineralogy, possible contact science. D3: targeted remote mineralogy, possible contact science.	Locality considered highest priority for sampling. Potential samples of pale and darker materials in distributary region, of bounding ridges, potentially uphill exposed bedrock.	D1: Unsuccessful downlink of regional mosaic. Revised D1: skip locality. Revised D2: skip locality. Revised D3: Context imaging, contact science, sample collection.	(1) Strategically inferred, high-priority samples sites of D1 and D2 show no exposure in ground operations. (2). Almost no indication of exposed outcrop, except alongside inflow channel. (3) Exposed outcrop consists of fluvial-deltaic deposits from local channel, with layers of varying grain sizes and varying degrees of lithification (likely from surficial fluids). Grains inferred to be derived from high-relief rock units.	ROI did not produce the strategically determined samples. Sample: Three fluvial-deltaic samples with different grain compositions and cementation.

ROI = region of interest.

could be identified was not set by the current state of knowledge of the area, as those data were not provided to any GHOST team members, but by a combination of the prefield assessment of remote imaging and ground-based observations of the Tiger team during 2 days of field work.

This provides a more direct understanding of what a remote science team might reasonably expect to accomplish when interrogating a site for the first time, as is nearly always the case on Mars.

6.1. Sol cost

The sols spent by each team to interrogate the field site are recorded in Table 5, which shows that the Tactical team used 56 sols to complete the four campaigns, while the Strategic team used 55 sols—a difference of <1%. As was instructed at the start of the test, the time line for the mission was tightly constrained, and both teams remained within the estimated 60 sols expected for the test. Although the Strategic team dramatically changed the focus of its observations from its strategic plan, it worked to remain within the tight sol path defined by the test. The Tactical team also stayed within the sol constraints defined by the test, although it voiced a greater level of comfort with the constrained path.

The Tactical team felt confident that it could produce robust context and samples appropriate to science goals, even if it did not achieve an equivalent level of science in each of the four ROIs. By contrast, the Strategic team felt constant pressure to keep moving forward to the final ROI, which had been strategically determined to be the highest priority. As a result, the Strategic team ended up choosing a sampling site using only 2 sols' worth of data. Whereas the likelihood of making such a high-resource decision so rapidly is low on Mars, making such a decision with spotty contextual data is a common necessity in rover traverse planning, and the Strategic methodology, in particular, depends upon science teams preplanning sols days to weeks before *in situ* data are available.

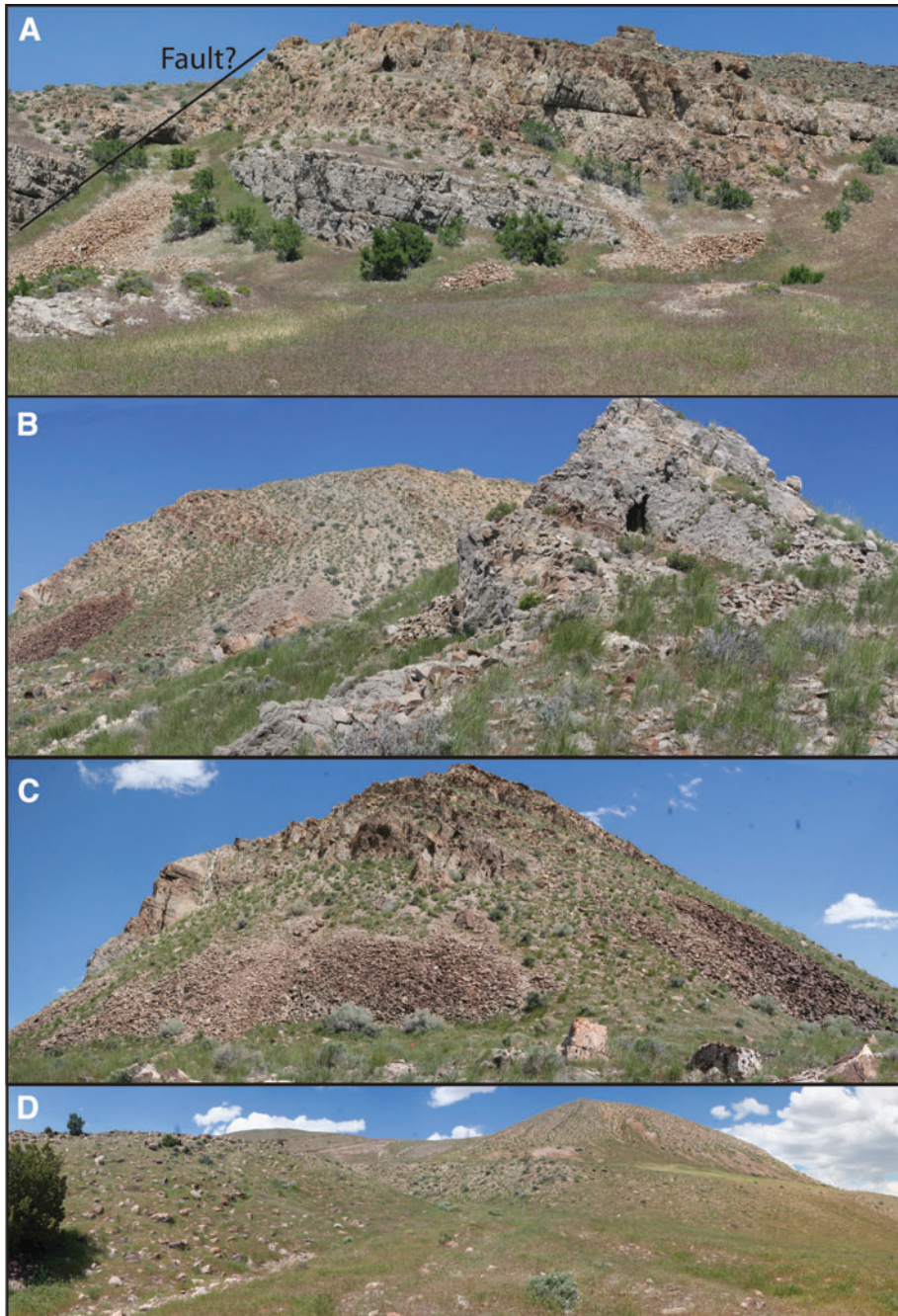
In short, the results in this study suggest that the Strategic planning scenario yields negligible sol savings over an equivalent Tactical planning scenario, while decreasing the confidence in resulting interpretations and sample selection.

6.2. Science return

Table 6 shows the science results reported by the Strategic and Tactical teams, including feature identification and interpretation, compared with the key features assessed by the Tiger team.

Both the Strategic and Tactical teams were able to identify the geologic features key to interpreting the nature of the site as recording a sequence of lake advances and retreats. Interestingly, however, these interpretations, which were determined through the prefield analysis of orbital data, were not translated—by either team—to the ground-based analysis. As a result, both the Strategic and Tactical teams focused their efforts on the local bedrock outcrops. Although this was productive in terms of science output, it was a very different outcome than what was initially planned for the test. The Strategic and Tactical teams found that the data key to decision-making for follow-on sols included





**FIG. 7.** Tactically responsive team panorama mosaics for each campaign area. (A) Mosaic of region upon entering ROI A; box is  $\sim 5 \text{ m}^2$ . (B) Mosaic of ROI B from west edge, with cave highlighted (box is  $\sim 3 \text{ m}^2$ ). (C) Mosaic of ROI C from west edge of the area showing dark-colored patches of debris on the flanks of the hillside (in box,  $\sim 2 \text{ m}^2$ ). (D) Panorama looking northward from the basin in ROI D (box at gully rim is  $\sim 2 \text{ m}^2$ ).

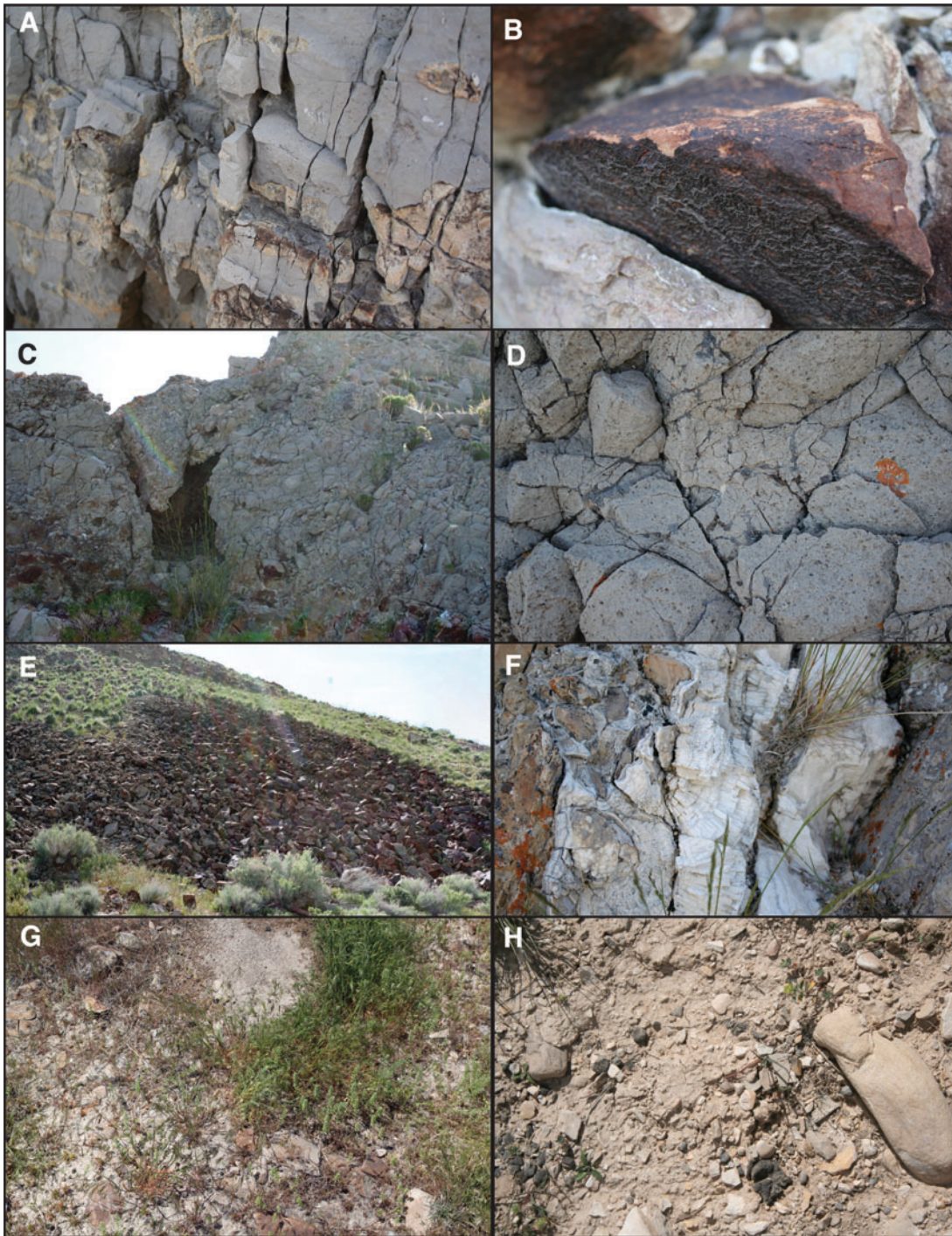
cm-scale and  $\mu\text{m}$ -scale imaging (*e.g.*, for characterizing layers and grain characteristics), and confirmation (or not) of orbitally derived geochemistry.

This dramatic difference in science focus was not apparent from the Tiger team results. Using traditional field geology methods, unrestricted by rover traversability and the time-limited constraints faced by the other teams, field assessment by the Tiger team resulted in observations of local bedrock that were nearly identical to those of the Strategic and Tactical teams. All groups recognized bedrock outcrops as consisting primarily of chert bearing dolomite with minor sandstone interbeds, as well as the presence of regional fractures which, at times, for instance in ROI C, contain mineralized fill. As was determined by the Strategic

and Tactical teams, the Tiger team noted that these features likely record an extended geologic history of aqueous deposition, diagenesis, and secondary alteration.

The Tiger team, however, was also able to resolve a range of geologic materials associated with the geomorphic evidence for late-stage lacustrine deposition. This included most prominently the presence of tufa deposits in ROI B. Note that there is considerable controversy regarding the nomenclature for freshwater/continental/nonmarine carbonates. In this work, we refer to the carbonate coating bedrock as tufa, which is consistent with historic documentation, including the original descriptions of this region by King (1878) and Gilbert (1890) and the use of this term in the literature specifically associated with the Provo shoreline (*e.g.*, Felton *et al.*, 2003).





**FIG. 8.** Context and hand-lens scale images acquired by the tactically responsive team. **(A)** Hand-lens scale image chert nodule-hosting dolostone in ROI A (scale bar is 500  $\mu\text{m}$ ). **(B)** Hypothesized sandstone beds within dolostone in ROI A (scale bar is 500  $\mu\text{m}$ ). **(C)** Cave feature present in bedrock (cave is 1.5 m high); ROI B. **(D)** Hand-lens scale image of dolomitic material near the cave rim in ROI B (scale bar is 500  $\mu\text{m}$ ). **(E)** Dark patch of debris in ROI C (scale bar is 30  $\text{cm}^2$ ). **(F)** Calcite vein in ROI C (scale bar is 10 cm). **(G)** Context image of basin deposits in ROI D (scale bar is 20 cm). **(H)** Context image of light-colored material forming the western shoreline in ROI D (scale bar is 20 cm).

Most studies propose a biological role of algae or photosynthesizing cyanobacteria in the formation of tufa, while others suggest that degassing associated with wave action may be sufficient to drive carbonate precipitation (Felton *et al.*, 2003; Capezzuoli *et al.*, 2013). Despite the uncertainty in whether tufa formation was driven by biotic or

abiotic processes, the association of tufa with biological processes resulted in this facies being classified by the Tiger team as the highest priority for sampling. In comparison, the Strategic team prioritized ROI D, and the Tactical team prioritized ROI D and the layered units in ROI A, without identifying the tufa deposit.



TABLE 4. TACTICAL TEAM DATA AND INTERPRETATIONS

	<i>Initial plan</i>	<i>Sample strategy</i>	<i>Acquired plan</i>	<i>Key observations and interpretation</i>	<i>Final sample(s)</i>
ROI A	180° Panorama from southwest (SW) corner. Hypothesized bedrock (1): workspace mosaic, remote mineralogy, possible contact science. Hypothesized fluvial material (2) in the northeast quadrant: workspace mosaic, remote mineralogy, possible contact science.	Decide whether bedrock should be sampled.	Panoramic mosaic from SW. Southern cliff face mosaic, remote and contact science. Second cliff face mosaic, remote and contact science. Hypothesized fluvial material not evident in northeast (NE) quadrant.	(1) Finely layered, inferred lacustrine deposits not exposed in this region. (2) Erosionally resistant bedrock has dips suggesting isolated bedrock block; compelling strata. (3) Remote mineralogy and contact science suggest that bedrock consists of both detrital (siliciclastic) and chemical (carbonate) sediments; with evidence of multiple stages of early-to-late diagenetic fluid flow ( <i>i.e.</i> , occurrence of chert and dolomite in carbonate phases).	Sample: Fine-grained dolomite from middle bedrock unit
ROI B	180° Panorama from west (W) corner, clast survey, remote mineralogy. Workspace mosaic of east-west (E-W) trending lineation (2) and light-colored resistant material (1) Possible remote mineralogy and contact science.	Potential for samples of light-toned layered bedrock and darker toned materials that appear to cross-cut light-toned strata.	180° Panorama from SW corner Clast survey and workspace mosaic of dipping bedrock. Contact science of karst feature in bedrock.	(1) Light-toned layered bedrock is not exposed in this region; clast survey suggests that light-toned material consists of nonlithified clay-sized material that encompasses clasts derived from local erosionally resistant bedrock. (2) No good targets to sample or investigate mechanisms of fluvial transport. (3) North-dipping fault surface is responsible for E-W trending lineation at (2). (4) Workspace imaging of cave-like feature reveals extensive karst ( <i>i.e.</i> , dissolution of primary lithologies) and cementation. Host bedrock is bedded and shows variation in chert content, bed thickness, and detrital siliciclastic content, indicating variations in depositional conditions.	Sample: Karst-associated dolomite.
ROI C	180° Mosaic, clast survey, untargted remote mineralogy, possible contact science to investigate bright material at (3). Vertical mosaic to assess circular dark patches at (2). 270° Mosaic, untargted remote mineralogy, possible contact science of high-relief bedrock (1).	Highest priority for sampling is the light-toned and overlying dark-toned strata at (2) and (3).	Panoramic mosaic from W, clast survey. Targeted mastcam image of (2); contact science. Cliff-face mosaic. Contact science to investigate vein and dissolution features at (1).	(1) Clast survey suggests that light-toned material consists of nonlithified clay-sized material that encompasses clasts derived from local erosionally resistant bedrock. (2) Dark-toned material is locally derived bedrock lithologies. Apparent rims of light-colored material could not be confirmed. (3) Prominent vein cuts dolomitic bedrock lithology; calcite vein-fill shows evidence for multiple stages of postdepositional fluid flow. (4) Evidence for multiple stages of water activity—perhaps multiple intervals of habitability.	Sample: Bedded dolomite Sample: light-colored muddy material at the western edge of area
ROI D	270° Mosaic, clast survey. Targted remote mineralogy of pale and darker materials, possible contact science. 180° Mosaic facing up-channel. Targted remote mineralogy. Workspace imaging on light-colored material; targted remote mineralogy, possible contact science.	High priority for sampling hypothesized lacustrine material. Potential samples of pale and darker materials in tributary region of bounding ridges.	Clast survey, workspace mosaic, scuff in basinal area (2). Remote mineralogy, contact science. Vertical mosaic: facing up-gully (north). Belly mosaic of bright material (1) rimming the basin. Remote mineralogy, contact science.	(1) Targted remote mineralogy and contact science reveal weathering products (clay minerals, dolomite) forming a fine-grained layer on the basin floor. (2) Evidence for intermittent high-energy water flow into the basin. Grains inferred to be derived from high-relief rock units. (3) Bright material along basin rim also consists of weathering products (clay minerals, calcite); clasts in mud matrix. Interpreted as shoreline deposit.	Sample: basin deposit: silty material Sample: rimming material; clay+calcite

TABLE 5. COMPARISON OF SOL COSTS

	<i>ROI A</i>	<i>ROI B</i>	<i>ROI C</i>	<i>ROI D</i>	<i>Total sols</i>
Strategically constrained team					
Driving					
Drive plus imaging	2	1	2	2	7
Bump or bump plus imaging	1	0	1	1	3
Contact science					
Drilling (5 sols required)	5	5	10	15	35
Contact science	1	1	2	1	5
Remote imaging					
Panorama	1	1	1	2	5
Strategically constrained team total sols					55
Tactically responsive team					
Driving					
Drive plus imaging	1	1	3	0	5
Bump or bump plus imaging	3	2	4	4	14
Contact science					
Drilling (5 sols required)	5	5	15	0	25
Contact science	2	2	3	2	9
Remote imaging					
Panorama	1	1	1	1	4
Tactically responsive team total sols					56

Drive sols between campaign locations are not book kept, as the path between each campaign (and thus the sol cost) was identical for both teams.

In ROI B, the Tiger team recognized a decameter-wide, gravel-covered topographic platform, inferred to represent a primary, wave-eroded shoreline. Along this platform, the Tiger team recorded evidence of carbonate phases that both cemented and draped eroded and brecciated bedrock exposures. Both the Strategic and Tactical teams recognized brecciation of bedrock in these regions, which they identified as potential karst dissolution, consistent with subaerial exposure at a lacustrine shoreline, but did not recognize draping elements of the tufa, which were exposed primarily along the uphill side of the ROI.

In a postexercise analysis, it was determined that the Strategic team did, in fact, image the tufa deposits (Fig. 9), but the imaging was completed as part of a routine 360° mosaic upon leaving the ROI and was not received until after the team left the ROI (a common occurrence when time is constrained; less critical data will be placed in a

lower priority in the downlink packet, and would arrive subsequent to the planning cycle for the next sol). As a result, timely analysis of the mosaic did not occur.

Similarly, in ROI D, the Tiger team recognized both the downhill bounding gravel bar deposit associated with a prominent shoreline deposit, and an uphill inlet channel. As with the prefield assessment by the Strategic and Tactical teams, these features were inferred to represent the past presence of an ephemeral to perennial lake. Analysis by the Tiger team suggests that this lake (or potentially lake-side lagoon) may have been 5–8 m deep. Deltaic deposits were identified by gravel-rich layers based on the rounding of pebbles and gravel. Whereas this site was considered high priority for sampling by both teams due to its likely lacustrine history, the Tiger team did not identify any obvious locations where the evidence for life would be easy to sample.

TABLE 6. SCIENCE RETURN ASSESSMENT USING TIGER TEAM ANALYSIS AS A BASELINE FOR DETERMINING KEY FEATURES OF INTEREST

Key features	<i>Strategic</i>	<i>Tactical</i>	Interpretation	<i>Strategic</i>	<i>Tactical</i>
	Cyclic carbonate-rich layers	Y		Y	Marine environment
Chert nodules	Y	Y	Diagenetic phase	Y	Y
Chert nodules	Y	Y	Biomarker potential	Y	Y
Tufa	N	N	Potential biotic activity	N	N
Shoreline	PFH	PFH	Subaerial exposure	Y	Y
Cross-cutting bedrock fractures	Y	Y	Later diagenetic feature	Y	Y
Mineralized fracture fill	Y	Y	Subsurface fluid activity	Y	Y
Lake and inlet deposits	Y	Y	Late timing	PFH	PFH
Appropriate samples	Y	Y	Sample context	Y	Y

If the operational team identified the feature or correctly interpreted it, that category is marked with a Y; if not, the category is marked N. If the correct identification or interpretation was one hypothesized during prefield assessment, it is marked with PFH; future data mining or discussion could potentially have led to the further interpretation of test data.





**FIG. 9.** Tufa deposits in ROI B. Both strategically constrained and tactically responsive teams identified potential karst-related features on the downhill side (black arrow) of the main bedrock exposure (asterisk). Notably, neither team recognized outcrop coating associated with precipitation of tufa (white arrows). Abundant fracturing of the tufa suggests primary deposition as a hydrated carbonate phase. Protuberance under the black arrow is  $\sim 2$  m above the grassy surface.

In addition, the high stratigraphic placement of the lake deposit (above the more ancient deposits) indicated that the facies was likely not within the time frame of interest.

Notably, ROI D was considered the highest priority for sampling by both teams, while the feature deemed by the Tiger team to be of highest priority for sampling was not recognized by either operational team. As earlier noted, neither the Strategic nor the Tactical team chose to do a reconnaissance pass before choosing sample sites; in both cases, the teams felt that the small size of each ROI, coupled with the tight time constraints, would make the reconnaissance pass unnecessary. However, in this case, the diagnostic features were only visible if the team moved to a different viewing angle than the initial approach to the ROI; if either team had chosen to do a reconnaissance pass of the site first, the tufa deposit might have been identified.

The Strategic team did, in fact, capture these deposits in images on the way to the next ROI, but these images were not assessed at a point where it was possible to make a decision to revisit this unique lithology without committing to completely break the strategic timeline. Thus, although both teams noted that, in the interests of time, they did not conduct a reconnaissance pass that included regional imaging from different viewpoints, both teams confirmed that they would have actually saved time by using the reconnaissance pass methodology. The teams believed this was due to the nature of rover-driven science, which consists of a series of very narrow views separated by broad swaths of uninterrogated landscape; the reconnaissance pass allows more of these gaps to be filled by data.

The conclusion is that the reconnaissance pass is an essential tool for science operations, especially in cases where the sol path is highly constrained. In such cases, the reconnaissance pass provides contextual data systematically, without the time pressure of having to make rapid decisions based on an incomplete picture of the site.

The other geologic feature that was not identified or interpreted correctly by either team was the series of shorelines of the former Lake Bonneville. The reason for this is likely that the shorelines, although potentially viewable

from orbital images, were subtle when seen from the ground and were best inferred by having multiple different wide-field views of the region. The Tiger team, in fact, did not identify a clear shoreline until ROI B, and only after spanning several scales at both the ROI A and B locations. With respect to identifying similar features on Mars, we note that shorelines are subtle macrofeatures and potentially extensive across the landscape, meaning that consistent observations made over time could help build the evidence needed to identify and test the hypotheses regarding possible shorelines.

Finally, the sample acquisition profile for each team is similar for ROIs A to C; both teams acquired a single sample at each of the first two ROIs and multiple samples at ROI C. Sampling, however, reflected different stated goals, in which the Strategic team focused on understanding the nature and timing of fluid events within the overall geologic history of the region, and the Tactical team focused on understanding the range of dolomite/chert deposits and how they were deposited. Both teams, however, focused their efforts on acquiring samples that had a high likelihood of retaining biosignatures, and so, both teams acquired multiple dolomite/chert samples.

Additional samples were chosen to contextualize the geologic story in space and time.

## 7. Discussion

In terms of sol cost, although sol conservation drove every planning period of the Strategic team, there was in the end a  $<1\%$  difference in sol cost between the two methods or  $<10$  sols over a 1000-sol period. Although the driving rationale each team chose was different, science return from each method, as judged by correct interpretations, was similar. In terms of the appropriateness and science value of sample caches, both caches provided important information about biosignatures, but neither team focused its efforts on the highest priority site identified by the Tiger team. We note, however, that only the Strategic team noted that it had purposely saved sample slots for the last ROI (ROI D),

and thus it judged the sample collection to be nonoptimal. Three of the Strategic team's samples were acquired at ROI D, while no samples were acquired by the Tactical team in this location.

The Strategic team chose to collect these samples as a way to acquire information about the site when few sols were left, the rationale being that the samples could be assessed after the fact. This difference is thus a product of sol constraints, rather than an indication of qualitative differences in hypothesis development. The preplanning experience (with orbital data) was straightforward and valuable for team building, but the first day of surface operations revealed surprises that necessitated a restructuring of the plan. This indicates that preplanning is a crucial first step that is nevertheless limited in its long-term usefulness in many real-world cases. For this case, the inescapable conclusion is that the additional effort the Strategic team put into preplanning sol paths did not result in a sol cost savings, improved science return, or optimal biologically relevant sample collection and was thus not an efficient investment.

We confirm the previous conclusions (Yingst *et al.*, 2016, 2018, 2020) that any methodology is more likely to result in missed important science without utilizing the reconnaissance pass. Decisions regarding whether to further interrogate or sample a target are of necessity made regularly without an optimal context in a rover operation scenario. In addition, time pressures often necessitate driving away from a site before all data from that site have been downlinked to Earth, meaning that data will not be down in time to inform the next sol's plan.

It is thus important to acquire observations designed to build up contextual information, while avoiding the situation where downlink of that data is too late to feed into the next plan. This is crucial for reasonable hypotheses to be posed and tested; it allows more localized or discrete features to be identified and provides more opportunities for subtle clues to be captured, strategically if not tactically.

In those situations where a reconnaissance pass is impractical or impossible, systematic observations—a sequence of repeatable, resource-light remote observations—may provide an additional context that increases confidence in identifications and interpretations. That confidence is crucial when approaching a tactical decision point to remain on the planned sol path or change it. This is true not only in support of hypothesis development, but also in virtual returns to sites long passed, to evaluate them with subsequently acquired data and knowledge. Because detailed observations are only made in limited locations with a rover, the acquisition of systematic images provides not only context but also insurance against returning to a location by providing data in-hand to investigate new hypotheses without the need to drive back.

Such observations might include placing reconnaissance stations in areas where orbital data may not reveal obvious targets of interest or acquiring 360° color imaging regularly to determine whether additional data collection is warranted. Such imaging can be planned to minimize data volume resources and can allow scientists to survey the terrain and evaluate any rock texture changes.

Facies variations can be subtle and below the resolution limit of orbital image and spectral data. For example, Stations A3 and C1 both contained geologically interesting

targets that were not recognizable in orbital data. Although Station A3 was identified as of interest in the strategic plan, this location was not well evaluated during the test; the Tactical team reviewed the data only after the rover exited the area and the Strategic team descoped the site to save time. The light-toned vein in ROI C was evident in imaging at C1, but not detectable in orbital data. This vein contained the only clues to the secondary fluid event that formed it, but it would have been missed without panoramic ground-based imaging.

Both teams used orbital data extensively to design the hypotheses and plan traverses that would test those hypotheses. In addition, *in situ* confirmation and detailed expansion of expected geochemistry were an important discriminator in the Strategic team's decision to alter its original strategy. However, once this initial assessment was completed, the team found that the geochemistry was well predicted by the revised hypotheses, and it was not used again to refine or change the planned traverse. The Tactical team results were similar; they determined from initial data in each ROI that the geochemistry was as expected. And while it was used to confirm or refine hypotheses, it was not used to alter the course.

The point-to-point nature of data collected remotely means that often the dots do not connect fully, or there are not enough dots to resolve a diagnostic feature. As noted above, this problem is exacerbated by downlink limitations, such that many images of some locations are downlinked after departure. The ability to rapidly assess from micro- to macro- to mesoscale as needed was an ability the Tiger team used to great effect to understand how regional features connected to the local geology; it was this that allowed the subtle shorelines to be identified and interpreted.

The rover-based teams, on the contrary, were unable to make this jump, and thus, the connection between orbital data originally interpreted as rock layers and the discontinuous expressions of these features on the ground was missed. The rover teams also struggled to determine the nature of the fractures at ROI C as tectonic or evidence of a larger intrusive structure such as a dome; by comparison, the relationship of the large fracture with the surrounding layers made the tectonic origin clear to the Tiger team. The use of three-dimensional immersion tools, such as hololenses used to project data from the Gale crater into a virtual reality environment (Abercrombie *et al.*, 2019), would give remote geologists a way to span scales as terrestrial field geologists can do. This would save time and improve science output.

Several lessons were learned during this field excursion:

- (1) Strategic science planning must account for terrain. Site characterization is particularly difficult when climbing upslope due to the limited viewshed. In addition, while rover traverses nearly always head uphill, certain features, such as tufa deposits, will often be best viewed by looking back along the traverse route.

We note also that the results of this test indicate that the spacing of panoramas is highly dependent on the relief in the region and the frequency of other image data acquisitions (*e.g.*, drive direction mosaics). Where there are topographic obstructions, imaging opportunities should be increased to less than the distance between obstructing topographic features. In flat regions, >50 m spacing may be

sufficient to assess the textural variability; note that the MSL adopted an  $\sim 300$  m spacing for  $360^\circ$  panoramas for a large portion of the mission (Vasavada *et al.*, 2014). To minimize resources used, data can be prioritized to image outcrops, with care taken to obtain multiple viewing angles (*e.g.*, front and back views).

- (2) Caution must be exercised in planning traverses around chemical signatures at the expense of systematic site characterization. Spectral data tend to provide insight into chemical components to be expected at the site as a whole. It should be noted, however, that minor abundances of certain minerals can dominate spectra but do not always imply an unusual abundance of the mineral itself (*e.g.*, Rampe *et al.*, 2020); this can lead to deviations from the sol path that may not result in furthering mission goals. In this test, it was not a specific technique that was of concern, but the general problem that arose when a signal of interest did not yield the expected mission-specific result *in situ*. In addition, in some cases, a strong spectral signature is more indicative of a relatively dust-free location rather than a lithologic difference from the surrounding areas.

Spectral data provide important insight to the strategic formulation of the geologic hypotheses to be tested *in situ* but should not necessarily be the sole reason for selecting a route in cases similar to the one tested here.

- (3) We recommend that, in a highly constrained timeline, orbital data be considered secondary to *in situ* data when traverse planning, once on the ground. Mars analog tests over several years (Yingst *et al.*, 2015, 2016, 2018) indicate that, when traversability is not a variable, teams rarely deviate more than a few 10s of meters from the traverse planned using orbital data, meaning that teams are relatively skilled at picking out the scientifically richest places to explore *in situ*. The time and effort invested in understanding and utilizing orbital data to preplan traverses are crucial; but once that investment is made, the usefulness of the data for traverse planning is minimal. Thus, we suggest not investing significant time in fine-tuning traverse planning from orbital data; *in situ* data are more valuable.

As a final note, this and previous tests represent complex geology real-world tests, in which it was expected that remote teams would miss important clues to the geologic story. Nevertheless, the GHOST tests consistently reinforce the fact that, although remote geologists will miss some geologic clues, the first-order story tends to be accurately inferred from data gathered through rover-enabled methods, regardless of the method used. It is hoped that the lessons learned here will help future remote science teams continue to refine their methods, and thus miss fewer and fewer of those clues.

### Acknowledgments

The authors wish to gratefully acknowledge the careful work of two anonymous reviewers of this article.

### Author Disclosure Statement

No competing financial interests exist.

### Funding Information

This work was funded by the NASA PSTAR program (grant NNX15AI89G). Additional support was provided by the Utah Geological Survey (UGS).

### Supplementary Material

Supplementary Material

### References

- Abercrombie SP, Menzies A, Abarca HE, *et al.* (2019) Multiplatform immersive visualization of planetary, asteroid and terrestrial analog terrain [abstract 2132]. In *50th Lunar Planetary Science Conference Abstracts*. Lunar and Planetary Institute, Houston, TX.
- Arvidson RE, Squyres SW, Baumgartner ET, *et al.* (2000) FIDO prototype Mars rover field trials, Black Rock Summit, Nevada, as test of the ability of robotic mobility systems to conduct field science. *J Geophys Res* 107:FIDO 2-1–FIDO 2-16.
- Bass DS, Wales RC, and Shalin VL (2005) Choosing Mars time: analysis of the Mars Exploration Rover experience. In *Proceedings of IEEE Aerospace Conference*. IEEE, Big Sky, MT, p 116.
- Black SR and Hynke BM (2018) Characterization of terrestrial hydrothermal alteration products with Mars analog instrumentation: implications for current and future rover investigations. *Icarus* 307:235–259.
- Capezzuoli E, Gandin A, and Pedley M (2013) Decoding tufa and travertine (fresh water carbonates) in the sedimentary record: the state of the art. *Sedimentology* 61:1–21.
- Chen CY and Maloof AC (2017) Revisiting the deformed high shoreline of Lake Bonneville. *Quat Sci Rev* 159:169–189.
- Clark DL and Oviatt CG (2018) Interim geologic map of the Bonneville Salt Flats and east part of the Wendover  $30' \times 60'$  quadrangles. Scale: 1:62,500. In *Open-File Report 692*. Utah Geological Survey, Salt Lake City, UT. Available online at [https://ugspub.nr.utah.gov/publications/open\\_file\\_reports/ofr-692.pdf](https://ugspub.nr.utah.gov/publications/open_file_reports/ofr-692.pdf) (accessed February 11, 2021).
- Clark DL, Oviatt CG, and Dinter DA (2017) Interim geologic map of the Tooele  $30' \times 60'$  quadrangle, Tooele, Salt Lake, and Davis Counties, Utah. In *Open-File Report 669DM*. Utah Geological Survey, Salt Lake City, UT.
- Cohen BA, Lim DSS, Young KE, *et al.* (2015) Pre-mission input requirements to enable successful sample collection by a remote field/EVA team. *J Hum Perf Extrem Environ* 12. DOI: 10.7771/2327-2937.1071.
- Downs RT and Hall-Wallace M (2003) The American Mineralogist Crystal Structure Database. *Am Min* 88:247–250.
- Ehlmann BL, Mustard JF, Fassett CI, *et al.* (2008) Clay minerals in delta deposits and organic preservation potential on Mars. *Nat Geosci* 1:355–358.
- Eppler DB, Adams B, Archer D, *et al.* (2013) Desert Research and Technology Studies (DRATS) 2010 science operations: operational approaches and lessons learned for managing science during human planetary surface missions. *Acta Astronaut* 90:224–241.
- Felton A, Jewell, PW, Chan MA *et al.* (2003) Controls of tufa development in Pleistocene Lake Bonneville, Utah. *J Geol* 114:377–389.
- Fong T, Abercromby A, Bualat MG, *et al.* (2010) Assessment of robotic recon for human exploration of the Moon. *Acta Astronaut* 67:9–10.
- Gilbert GK (1890) Lake Bonneville. U.S. Geological Survey Monograph 1, 438 p.



- Gouge TA, Mohrig D, Cardenas BT, *et al.* (2018) Stratigraphy and paleohydrology of delta channel deposits, Jezero crater, Mars. *Icarus* 301:58–75.
- Graham L, Graff TG, Yingst RA, *et al.* (2015) 2012 Moon Mars analog mission activities on Mauna Kea, Hawai'i. *Adv Space Res* 55:2405–2413.
- Greeley R, Basilevsky AT, Kuzmin RO, *et al.* (1994) Science results from the Marsokhod tests, Amboy lava field, California. Paper presented at the International Planetary Rover Symposium, Moscow, Russia.
- Hintze LF and Kowallis BJ (2021) *Geologic History of Utah*, 2nd ed., Brigham Young University Geology Studies Special Publication 10, Provo, UT, 266 p.
- Horgan BHN, Anderson RB, Dromart G, *et al.* (2020) The mineral diversity of Jezero crater: evidence for possible lacustrine carbonates on Mars. *Icarus* 339:113526.
- King C (1878) Systematic geology: U.S. geological explorations of the fortieth parallel (Vol. 1). Government Printing Office, Washington, DC, 579 p.
- Kokaly RF, Clark RN, Swayze GA, *et al.* (2017) USGS Spectral Library Version 7, U.S. Geological Survey Data Series 1035, U.S. Geological Survey, Reston, VA, 61 p.
- Lee P (2007) Haughton-Mars Project 1997–2007: a decade of Mars analog science and exploration research at Haughton Crater, Devon Island, High Arctic. 2nd International Workshop-Exploring Mars and its Earth Analogues, Trento, Italy.
- Lee P, Braham S, Boucher M, *et al.* (2017) Haughton-Mars Project: 10 years of science operations and exploration systems development at a Moon/Mars analog site on Devon Island, High Arctic [abstract 2426]. In *48th Lunar Planetary Science Conference Abstracts*. Lunar and Planetary Institute, Houston, TX.
- Lim DSS, Abercromby AFJ, Kobs Nawotniak SE, *et al.* (2019) The BASALT Research Program: designing and Developing Mission Elements in Support of Human Scientific Exploration of Mars. *Astrobiology* 19:245–259.
- Milkovich SM, Stack KM, Sun VZ, *et al.* (2022) Balancing predictive and reactive science planning for Mars 2020 perseverance. IEEE Conference. 5–12 March, Bigsby, MT. DOI: 10.1109/AERD53065.2022.
- Milliken R, Grotzinger J, and Thomson B (2010) Paleoclimate of Mars as captured by the stratigraphic record in Gale Crater. *Geophys Res Lett* 37. DOI: 10.1029/2009GL041870.
- Osinski G, Battler M, Caudill CM, *et al.* (2019) The CanMars Sample Return analogue mission. *Planet Space Sci* 166:110–130.
- Oviatt CG and Schroder JF, Eds. (2016) *Lake Bonneville—A Scientific Update*. Elsevier, New York, 696 p.
- Rampe EB, Bristow TF, Morris RV, *et al.* (2020) Mineralogy of Vera Rubin Ridge from the Mars Science Laboratory CheMin instrument. *J Geophys Res* 125. DOI: 10.1029/2019JE006306.
- Reheis MC, Adams KD, Oviatt CG, *et al.* (2014) Pluvial lakes in the Great Basin of the western United States—a view from the outcrop. *Quat Sci Rev* 97:33–57.
- Squyres SW, Arvidson RE, Baumgartner ET, *et al.* (2003) Athena Mars rover science investigation. *J Geophys Res* 3:806.
- Stoker CL, Cabrol NA, Roush TR, *et al.* (2001) The 1999 Marsokhod rover mission simulation at Silver Lake, mission overview, data sets, and summary of results, California. *J Geophys Res* 106:7639–7664.
- Stoker CL, Roush TL, Arvidson RE, *et al.* (2002) Two dogs, new tricks: a two-rover mission simulation using K9 and FIDO at Black Rock Summit, Nevada. *J Geophys Res* 107: FIDO 8-1–FIDO 8-13.
- Vasavada AR, Grotzinger JP, Arvidson RE, *et al.* (2014) Overview of the Mars Science Laboratory mission: Bradbury landing to Yellowknife Bay and beyond. *J Geophys Res* 119: 1134–1161.
- Whittaker W, Bapna D, Maimone MW, *et al.* (1997) Atacama Desert trek: a planetary analog field experiment. In *Proceedings of the 7th International Symposium on Artificial Intelligence, Robotics, and Automation for Space*, Tokyo, Japan.
- Wilson M, Trosper J, and Abilleira F (2017) NASA Mars 2020 landed mission development. Tech. Rep., Jet Propulsion Laboratory, California Institute of Technology. trs.jpl.nasa.gov (accessed December 8, 2022).
- Wray JJ, Murchie SL, Bishop JL, *et al.* (2016) Orbital evidence for more widespread carbonate-bearing rocks on Mars. *J Geophys Res* 121:652–677.
- Yingst RA, Cohen BA, Crumpler L, *et al.* (2014) Testing Mars-inspired operational strategies for semi-autonomous rovers on the Moon: the GeoHeuristic Operational Strategies Test in Alaska. *Acta Astronaut* 99:24–36.
- Yingst RA, Russell P, ten Kate IL, *et al.* (2015) Designing remote operations strategies to optimize science mission goals: lessons learned from the Moon Mars Analog Mission Activities Mauna Kea 2012 field test. *Acta Astronaut* 113:120–131.
- Yingst RA, Berger J, Cohen BA, *et al.* (2016) Determining best practices in reconnoitering sites for habitability potential on Mars using a semi-autonomous rover: a GeoHeuristic Operational Strategies Test. *Acta Astronaut* 132:268–281.
- Yingst RA, Bartley JK, Chidsey TC, Jr., *et al.* (2018) Testing the efficiency of rover science protocols for robotic sample selection: a GeoHeuristic Operational Strategies Test. *Acta Astronaut* 146:300–315.
- Yingst RA, Bartley JK, Chidsey TJ, Jr., *et al.* (2020) Is a linear or a walkabout protocol more efficient when using a rover to choose biologically relevant samples in a small region of interest? *Astrobiology* 20:3.

Address correspondence to:  
R. Aileen Yingst  
10 Julias Way  
Brunswick, ME 04011  
USA

E-mail: yingst@psi.edu

Submitted 11 November 2021

Accepted 2 July 2022

#### Abbreviations Used

DEM	= digital elevation model
GHOST	= GeoHeuristic Operational Strategies Testing
MER	= Mars Exploration Rover
MSL	= Mars Science Laboratory
PDI	= postdrive imaging
PFH	= hypothesized during prefield assessment
ROI	= region of interest
VNIR	= visible to near-infrared
XRD	= X-ray diffraction
XRF	= X-ray fluorescence