nature astronomy

Article

https://doi.org/10.1038/s41550-022-01882-x

Orbit-to-ground framework to decode and predict biosignature patterns in terrestrial analogues

Received: 8 December 2020

Accepted: 16 December 2022

Published online: 06 March 2023

Check for updates

Kimberley Warren-Rhodes[1,2] ✉, Nathalie A. Cabrol[1,2], Michael Phillips[3], Cinthya Tebes-Cayo[4], Freddie Kalaitzis[5], Diego Ayma[4], Cecilia Demergasso[4], Guillermo Chong-Diaz[4], Kevin Lee[6], Nancy Hinman[7], Kevin L. Rhodes[8], Linda Ng Boyle[9], Janice L. Bishop[1,2], Michael H. Hofmann[7], Neil Hutchinson[10], Camila Javiera[4], Jeffrey Moersch[3], Claire Mondro[3], Nora Nofke[11], Victor Parro[12], Connie Rodriguez[4], Pablo Sobron[1,13], Philippe Sarazzin[1,2], David Wettergreen[14], Kris Zacny[15] & the SETI Institute NAI Team*

In the search for biosignatures on Mars, there is an abundance of data from orbiters and rovers to characterize global and regional habitability, but much less information is available at the scales and resolutions of microbial habitats and biosignatures. Understanding whether the distribution of terrestrial biosignatures is characterized by recognizable and predictable patterns could yield signposts to optimize search efforts for life on other terrestrial planets. We advance an adaptable framework that couples statistical ecology with deep learning to recognize and predict biosignature patterns at nested spatial scales in a polyextreme terrestrial environment. Drone flight imagery connected simulated HiRISE data to ground surveys, spectroscopy and biosignature mapping to reveal predictable distributions linked to environmental factors. Artificial intelligence–machine learning models successfully identified geologic features with high probabilities for containing biosignatures at spatial scales relevant to rover-based astrobiology exploration. Targeted approaches augmented by deep learning delivered 56.9–87.5% probabilities of biosignature detection versus <10% for random searches and reduced the physical search space by 85–97%. Libraries of biosignature distributions, detection probabilities, predictive models and search roadmaps for many terrestrial environments will standardize analogue science research, enabling agnostic comparisons at all scales.

In extreme environments, the distribution of biosignatures is tightly controlled by a complex interdependence of geological, physicochemical and biological interactions[1–5]. In such environments, microbial populations often occur in non-random spatial distributions closely tied to their physical settings, key morphologies (substrates, 'habitats') and/or requirements for life, especially water[6–10]. To date, few studies have systematically examined such linkages across integrated spatial scales (Fig. 1) or applied machine learning to test the predictive power

A full list of affiliations appears at the end of the paper. ✉e-mail: krhodes@seti.org

https://doi.org/10.1038/s41550-022-01882-x

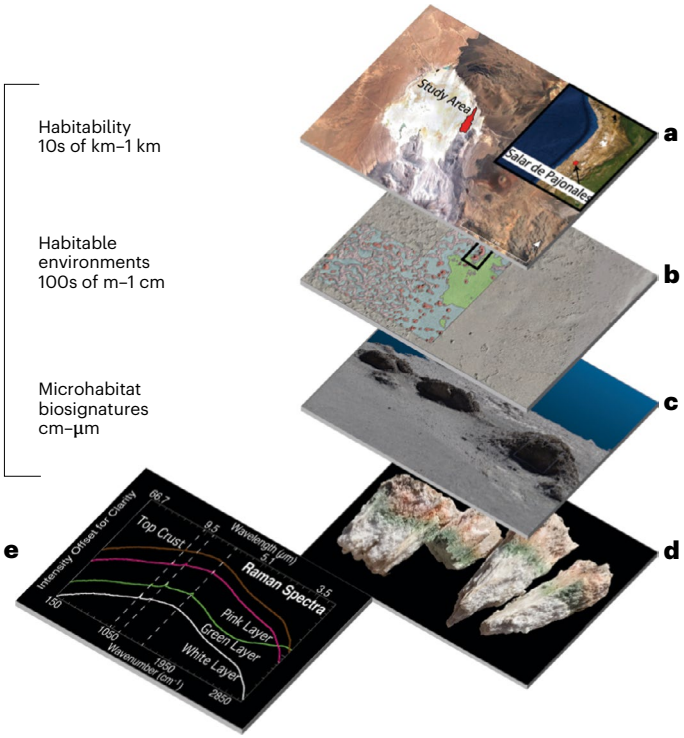


Fig. 1 | Orbit-to-ground scales of investigation. a, Orbital view of Salar de Pajonales, Chile. The study area (8,193 m^2), composed of a field of gypsum domes, is shown in red. **b**, UAV view of the Dome Field in the southern portion of the Pajonales location, the focus of this study, with macrohabitats shown in false colour: red, domes; pink, ridges/polygonal networks; blue, aeolian cover; green, patterned ground. The black box delineates the dome complex for the fine-scale study. **c**, Detail of the fine-scale study area with the three-dimensional representation of dome macrohabitats obtained from drone imaging, comparable to a rover 'Mastcam' scale. **d**,**e**, Coloured bands (biosignatures) of endolithic photosynthetic communities colonizing a gypsum type II crystal microhabitat in the domes (**d**) and Raman spectra of the distinct biosignature layers from **d** (**e**), comparable to 'SuperCam' Raman spectra on the Mars 2020 rover Perseverance.

and probabilities of detecting life at the extremes of habitability[11–13]. In this Article, we present a proof-of-concept study that quantifies such probabilities and constructs predictive models for geologic feature recognition and biosignature detection.

Our process combines statistical microbial landscape ecology[14,15] with remote sensing from unmanned aerial vehicles (UAVs/drones) and machine learning to map and explain (Figs. 2 and 3), model and predict (Figs. 4 and 5) the distributions of biosignatures in a Mars-relevant evaporitic palaeolake setting. Our approach captures microbial spatial distributions and their controls across a continuum of scales and demonstrates that, even amidst near-uniform mineralogy, biosignature patterns are discernible, non-random and predictable. We improve the probability of detecting biosignatures[13] from <10% (for random search) to up to 87.5% by augmenting macro- and microhabitat targeting with deep learning. Simultaneously, we shrink the search space by up to 97%. Ultimately, our goal is to provide a broadly applicable and adaptable methodological framework to standardize terrestrial analogue research. With this approach, the astrobiology community can build a database, akin to spectral libraries, of biosignature distributions, detection probabilities and predictive models to guide and inform planetary exploration[16–19].

Habitat and water control biosignature distributions

We quantified the distribution of photosynthetic communities in Salar de Pajonales (Pajonales), a 104 km^2 ancient lakebed located at 3,500 m elevation at the margin of the hyperarid Atacama Desert and Chilean Altiplano (Fig. 1)[20]. Exceptionally high ultraviolet radiation[21], hypersalinity and low temperatures distinguish Pajonales as an analogue for the evaporitic basins of Mars. Polyextreme conditions pose serious challenges for life but complex microbial populations can inhabit and have high potential to be preserved in salt crusts and sediments[16,22–24]. These properties make salt-encrusted basins and their microbiomes ideal terrestrial analogues to map extant or extinct biosignatures and promising future targets for missions to Mars[16–18].

Pajonales exhibits features from both physical and biological processes highly relevant to biosignature search on Mars[20]. Fractal-like ridge networks, patterned ground and shrinkage crack terrains of abiotic and/or biotic origin (Extended Data Fig. 1) are morphologically similar to the polygonal terrain observed from orbit in salt-encrusted basins on Mars[20,25]. The Pajonales palaeolake basin contains hydrated Ca sulfate and chloride minerals and modern and fossilized stromatolite and microbialite formations that serve as protective refugia for soil crust or endolithic assemblages[26–29] (Extended Data Fig. 1). These sulfate–chloride habitats have high-to-very-high preservation potential and are cousins to the better studied 'microbial structures, textures and microfossils preserved in carbonates, silica and clay-rich deposits'[30,31].

At Pajonales, we studied a 2.78 km^2 landscape of surficial polygonal networks and domical structures ('Dome Field') composed of near-uniform mineralogy (gypsum, $CaSO_4{\cdot}2H_2O$)[20,22]. Landscape morphology was classified and quantified by UAV and with five in situ nested random ecological studies (Methods, Extended Data Fig. 2, Extended Data Tables 1 and 2 and Supplementary Fig. 1). The presence/absence of biosignatures was characterized with standard visual assessments (7,765 images, n = 1,154 random samples) of two distinct biological morphotypes comprising photosynthetic microorganisms: (1) endolithic colonization, that is, layered communities within salt structures[23], and (2) biological soil crusts[2,32] (BSCs), that is occupying surficial habitats (Extended Data Fig. 1). Raman and visible short-wave infrared (VSWIR) spectroscopy, multiplex microarray immunoassays and 16S ribosomal RNA gene sequencing of pigmented samples revealed distinctive markers for life. Markers included carotenoid biosignatures expressed as orange–pink layers produced by halophilic, radiation-resistant and desiccation-resistant bacteria and archaea; and chlorophyll biosignatures, as green layers produced by cyanobacteria (Extended Data Fig. 1). In all samples, absorption features in the reflectance spectra corresponded to biosignatures in the green and orange–pink endolithic bands.

Biotic abundance was quantified as the probability of colonization, defined as percentage colonization (%Col = number of samples with biosignatures of photosynthetic communities ÷ total number of samples × 100, Methods). Assessments were validated by VSWIR and Raman spectroscopy, microscopy, mass spectroscopy, genomics and proteomics (Methods, Extended Data and Supplementary Fig. 2). Mineralogy and geochemistry were assessed in the field using VSWIR and in the laboratory using X-ray diffraction (XRD) and micro X-ray fluorescence spectroscopy (Methods).

Biosignatures cluster by spatial scale and macro- and microhabitat

The spatial distribution of microbial biosignatures at Pajonales varied significantly. Specifically, 9.2% of the location was colonized, with communities clustered in isolated, non-random patchy spatial patterns (Fig. 2, Extended Data Table 2 and Supplementary Data Fig. 3).

UAV and ground survey data classified geomorphological features into four macrohabitats (metre to kilometre scales) and six microhabitats (centimetre scale), each with distinct biosignature probabilities (Table 1 and Fig. 3). Every dome supported at least one endolithic biosignature, with a mean probability across all dome samples of 40.8% (Table 1 and Extended Data Table 2). Ridges and aeolian cover had

https://doi.org/10.1038/s41550-022-01882-x

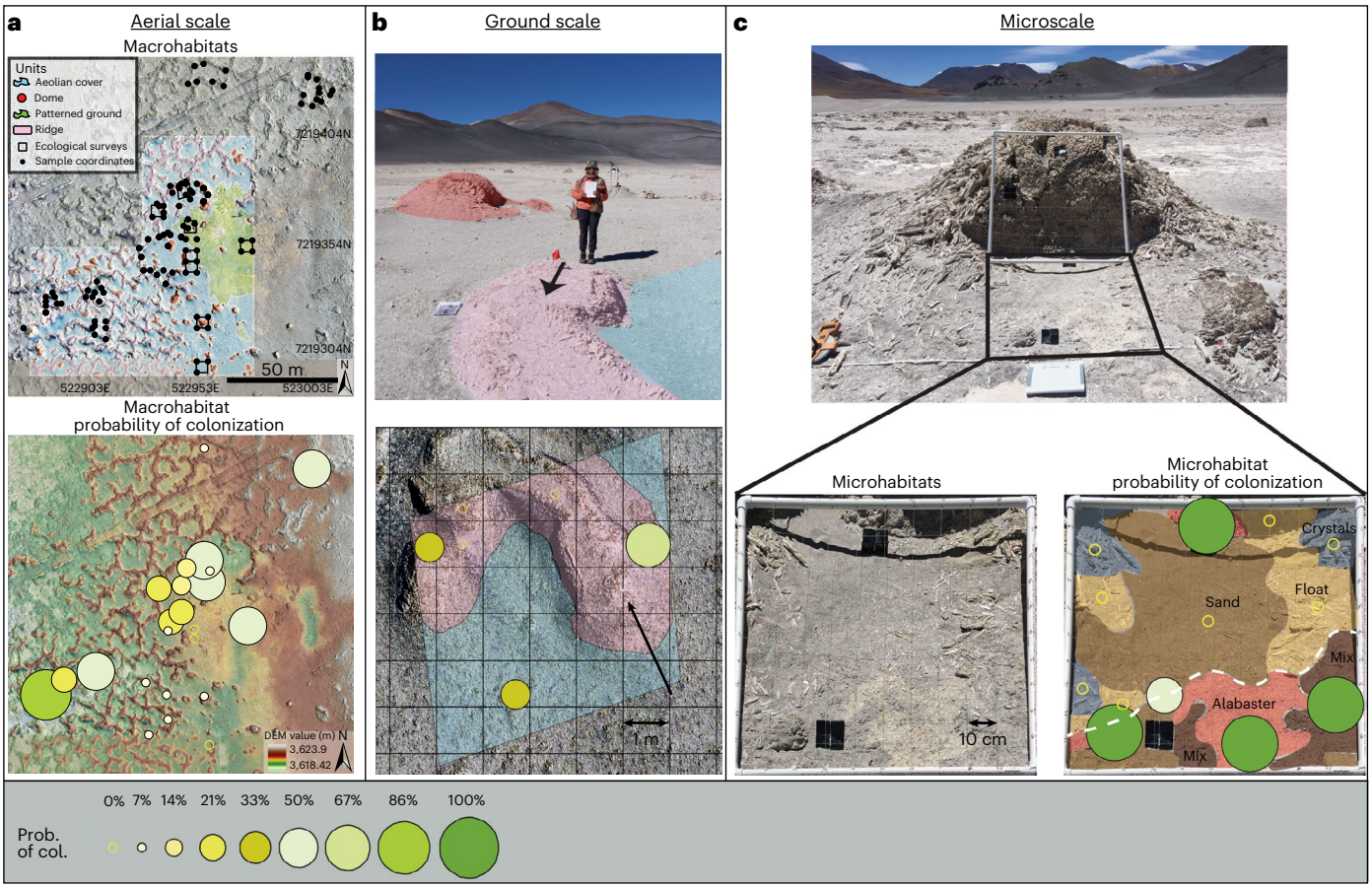


Fig. 2 | Biosignatures exhibit significant spatial heterogeneity and occur in non-random clustered distributions at hierarchical scales. a, Top: aerial view of Dome Field macrohabitats in UAV false colour: domes, ridges/polygonal networks, aeolian cover and patterned ground. Black squares show eight randomly selected survey areas (5 m × 5 m); black dots show individual sample sites (Methods, Extended Data Fig. 2, Extended Data Tables 1 and 2 and Supplementary Figs. 1 and 3). Bottom: spatial survey results (for example, areas, quadrants, Extended Data Table 2) for probability of colonization (as presence/absence of biosignatures, that is, percent colonization, %Col, circles) for pigmented endolithic and BSC communities. DEM, digital elevation model. **b**, Top: visible image of dome, ridge and aeolian cover macrohabitats at ground scale. Bottom: example of spatial survey results for study area 35 of quadrant 6; each %Col circle is a 'site' mean of three random 10 cm^2 'microsite' spatial replicates (ecological study 1 (ES-1), Extended Data Fig. 2). The black arrow indicates the same perspective as in **b**, top. **c**, Top: close-up image of dome visible in **b**, the site of the fine-scale ES-3 study (n = 800 samples; white and black boxes are 1 m^2 quadrats; Extended Data Fig. 2). Bottom left: close-up visible image of microhabitats. Bottom right: biosignature and microhabitat mapping results, in the same visible image, with %Col for each microhabitat. Note the microbial hotspots (green circles) near the microscale habitat transition zone outlined by the white dashed line. Black tiles, ground-truthed HySpex Mjolnir spectral scans.

significantly lower colonization. Patterned ground covered 40.3% of the study area's salar surface, with 50% colonized by BSC (Table 1 and Fig. 3).

Variation in %Col between microhabitats (that is, substrate materials) was more variable than that between macrohabitats (Table 1 and Extended Data Table 2). Alabaster (Table 1) microhabitats were almost universally inhabited and were the most reliable predictor of biosignatures (Fig. 3, Table 1, Extended Data Fig. 3, Extended Data Table 2 and Supplementary Tables 1 and 2). Float and sand were rarely colonized, except when sand was near alabaster (Supplementary Table 2 and Supplementary Fig. 4). Crystal colonization was macrohabitat dependent and ranged from <5% in aeolian cover to 33% in domes. Fine-scale mapping (Extended Data Fig. 2) identified two types of crystal (type I or type II). Type II crystals are distinguishable by a powdery alabaster layer topping the crystal, and this textural difference was significantly correlated with higher biosignature probabilities (Supplementary Data, fine-scale statistical analyses). Alabaster was thus strongly associated with the presence of biosignatures regardless of macrohabitat. Patterned ground colonization was binary: all microstructures supported photosynthetic communities, whereas all bare salar surface samples contained no biosignatures (Fig. 3 and Table 1).

Water controls habitability, spatial patterns at all scales

We hypothesized that macro- and microhabitats are proxies for water availability at Pajonales and that the underlying spatial heterogeneity of water content acts as the primary driver for non-random habitat (for example, alabaster) and biosignature distributions. To confirm this hypothesis, we completed in situ microclimate monitoring and wetting experiments (Extended Data Table 3). These data showed that alabaster, type II crystals and microstructures capture and/or retain liquid water up to twice as long as other microhabitats. Type II crystals are often located near open cracks in domes, where efflorescence and infiltration are greatest[33]. Our data revealed a significantly higher probability of endolithic biosignatures in these microhabitats (Supplementary Data, fine-scale statistical analyses). Likewise, alabaster typically encompasses the bases of ridges and domes, where humidity is high. Alabaster and type II crystals contain a well connected three-dimensional network of small pores (average size <6 µm; Extended Data Fig. 4) that are absent from other microhabitats. The porous network transports brines upwards (that is, efflorescence), explaining the higher water availability and presence of biosignatures.

https://doi.org/10.1038/s41550-022-01882-x

Table 1 | Macro- and microhabitats exhibit significantly different probabilities for containing endolithic or BSC biosignatures

		Surface geomorphological features	Probability of biosignatures
Habitat type	**Dimensions**	**Description**	**(%Col)[b]**
Macrohabitats[a]			
Domes	1–7.5 m, ≤20-cm-thick crust	Subcircular, convex-upward, hollow; tumuli[33]; topographic highs (~0.1–2.0 m); 8.7% of total surface area of study location (8,193 m^2)	40.8
Ridges (or polygon ridges)	1–100s of m long	Continuous topographic gypsum features (height 0.1–1.87 m); ridges are the exteriors of large-scale polygonal networks (diameter: 5–15 m); 28.2% of total surface area; 'polygon ridges' or 'pressure ridges'[33]	18.4
Aeolian (surface) cover	Salar wide	Gypsum regolith comprising intermixed gypsum grains (silt to sand sized) and wind-blown volcanic grains; 48.5% of total surface area of study location; quasiflat aeolian surface cover comprises the interior of large-scale polygonal networks	10.2
Patterned ground	Salar wide	Bare gypsum quasiflat surface covered by microstructures (height 1–3 cm); fractal-like patterned ground (Fig. 3a, green) among other habitat types (Fig. 3a, blue, pink, red); 14.5% of total surface area	50
Microhabitats (≤10 cm^2)[c]			
Alabaster	<200 µm	Powdery, finely crystalline gypsum with abundant intercrystalline pores within domes, ridges and aeolian cover	87.5
Crystals	<18 cm	Large gypsum crystals (selenite) with abundant intercrystalline pores within domes, ridges and aeolian cover	15.7
Float	<10 cm	Mixture of loose sand and eroded crystals, typically lying horizontally on the surface	0
Sand	<2 mm	Gypsum grains (locally mixed with quartz, feldspar, haematite, magnetite grains)	0[d]
Bare salar surface	Not applicable	Flat-lying surface with gypsum crystals (001) oriented perpendicular to the surface; interiors of micropolygons	0
Microstructures	<2 cm	Narrow, sinuous networks of cemented, finely crystalline gypsum with positive relief; rims of micropolygons	100

[a]Detected in drone and simulated satellite imagery and characterized with VSWIR and Raman spectroscopy. [b]Probability of biosignatures, as %Col (for χ^2 tests and detailed statistics, see Extended Data Table 2 and Supplementary Tables 1–3). [c]Detected in situ in ground surveys and analysed via VSWIR and Raman spectroscopy in the field and laboratory. [d]Sand %Col ranged from 0 to 7% (Extended Data Table 2).

Other environmental factors, such as light and nutrient content, were not predominant controls on microbial distributions at Pajonales based on microclimate data, fine-scale studies or laboratory analyses (Supplementary Data, water availability and controls). Nonetheless, it is important to reiterate that these factors shape endolithic and BSC spatial distributions, with their effects well documented for many hypersaline and desert environments[1,2,4,23,34–39].

Predictability: habitability and biosignature probability maps

To predict biosignature distributions in Pajonales we first compiled field-based data to extract scalable rules, statistical probabilities and quantitative correlations (Fig. 4). These data were used to train convolutional neural networks (CNNs)[40] and spatial generalized additive models (GAMs) to predict macro- and microhabitat types, that is, biosignature-bearing geologic materials (Fig. 5).

CNN model (Fig. 5b,c 'prediction') and GAM (Fig. 5d) results aligned well with ground-truth data (Fig. 4; 'true' in Fig. 5b,c). At the aerial scale (Extended Data Table 1), the influences of physical and hydrological processes were evident amongst macrohabitats, each with distinct biosignature detection probabilities (Figs. 4 and 5a, Table 1 and Extended Data Tables 2 and 4). At UAV and HiRISE spatial resolutions (Fig. 5a), a fully connected CNN model classified, on a per-pixel basis, the four main types of macrohabitat across Pajonales. Identification confidence values ranged between 78.8% and 95.1% (UAV, aerial scale) and between 54.8% and 87.0% (HiRISE, orbital scale).

At ground to fine (metre to centimetre) scales, CNN and GAM results reflected ecological statistical results (Figs. 4 and 5) and confirmed that microhabitat type is a credible and diagnostic predictor for the presence of biosignatures. GAMs identified alabaster as the most effective (positive) predictor for biosignatures regardless of macrohabitat (Figs. 4f and 5d). These models also flagged float and sand as having the weakest associations (negative predictors, Supplementary Tables 3 and 4). CNN models were trained on images of geological materials (for example, alabaster, type II crystal) within dome and aeolian cover macrohabitats and performed well on validation and test datasets (that is, images withheld during network training, Fig. 5b,c and Supplementary Figs. 5–8). For instance, alabaster was predicted with a median accuracy of 76% and a range of 65–90% across 10 runs with random train/validation/test partitions (or 'splits'; see Methods for accuracy definition and details of model runs, Supplementary Fig. 6). Thus, CNN models demonstrated predictive capability for the presence of geological materials with high probability for containing biosignatures.

Together, the empirical conditional probabilities and spatial biosignature distributions quantified above are representative of Pajonales as a whole. However, as with any statistical model, further honing of the CNN model and GAM to capture other effects will require more training data. Future work on terrestrial-trained CNN models will expand capabilities to include the physical scale of the features of interest and further instrumentation. In addition to CNNs, generative models (for example, cellular automaton simulations) could explore interactions between life and its physical environment.

Future directions in the search for biosignatures on Mars

The framework and models presented here demonstrate how quantitative systems ecology melded with deep learning feature recognition and prospecting can be a powerful tool to expedite the search and detection of biosignatures in terrestrial analogues. For the Pajonales test site, the pairing of empirical biosignature probabilities with artificial intelligence–machine learning models enhanced success through identification and prediction of macro- and microscopic

https://doi.org/10.1038/s41550-022-01882-x

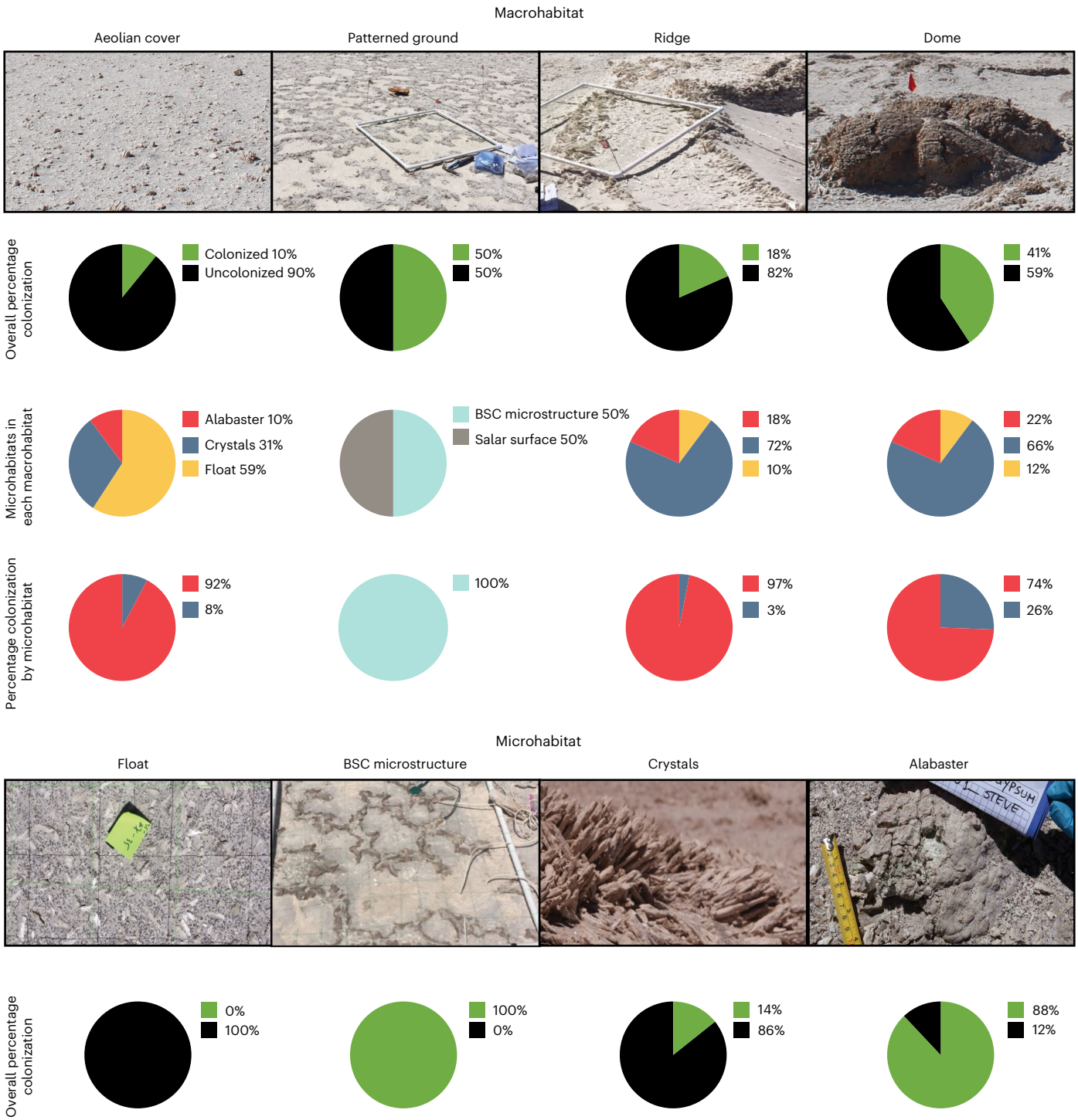
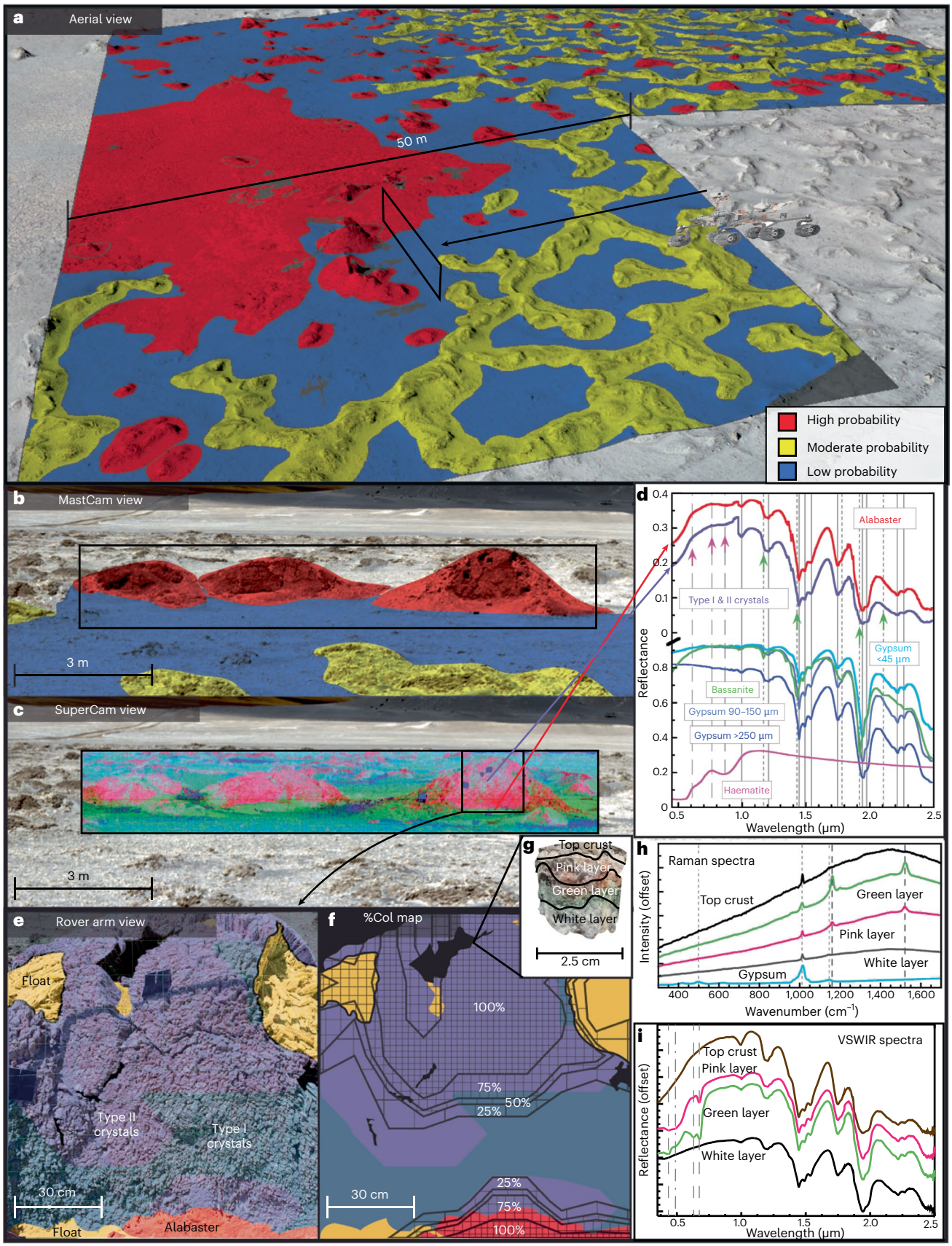


Fig. 3 | Macro- and microhabitat composition and biosignature detection probabilities. Columns are by habitat and rows by indicator. For example, 90% of aeolian cover macrohabitat samples were uncolonized, while 10% had photosynthetic community biosignatures. In the aeolian cover macrohabitat, 59% of samples were float (0% of float samples were colonized, that is, no samples contained biosignatures), 10% were alabaster (92%Col) and 31% were crystals (8%Col). By microhabitat, 0% of all float samples had biosignatures. Probabilities, standard errors and methods are further detailed in Methods, Extended Data and Supplementary Information.

biosignature-bearing targets. Whereas a random search yielded a 9.2% probability of detecting biosignatures, the targeted search of macro- and microscale geologic features and materials, guided by machine learning models, delivered up to an ~87.5% chance of locating biosignatures in the first sample while reducing the initial search space by 85–97%. In light of these advances, our new approach represents a path forward for terrestrial analogue research to produce scientific results that are objective, comparable and replicable, as well as novel in codifying expert experience.

Terrestrial analogues are foundational to successful exploration on Mars[16–19,30,31,41–45]. Our study provides a clear and defined general path (at different scales) for astrobiological exploration, focused on Mars and tested and applied at a specific analogue site. The sulfate–chloride palaeolake system at Pajonales has arisen from evaporation

https://doi.org/10.1038/s41550-022-01882-x



over millennia in a climate that may be one of the best available Earth analogues to the Noachian–Hesperian transition on Mars[20]. The macro- and microscopic structures and biotic systems modelled in this study represent an important addition to the catalogue of terrestrial Mars analogues in the literature[1,4,16,30,31,38] (Extended Data Fig. 1) with quantitative probabilities for biosignatures.

Fig. 4 | Nested spatial scale biosignature probability and habitat maps. **a**, Aerial-scale biosignature probability map by macrohabitat (Table 1 and Extended Data Table 2). **b**, Rover ground view (Mastcam) of macrohabitat biosignature probability map. **c**, Rover ground view (SuperCam) showing HySpex Mjolnir-acquired spectral map generated with the first three eigenimages of the minimum noise fraction transform of the hyperspectral image cube in the R, G, B channels (Methods). In general: pink/white, dome type I and II crystals; green, alabaster. **d**, HySpex Mjolnir VSWIR spectra of dome type I and II crystals and alabaster. Laboratory spectra of gypsum (three grain sizes), bassanite $CaSO_4{\cdot}0.5H_2O$ (USGS spectral library) and <125 µm haematite (Fe_2O_3)[121] are plotted for reference, with notable mineral absorption features indicated with vertical lines. **e**, Rover arm scale microhabitat map, based on ecological survey data (red, alabaster; yellow, float; blue–grey, type I crystals; purple, type II crystals). **f**, Contour map of endolithic biosignatures (probability, 0–100%Col) overlain on microhabitat map of the dome shown in **e. g**, Image of biosignatures in a type II crystal from the dome (comparable to Remote Micro Imager) showing brown/tan surface crust (top crust), pink and green endolithic biosignature layers and a bottom white gypsum layer. **h**, Raman spectra (comparable to the Raman Laser Spectrometer on ExoMars) for type II crystals with a strong gypsum ν_1 peak at 1,013 cm^{-1} and weak peaks at 499 cm^{-1} for ν_2 and 1,143 cm^{-1} for ν_3 (ref. [122]), including colonized pink and green layers with Raman peaks at 1,162 and 1,523 cm^{-1} indicative of carotenoids[123]. **i**, Analytical Spectral Devices (ASD) VSWIR spectra (comparable to Mars 2020 SuperCam VISIR spectra) of type II crystals, indicating gypsum mineralogy. Note absorption features characteristic of chlorophyll near 0.625 and 0.675 µm (dashed vertical lines) in the green spectrum and of chlorophyll near 0.675 µm and carotenoids near 0.5 µm (dash–dot vertical line) in the pink spectrum. These ASD data suggest that the detector is at least as sensitive as the laboratory cell counts reported for the coloured band layers (Supplementary Fig. 2).

Our study generated predictive models (Fig. 5) and exploration roadmaps (Fig. 4; example roadmap, Supplementary Data) for extant biosignatures within a polyextreme environment. Artificial intelligence–machine learning models successfully learned to identify macro- and microgeological features with the highest probability of containing biosignatures. This proof-of-concept deep learning-based feature recognition[46] for predictive biosignature exploration in a specific analogue relevant to Mars exploration was selected to highlight and test the fundamentals of this method within logistical and budgetary constraints. However, the method is not limited to this analogue and we encourage the astrobiology community to consider our approach for standardizing research across terrestrial analogue sites, features, instrument suites and (bio)signatures. Spectral analyses on Mars rely on libraries of reference spectra first compiled in controlled settings on Earth[47]; similarly, the interpretation of Martian environments could be advanced through a library of quantitative reference information on biosignature distribution and ecological data from a multitude of Earth-based analogues compiled with a coherent methodological framework. Such a library could assist future Mars mission scientists in the selection of facies, mineral assemblages and structures with the highest chance of containing biosignatures. Ultimately, we hope the approach will facilitate compilation of a databank of biosignature probability[13] and habitability algorithms, roadmaps and models and serve as a guide for exploration on Mars. The framework may also have applications to other astrobiology targets, such as the surface of Titan, the plumes of Enceladus or the ice cover of Europa[48].

Methods

Study location

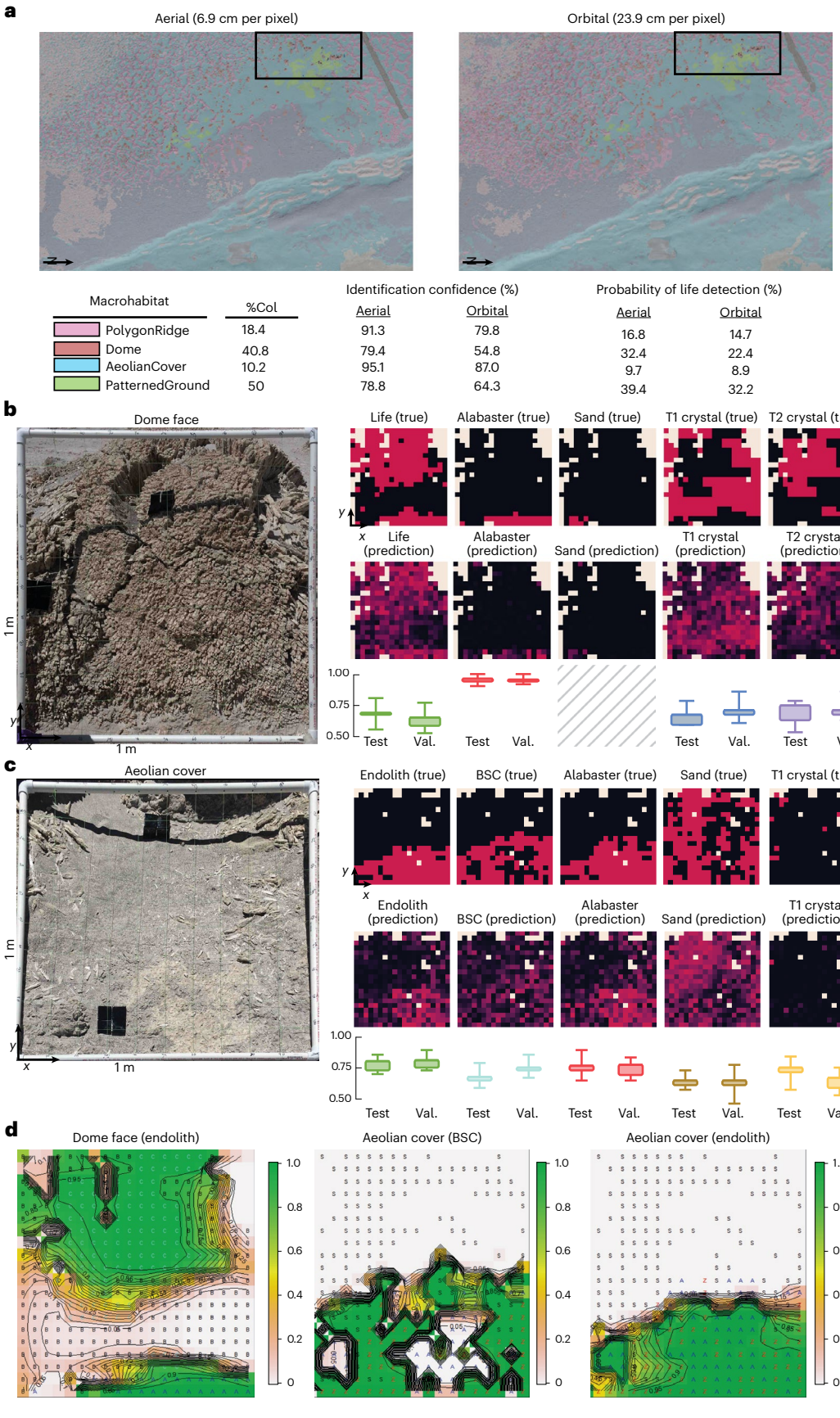
The ~3 km^2 field study location (25.14149° S, 68.77215° W) is situated in the eastern–central portion of the 104 km^2 Salar de Pajonales (Pajonales) basin, a 4.09-Myr-old lakebed with active lagoons and fossil salar structures (Fig. 1)[49,50]. There, low precipitation (80–150 mm yr^{-1}) and high evaporation potential (1,350 mm yr^{-1}) are typical of the arid conditions of the Atacama Desert and Altiplano, and have strengthened since the Miocene[51]. Average air temperature is 5 °C, with values as low as −20 °C recorded. Seasonal precipitation largely occurs via austral summer rains from the east (Altiplanic winter), and southerly winter storms and occasional snow. Despite significant annual precipitation, high winds, low humidity and strong insolation combine to keep Pajonales mostly ice free, with water availability remaining low in much of its surface environment. The water table is shallow in portions of the salar, where at least 14 perennial lagoons exist. The bedrock underlying and adjacent to the salar is associated with Cenozoic volcanism, still active today, including nearby thermal springs. Volcanic rocks are mainly of intermediate to felsic composition (andesites and rhyolites). Carbonates (calcite), borates (ulexites), chlorides (halite), quartz and minor feldspars and other minerals, locally abundant and associated with gypsum, compose the main mineral assemblage[49,50].

Physical environment and habitats

UAV mapping of habitats/geomorphology/digital elevation model. Our study defines a habitat as the physical environment in which a particular organism or community of organisms may live. Macrohabitats were classified first on the basis of field observations and UAV images and subsequently refined and ground-truthed with field assessments. Initial UAV imagery and ground-based assessments revealed several surficial geomorphologic features with potential habitability, that is, ecological macrohabitat types. These macrohabitat (Table 1) features included (1) ridge networks[20,33] (~1–100 m length), (2) domes (≤1–7.5 m diameter; tumuli)[20,22,33], (3) fractal-like patterned ground (square centimetres to hundreds of square metres in area), comprised of bare gypsum salar surface ('salt pan') covered by microstructures (1–3 cm height)[20,52] and (4) aeolian surface cover ('gypsum regolith')[20,33] between and around these features (square centimetres to hundreds of square metres in area). These four distinct geomorphic units were mapped in Esri ArcGIS software using an orthophotomosaic and digital elevation model, both derived from UAV images with the software package Pix4D. Ridges were identified as elongate, continuous, positive topographic features with length:width ratio (L:W) values of >2:1. Domes were identified as circular to subcircular, isolated, positive topographic features with L:W values of <2:1. Patterned ground was distinguished from other flat surfaces on the basis of its tan colour and distinct polygonal pattern. All other flat surfaces comprised aeolian cover, identified from its sandy texture, beige–grey colouration and occasional dark and light rippling on the leeward sides of ridges and domes. A fully convolutional neural network (DeepLab v3+ model[22] with a ResNet50 backbone[46]) was trained to classify, on a per-pixel basis (that is, semantic segmentation), the terrain in the orthophotomosaic of the field site and surrounding area at UAV and HiRISE ground sampling distances (6.9 and 23.9 cm per pixel respectively). Training images were labelled by a human analyst with on-the-ground experience at the field site using the criteria delineated above for ridges, domes, patterned ground and aeolian cover. Additional categories were added as necessary to incorporate terrain not found within the confines of the study site. Network training was carried out until the loss rate stabilized, typically <40 epochs (one epoch = one complete pass of training data through the network), with final test accuracies (identification confidence) presented in Fig. 5a for each class. See also CNN methods below for details of training, validation and testing datasets.

Microhabitat (Fig. 3, Table 1 and Extended Data Fig. 1) categories were identified and refined during nested spatial scale studies (described below) by visual assessments of texture, coherence and colour in the field and spectral imagery acquired on the ground. We collected hyperspectral panoramic images of the dome habitat using a HySpex Mjolnir hyperspectral camera mounted on a scanning tripod and placed approximately 30 m from the dome target. From each scan the pushbroom-style dual camera collects spectral data across the visible–near-infrared (VNIR) to short-wave infrared (SWIR) spectral

https://doi.org/10.1038/s41550-022-01882-x



range of 400–3,500 nm at a bandwidth sampling interval of 3.0 or 5.1 nm as it scans horizontally across the scene at a constant pan rate. We collected three scans under clear skies in mid-afternoon lighting conditions, while calibrating the frame period and integration time of the HySpex camera to the lighting conditions to maximize reflectance and eliminate saturated pixels. We then collected a final complete

https://doi.org/10.1038/s41550-022-01882-x

Fig. 5 | Habitat and biosignature probability maps from CNN models and spatial GAMs. a, Top: pixel-wise classification of images with ground sampling distances of 6.9 cm per pixel (aerial) and 23.9 cm per pixel (orbital), generated by a fully connected CNN (details in Methods). Bottom: colours corresponding to macrohabitats in classification scenes (top), and probability of biosignatures, CNN identification confidence values and biosignature detection expected values (identification confidence × %Col). **b**, Dome macrohabitat fine-scale (cm) CNN results for classification at the grid-cell image level. Top row: 'true' shows ES-3 ground-truth data. Middle row: CNN 'predictions' for presence of endolithic biosignatures and microhabitat type. Note that predictive heatmaps represent CNN predictive capability over the macrohabitat images (split into training/validation/test sets) and belong to just one of the 10 randomized runs, each with a random training/validation/test split (see Methods for randomized run definition and Supplementary Fig. 6 for recall/precision curve per randomized split). Bottom row: classification accuracy (*y* axis) across 10 randomized runs, summarized by boxplots, each defined by the following numbers: mean, s.d. σ (minimum, maximum). **c**, Aeolian cover macrohabitat CNN results. Top row: 'true' shows ground-truth data. Middle row: CNN predictions. Bottom row: classification accuracy (*y* axis) across 10 randomized runs, summarized by boxplots, each defined by the following numbers: mean, s.d. σ (minimum, maximum). **d**, Left: GAM contour map of probability of endolithic biosignatures in dome (see Methods and Supplementary Tables 3 and 4 for model details including parameter estimates and odds ratios). A, alabaster; B, type I crystal; C, type II crystal. Note biosignature-bearing 'transition zones' (tz) at boundary of type I and type II crystals or alabaster. Centre: GAM probability contour map of BSC biosignatures in aeolian cover. A, alabaster; S, sand; Z, alabaster + sand. Right: GAM contour probability map of endolithic biosignatures in aeolian cover (Supplementary Tables 3 and 4). Note that probabilities of biosignatures in sand in the upper half are zero, whereas sand (orange, yellow and pink colours) with alabaster (green colour) in transition zones is colonized. Colour ramp: red (probability 0) to green (probability 1).

panoramic scan using the calibrated parameters. The final scan produced a hyperspectral image cube with a spatial resolution of ~0.5 cm per pixel in the VNIR spectral region and ~1.0 cm per pixel in the SWIR spectral region. The spectral image covers a panoramic view of the primary dome target, three smaller domes next to it and foreground primarily composed of aeolian cover.

Raw data values in the hyperspectral image cubes acquired from the tripod were calibrated to reflectance values using in-scene Spectralon calibration targets. Each scene consists of two spatially overlapping cubes, one in the VNIR wavelength range and the other in the SWIR. Because the optical apertures for the two subsystems that collect these cubes are slightly offset, the pixels in the two component cubes are not perfectly registered in the along-scan (horizontal) direction. Thus, extraction of reflectance spectra of the alabaster and type I and type II crystal units (Fig. 4d) was accomplished by manually selecting the corresponding pixels from the two component cubes by referencing features observed in both. The overwhelming majority of spectra examined contain the signature of gypsum (Fig. 4d), and only subtle spectral differences between different surfaces on and around the domes were noted. Weak shoulders and slight shifts in band position were found for the type I and type II crystal material that are consistent with the presence of bassanite[53]. These are highlighted by green arrows (Fig. 4d). A weak shoulder near 0.6 µm, a reflectance maximum near 0.76 µm and a weak minimum near 0.88 µm are attributed to haematite. These features are indicated by purple arrows and are enhanced in the alabaster-type material. To highlight this subtle spectral variability and reveal any spatial patterns present in that variability, we processed the VNIR and SWIR cubes separately using the minimum noise fraction transform principal component analysis algorithm[54]. The false-colour image shown in Fig. 4c was created by loading the first three bands representing most significant spectral variance from the SWIR cube processed using minimum noise fraction for that scene into the red, green and blue channels of the image. The colour variations in this representation correspond to slight differences in the general gypsum spectrum and are seen to be spatially segregated. While we have not undertaken a separate study to explain the causes of these spectral variations, they are probably due to differences in grain sizes, degree of compaction and perhaps the presence or absence of different minor (non-gypsum) components.

Environmental controls: climate monitoring and wetting experiments. Environmental data for the study location and within domes and microstructures were acquired in situ from October 2018 to November 2019 with a HOBO microstation (H21-001) and Onset Computers logger array. One weather station was connected to the following smart sensors to monitor macroclimate parameters: (1) rainfall (1 mm resolution, HOBO RG3-M rain gauge metric data logger); (2) photosynthetically active radiation (S-LIA-M003), to quantify solar flux in the biologically relevant portion of the spectrum (400–700 nm; range 0–2,700 µmol m^{-2} s^{-1}); (3) wind speed and direction (S-WCF-M003; ±1.1 m s^{-1}, ±7°); (4) air temperature (*T*, °C) and relative humidity (%RH) at 1 m above dome surfaces (~3 m above salar surface; HOBO S-THB-M002; precision ±0.2 °C, ±2.5%RH; shaded by a solar radiation shield, RS3-B).

Microclimate data were acquired through arrays of both stand-alone external *T*/RH HOBO U23 ProV2 loggers (U23-002; precision ±0.21 °C, ±5%RH) and microstations (H21-001). To measure the presence of liquid water on surfaces and in rock interiors (endolithic microenvironments) from fog/dew/rainfall/snow, microstations were connected to (1) *T*/RH smart sensors (S-THB-M008), (2) leaf wetness sensors (S-LWA-M003; precision ±5%RH: reading of 0%RH is dry, 100%RH reflects a sensor grid completely covered by a thin layer of water) and (3) EC5 conductivity probes (S-SMC-M005; baseline >0: liquid water and/or 0.25 m^3 per m^3 soil saturated) placed on both the surfaces and interiors of domes ~0.5–1 cm below the surface where endolithic microbial communities layers occurred. They were also placed on salar surfaces and microstructures in all four cardinal directions ~0.5–1 cm below the surface where BSC layers occurred. On the basis of previous studies[55–58] and successful long-term use of RH and conductivity sensors for direct field measurement of water activity—the parameters controlling microbiological activity in hyperarid deserts[55,56]—a cutoff for liquid water was assumed at RH ≥ 95%. Although lichen communities can undergo photosynthesis at 70%RH, the 95%RH threshold reflects a mid-value for the predominantly cyanobacterial endolithic communities in the Dome Field location, which require liquid water to initiate photosynthesis (range 90–100%)[59–64]. For logger probes inserted into gypsum domes and microstructures, an electric drill was used to create small (~1 cm) boreholes in which *T*/RH sensors (S-THB-M008) were inserted and sealed with silicon epoxy. All sensors logged at 30 min intervals.

This environmental array allowed us to measure naturally available moisture sources to microbial communities, including surface precipitation (photosynthetically active radiation–cloud/storm cover; rainfall/snow–rain gauge, leaf wetness) and concomitant available moisture of interiors of domes, microstructures and bare salar surfaces (*T*/RH, electrical conductivity). In situ rain gauge data were corroborated with nearby station precipitation data, when available, and historical climate data were obtained from reference literature, regional observatories and the Chilean Meteorological Bureau. To ascertain the relative role of microstructures versus bare salar surfaces in capturing and retaining water during rainfall events, a simulated rainfall wetting experiment was completed within each of three (10 mm, 1 mm and control) 1 m^2 quadrats. The following sensors (see above for specifications) were emplaced: (1) a leaf wetness sensor; (2) four HOBO *T*/RH probes

(two in microstructures at the surface and 1 cm below the surface and two in bare salar surface and at 1 cm); (3) two EC5 conductivity probes in microstructures at the surface and 1 cm below the surface. Sensors were placed in the field to acclimatize for two days before the field experiment and left in the field for two days after simulated rain. Rain events were simulated with a gentle spraying of deionized water equally across the quadrat via a commercial sprayer (10 mm = 10 l m^{-2} and 1 mm = 1 l m^{-2}) at 9:00 on 21 April 2019. The control received no water. Visual documentation using a hand-held camera was also acquired every 20 min during daylight for 48 h from the application of rain events to the end of the experiment.

Mineralogy, organic carbon content and nutrient content analyses. The mineralogical composition of powdered specimens sieved under a mesh size of 20 µm was verified with XRD using a Bruker D8 Advance diffractometer with a graphite–monochromatic Cu Kα radiation source (Cu cathode of wavelength Kα = 1.54051 Å) operated at 30 mA and 40 kV. For quantitative analysis of the crystalline phases present in the samples, the powder diffraction file (PDF-4+ 2016) of the International Centre for Diffraction Data was used.

Field Raman spectra were recorded with a customized spectrometer developed to simulate operational concepts planned for the Raman Laser Spectrometer instrument on board the European Space Agency's 2022 ExoMars rover mission. This field instrument performs measurements under conditions similar to those of the Raman Laser Spectrometer on the ExoMars rover, including at least 20 spectra in each sample, with a 50 µm spot size at a working distance of 5 mm and an irradiance level of 0.6–1.2 kW cm^2 with a 532 nm continuous wave laser[65]. The spectrometer covers a spectral range from 100 to 3,800 cm^{-1} with a spectral resolution of 12 cm^{-1}. The instrument was calibrated before each measurement using TECAPET, the same calibration standard as on board ExoMars. Data products were analysed using a statistical procedure previously applied in macro techniques, such as XRD and FT-Raman[66]. Because the spot size of the instrument is only 50 µm and only a single or a few grains are analysed at a time, the statistical analysis was performed on the relatively few (20–30) Raman spectra acquired per sample. A detailed point-by-point analysis procedure is provided in ref. [67].

Reflectance spectrometry is useful for rapid, non-destructive, compositional identification and analysis. Phyllosilicates, carbonates and sulfates display distinct absorption features in the SWIR region (1–2.5 µm) and organic molecules absorb light in the VNIR (0.4–1 µm) wavelength region[68,69]. We used an ASD FieldSpec 4 spectroradiometer to collect solar-illuminated VSWIR reflectance spectra of freshly exposed samples in the field under cloud-free skies and clear lighting conditions, and in the laboratory using a contact probe. The instrument was calibrated to a dark current and a white reference target between each measurement. Reflectance spectra are derived through subtraction of the dark-current spectrum and division by the white-reference spectrum. Spectra were plotted and analysed using the Interactive Data Language Environment for Visualizing Images software suite. Reflectance spectra acquired with the ASD in the field are provided in Fig. 4i for four locations on a sample (Fig. 4g). These illustrate the dominant gypsum features including bands at 0.99, 1.21, 1.45–1.49–1.54 (triplet), 1.75, 1.94, 2.22–2.26 (doublet), 2.43 and 2.48 µm (ref. [53]). Two spectra acquired in the field with the HySpex Mjolnir hyperspectral camera are shown in Fig. 4d. The type I and type II crystal gypsum spectrum (purple spectrum at the top of Fig. 4d) includes shoulder features for bassanite near 1.17, 1.43, 1.92 and 2.1 µm (ref. [53]) that indicate that the type I and type II crystal gypsum unit is less hydrated and some of the gypsum crystals are converted to bassanite. Both the alabaster and type I and type II crystal gypsum spectra include a shoulder near 0.61 µm, a reflectance maximum near 0.76 µm and a weak minimum near 0.88 µm that are consistent with ferrihydrite or fine-grained haematite[70,71]. These features are stronger in the alabaster-type spectrum, indicating that more iron oxide/oxyhydroxide is present in that sample, although the abundance is probably only a few per cent in either gypsum-rich unit.

Porosity was estimated using petrographic image analysis of thin sections prepared with vertical orientation. A blue dye was injected into the epoxy to visualize porosity in the thin sections. After acquisition of the photomicrographs at 10× magnification using a Leica DM4500P LED polarizing microscope, the plane-polarized light photomicrographs (Extended Data Fig. 4a,c) were first normalized using the automatic white balance function in GIMP image processing software (GIMP 2.10.18; https://www.gimp.org/). In a second step a selection mask with an RGB threshold of 50 was applied to select blue areas (pores) in the individual images. The selection was exported to ImageJ (https://imagej.nih.gov/ij/). In ImageJ, binary images (Extended Data Fig. 4b,d) were created that subsequently were used to analyse the percentage area of pixels (black pixels in images Extended Data Fig. 4b,d) representing porosity.

Nutrient content was measured on powdered samples. Instrumentation for total carbon, total organic carbon and total nitrogen was an Exeter Analytical CE-440 elemental analyser. Inorganic carbon was removed from samples before measurement of total organic carbon by acid extraction. This introduces the possibility that some organic compounds are lost during the extraction and rinsing process. Total phosphorus was extracted from powdered samples in a two-step process: combustion at 500 °C followed by dissolution in 0.15 M HCl. The extracts were analysed with the Astoria Pacific A2 segmented flow analyser using the molybdate blue method. Accuracy and precision have a standard error range that is less than the minimum detectable limit.

Water-soluble anions and low-molecular-weight organic acids were analysed using ion chromatography. Two grams of ground samples were diluted in 10 ml of deionized water, sonicated (3 × 1 min cycles) and filtered through a 0.22 µm pore size. The filtrate was loaded into a Metrohm 861 Advanced compact ion chromatographer. Twelve anionic species were measured simultaneously: inorganic—fluoride, chloride, bromide, nitrate, nitrite, sulfate and phosphate, and five low-molecular-weight organic acids—acetate, propionate, formate, tartrate and oxalate. A Metrosep A Supp 7–250 column was used with 3.6 mM sodium carbonate as eluent.

Microbial ecology methods. As in most analogue studies, our research began with initial observations to generate scientific hypotheses. In a second step we amassed statistical data at hierarchical scales to test these hypotheses (Figs. 2 and 3, Extended Data Figs. 1 and 2 and Supplementary Fig. 1). We then quantified the probabilities for biosignatures and repeatedly tested and honed them, enabling the formation of new hypotheses, predictions and rules in an iterative cycle (Extended Data Fig. 2). The resultant robust set of quantitative biosignature, geomorphological and ecological data, rules and probabilities then formed the inputs to deep learning models to construct habitability and biosignature heatmaps and models (Figs. 4 and 5).

Overall approach to quantifying photosynthetic microbial community abundance (probability of colonization) and mapping spatial distribution. Photosynthetic microbial community colonization at Pajonales was investigated from orbit-to-ground scales with a multi-year, multidisciplinary approach that tightly integrated traditional molecular and ecological methods, UAV surveys and ground-based spectroscopic and panoramic camera instruments. We completed an initial field survey, three iterative microbial landscape ecology studies, microenvironmental monitoring and wetting experiments to quantify photosynthetic microbial colonization and identify the controls on microbial abundance and spatial pattern. The study collected presence/absence data, an approach commonly employed in ecology to reduce time and sampling effort, with data collected from, for example, specific points, grid cells and quadrats. Specifically, we quantified the abundance of photosynthetic microbial communities

at multiple scales (that is, location to microscales) as the probability of a sample containing a biosignature, or percentage microbial community colonization (%Col), which was defined as the number of samples (or macro- or microhabitats) containing biosignatures ÷ total number of samples (or macro- or microhabitats) × 100 (refs. [72–87]).

Initial field campaign and assessment. The first-year field campaign (October 2016) was dedicated to a pilot survey of the biogeochemical, biological, ecological, mineralogical and geophysical features of the Pajonales location. Two UAV flight surveys (north location 0.237 km^2; south location 0.198 km^2) and three days of 'walkabout'[88] surveys were completed by the team. It was during these surveys that the visible presence of cryptic endolithic photosynthetic microbial communities in geomorphological structures was discovered and confirmed in situ by VNIR and Raman spectroscopic instruments, revealing beta-carotene and chlorophyll in the banded orange and green endolithic biosignature layered bands, respectively.

Initial data analysis and landscape study design. Walkabout survey and UAV image data from the field campaign served as the baseline to (i) define relevant scales of interest for the study (Extended Data Table 1), (ii) identify and classify geomorphic features of the Pajonales landscape (see UAV mapping above and Extended Data Fig. 1) and (iii) generate initial hypotheses and a detailed research design and sampling plan, which included three comprehensive microbial landscape ecology studies (2016–2019; Extended Data Fig. 2). All three studies included both UAV and ground-based components and tested hypotheses and generated results, rules and probabilities for biosignatures that served as inputs to the iterative hypothesis testing and refinements of the subsequent study (Extended Data Fig. 2). Because of the remoteness and high altitude of Salar de Pajonales, each field study was restricted to 1–2 weeks in duration on the basis of stringent logistical, financial and safety constraints. Three ecological studies were conducted in 2018–2019 to define and map microbial colonization and allow probability functions to be quantified.

1. ES-1 estimated microbial abundance and patterns and the effect of spatial scale using aerial to microscale nested biosignature surveys; ES-1 (i) tested initial hypotheses (Extended Data Fig. 2), (ii) collected first biosignature data and refined macro- and microhabitat classifications and (iii) produced the first rules and biosignature probabilities. These data were then tested for predictability in ES-2.
2. ES-2 quantified microbial abundance and patterns and the influence of multiple sizes and types of habitat; the study tested and honed rules and probabilities from ES-1 and generated new rules and predictions for testing in ES-3 (Extended Data Fig. 2).
3. ES-3 mapped abundance and biosignature spatial pattern at fine (centimetre) scales within individual domes, aeolian cover and ridges (Extended Data Fig. 2). The study continued testing the rules, predictions and probabilities from ES-2 and further refined them. The resultant data were then the basis for generating habitability and biosignature probability heatmaps, spatial analyses (global Moran's) and deep learning models (see deep learning sections below).

To investigate the effect of spatial scale (ES-1) on the probability of photosynthetic microbial community biosignatures (%Col) the study location was subdivided into nested hierarchical scales of study, including quadrants (red grid, 50 m × 50 m), areas (black grid, 5 m × 5 m) and microsites (Extended Data Figs. 1 and 2). All samples in ES-1 were randomly chosen on the basis of spatial scale (that is, not on the basis of habitat types). UAV imagery was initially classified as 'raised' areas (ridges and domes) or 'flat' areas (aeolian cover (grey) and patterned ground (green)). Initial study hypotheses were (i) that %Col varied significantly by spatial scale (ES-1) and (ii) that raised geomorphic features (perhaps linked to preferential water availability or more suitable rock architecture for colonization)[12,89–91] were more habitable than flat features, that is, that geomorphic unit/ecological habitat type controlled colonization (ES-2). These habitat classifications were iteratively refined over the course of the project (ES-2 and ES-3) as an understanding of the features and associated microbial inhabitants evolved. Furthermore, the 'abiotic' patterned ground observed in 2016 using aerial imagery (green unit) was ascertained in 2018 ground surveys to be microstructures with living photosynthetic BSC communities. To accommodate these new findings, we created microstructure and bare salar surface microhabitat and BSC biological morphotype categories and separated these data from the endolithic colonization mode analyses.

Scales of study. Ecologically relevant scales of study (that is, scales of microbial communities and their interactions) can range from the planetary to the nanoscale. Spatial scales of choice for a study can also be circumscribed by the particular methods selected, instruments deployed and/or features of interest. The 2016 walkabout surveys and UAV imagery revealed surface geomorphologic units and features of varying scales with potential habitability. One initial hypothesis centred on these features as habitat units of ecological importance (for example, raised versus flat units) that could influence the abundance and spatial pattern of microbial communities across the landscape, as has been documented in earlier studies[90,91]. Extended Data Table 1 defines the nested hierarchical scales of physical features in the Pajonales location.

Metagenomic studies and multiplex immunoassay studies. The methods and results for the Pajonales project's metagenomic and immunoassay investigations will be detailed in future companion papers. However, for completeness, here we provide a summary of those studies. Briefly, to identify the microbial composition of endolithic photosynthetic communities, DNA was extracted from 2 g of each gypsum sample layer according to the FastDNA SPIN kit for soil method (MP Bio). We quantified the DNA and its integrity using a NanoDrop (NP1000) labelling the ratio of DNA/RNA 260/230 and the ratio of DNA/protein 260/280. In addition, microbial diversity investigations through Illumina MiSeq sequencing of hypervariable regions of 16S rRNA for bacteria (V1–V3) were carried out using universal bacterial (27F/519R) primers 16S rRNA gene amplification with standard protocols[92]. Amplicon sequencing was performed using MR DNA (www.mrdnalab.com) according to the manufacturer's guidelines. Multiplex fluorescent immunoassays were used for rapid estimation of the presence of microbial markers and main metabolic traits: 0.5 g of powdered samples were analysed with the SOLID-LDChip immunosensor (Signs of Life Detector–Life Detector Chip)[93–95], using a fluorescent sandwich microarray immunoassay as described elsewhere[95–97]. The LDChip is the core sensor of the SOLID instrument for detecting possible traces of life in the field of planetary exploration[96,98]. The LDChip used in this work contained 200 polyclonal antibodies produced against a wide range of immunogens (small organic molecules; peptides and proteins; exopolysaccharides and lipopolysaccharides; spores and whole cell lysates of bacteria and archaea, including 20 antibodies to different cyanobacterial strains)[94,96].

Genomic and multiplex immunoassay studies revealed phototrophic communities to be dominated by bacterial phyla that included Cyanobacteria, Proteobacteria, Actinobacteria, Thermi, Bacteroidetes and Chloroflexota. Leptolyngbya, Gloeocapsa, Cyanothece, Chroococcidiopsis and Calothrix, which are poikilohydric organisms that 'equilibrate rapidly to ambient water content'[99] and are dependent upon the presence of liquid water for photosynthesis and growth, comprised some of the main cyanobacterial members. LDChip revealed the presence of a diversity of microbial markers immunodetected with antibodies to several cyanobacterial strains, Bacteroidetes as *Salinibacter* spp., haloarchaea as *Halorubrum* sp., thermophiles as *Pyrococcus* spp. and some potential perchlorate-reducing bacteria such as *Azospira* spp., among

others. LDChip also showed positive immunodetection with antibodies to proteins involved in nitrogen fixation, and even the pink-coloured light-driven proton pump bacteriorhodopsin in one of the samples.

Detailed methods for microbial landscape ecology studies of photosynthetic microbial community biosignatures

ES-1 (spatial scale): quantifying abundance and nested scale spatial patterns. Extended Data Fig. 2 and Supplementary Fig. 1 show the spatial design and hierarchical sampling for ES-1. A random sampling scheme was adopted for each nested spatial scale[72–87] to provide an objective representation of the probability of biosignatures, as %Col. The data were sampled at resolutions from hundreds of metres to centimetres. To be considered representative, ≥10% of the Pajonales study location was mapped. Two randomly selected quadrants (50 m × 50 m) were chosen. A third quadrant was added for the habitat surveys for ES-2. Within each quadrant, we randomly selected four 'areas' (5 m × 5 m), nine sites (1 m × 1 m) within each area and three microsites (10 cm × 10 cm) within each site (n = 216 samples).

Within each microsite, the following data were recorded: (i) presence/absence of visible photosynthetic communities/biosignatures; (ii) aspect/orientation of visible biosignatures (four cardinal directions) and depth (to centimetres) and colour of colonized layers; (iii) microhabitat type(s); (iv) samples for mineralogy and nutrient content; (v) samples for gravimetric water content; (vi) samples for molecular biodiversity and biomarker profiling by multiplex immunoassays. %Col of photosynthetic microbial communities was assessed in each microsite by the visible presence/absence of orange and/or green pigmentation, that is, standard ocular estimate methods utilized in multiple previous studies[23,32,85,86,100–107]. Ocular estimates of the presence/absence of photosynthetic communities within the green and orange visible endolithic layers and BSC[32] were also further corroborated (i) in situ with VSWIR and Raman instruments and (ii) in the Centro de Biotecnología's laboratory (for example, Supplementary Fig. 2) with images and emission spectra with a confocal laser scanning microscopy wavelength λ-scan function[108] and with a scanning electron microscope according to the standards of the Centro de Biotecnología of the Universidad Católica del Norte and the Geology Department of the University of Montana.

ES-2 (habitats): quantifying photosynthetic microbial community abundance and habitats. ES-2 was designed to investigate the effect of spatial variation of specific habitats on colonization. Classification of initial macro- and microhabitat data from ES-1 was refined and a larger analytical scale of observation, the 'large-scale polygon/polygonal network', was added with further UAV imagery analysis. It revealed that ridges and aeolian surface cover units comprised the exterior edges and interior surfaces, respectively, of polygons at larger scales (1 to >30 m in diameter; Extended Data Fig. 1). Previous studies have shown these distinct types of network unit (of all sizes) to influence biotic parameters or habitat geochemistry[90,91,109]. In this regard, we quantified abundance in seven large polygons in ES-2 (5–35 m in diameter) and seven domes (1–5 m diameter) (Extended Data Fig. 2). For the large polygons, seven exteriors, that is, ridges in ES-1, and seven polygon interiors, that is, aeolian cover in ES-1, were mapped (n = 147 total samples). In each habitat type, seven samples were randomly chosen and the following data collected: (i) presence/absence of visible photosynthetic community biosignatures; (ii) aspect/orientation of biosignatures (four cardinal directions) and depth of colonized layers, if present; (iii) microhabitat type(s); (iv) samples for mineralogy and nutrient content; (v) samples for gravimetric water content; (vi) samples for molecular biodiversity and biomarker profiling by multiplex immunoassays.

ES-3 (fine scale): dome, ridge and aeolian cover fine-scale studies. In this third iteration, data and the rules and probabilities for biosignatures from ES-1 and ES-2 were further analysed, refined and tested as inputs and predictions in ES-3 (Extended Data Fig. 2). Microhabitat surface areas and photosynthetic biotic colonization were also mapped and quantified at much higher spatial resolution and detail to construct heatmaps for fine-scale (centimetre-to-millimetre) habitats and biosignatures. Microhabitats were also further precharacterized (and predictions generated) using ground-based VSWIR hyperspectral scans (see methods above) and assessments of panoramic visual imagery. Subsequently, for each macrohabitat unit 1 m^2 quadrats (5 cm grid, n = 400 samples per quadrat, Extended Data Fig. 2) were used to map, quantify and confirm fine-scale biosignature patterns. Macroscale habitat orientation was held constant by selecting unit areas adjacent to each other in the same orientation, thus controlling for macroscale climate variations including light, temperature and wind speed/direction. Visual and VSWIR hyperspectral panoramic and ground-based (that is, hand-held camera) visual images of the ridge, dome and aeolian cover macrohabitats were also acquired to overlay onto biosignature data and assess microhabitat visual and spectral fidelity. Following colocated scans, in each 5 cm^2 sample, the following data were recorded: (i) presence/absence of photosynthetic community biosignatures; (ii) type of colonization (endolithic and/or BSC); (iii) aspect/orientation of biosignatures (four cardinal directions) and depth of colonized layers, if present; (iv) microhabitat characteristics (for example, type, features, spatial location, presence of a transition, powdery consistency in crystals, proximity to vertical crack). Mineralogy and molecular biology samples were also collected. Data from ES-3 formed inputs for spatial statistical analyses and deep learning models.

Statistical ecology and deep learning methods

Microbial landscape studies (ES-1–2): spatial and habitat effects on colonization. Chi-square tests for independence were conducted to assess differences in the ratio of colonization among the macro- and microhabitats, and at each spatial scale (Extended Data Table 2).

Spatial randomness/aggregation analyses (ES-1–3; Ripley's *K*, global Moran's *I*). Spatial statistics are a broad suite of statistical methods enabling the identification and comparative analysis of spatial patterns utilizing data that have a spatial characteristic[82,110]. One such tool is Ripley's K function, which is used to identify randomness and clustering of points within a spatial group[82,110,111]. Ripley's K and global Moran's I were used to assess spatial pattern at UAV and fine scales for endolithic and BSC communities at Pajonales. Ripley's K analyses were conducted with the R package spatstat[112,113] using grid coordinates and presence/absence data for ES-1–3. Permutation tests (999 random permutations) for the global Moran's I statistic of the same data were conducted with the R package (https://link.springer.com/article/10.1007/s11749-018-0599-x). Spatial relationships were defined by inverse distance weighting (1/distance). For results see Supplementary Fig. 3.

Deep learning methods. At the fine scale, the predictability of (endolithic/BSC) biosignatures, and the association with particular microhabitats (proxies for water availability), were analysed with pivot tables (for rules) and tested for predictive power with CNNs (Fig. 5a–c) and spatial GAMs (Fig. 5d).

GAM methods

Predictive model for endolithic colonization of dome macrohabitat. A spatial logistic GAM was conducted using the mgcv package in R (version 4.0.4) with a restricted maximum likelihood for the parameter estimation:

$$\log\left(P/(1-P)\right) = \beta_0 + \beta_1\,\mathrm{T2} + \beta_2\,\mathrm{T1} + f(x\_\mathrm{coord}, y\_\mathrm{coord}), \quad (1)$$

where P is the probability that the response variable (endolith, binary, 0–1) occurs in a cell[114–116] and the term $P/(1-P)$ is referred to as the odds

ratio or the likelihood of observing endolithic biosignatures given the presence of the microhabitat T1 or T2. A smoothness function $f(x\_\text{coord}, y\_\text{coord})$ was included to account for the spatial coordinates x_coord and y_coord. Observations that are closer in distance will have a stronger correlation than observations that are further apart. The smoothness function was a bivariate P-spline.

Alabaster was not included in the model given that it was linearly dependent with endolith (correlation was close to 1) and inclusion in the model would have impacted the outcomes of other variables in the model. This was also observed for sand and float, which were both negatively correlated to endolith (close to −1) and thus were not included in the model.

The microhabitats considered in the model as binary outcomes (1, microhabitat exists; 0, does not exist) included type I crystal microhabitat (T1) and type II crystal microhabitat (T2). The model can be written as

$$(P/(1-P)) = \exp\left(\beta_0 + \beta_1\,\text{T2} + \beta_2\,\text{T1} + f(x\_\text{coord}, y\_\text{coord})\right). \quad (2)$$

Estimation and inference results related to the spatial logistic GAM in equation (1) are given in Supplementary Table 3. With this model, a probability contour map for endolithic life in the dome face macrohabitat was created (Fig. 5d, left panel).

Predictive models for endolithic and BSC biosignatures in aeolian cover macrohabitat. Spatial logistic GAMs were also conducted, with a restricted maximum likelihood for the parameter estimation, to separately estimate the probability of endolithic and BSC biosignatures in the aeolian cover macrohabitat. Both models were formulated as follows:

$$\log(P/(1-P)) = \beta_0 + \beta_1\,\text{sand} + f(x\_\text{coord}, y\_\text{coord}), \quad (3)$$

where P is the probability that the response variable (endolith/BSC, both binary, 0–1) occurs in a cell and $P/(1-P)$ is the likelihood of observing endolith or BSC biosignatures given the presence of sand. The smoothness function described earlier is also included in equation (3). Similar to the dome model above, alabaster was not included in the model given its linear dependence with endolithic colonization (correlation was close to 1) and that inclusion in the model would have impacted the performance of other variables in the model.

Pivot-table and decision-tree methods. Pivot tables count the number of events, where for Pajonales each event is some combination of microhabitat (for example, alabaster, sand) and biosignature, as %Col (endolith, BSC). We derive a decision tree from the pivot table, in which the order of yes/no questions centred on a microhabitat is chosen to maximize the likelihood of detecting a biosignature in the fewest questions possible. Pivot tables and decision-tree results are shown in Extended Data Fig. 3, Supplementary Tables 1 and 2 and Fig. 4.

CNN methods. A CNN is a machine learning model that specializes in the isolation and extraction of salient features from spatial data in a data-driven manner[117]. That is, a feature is deemed salient if it is predictive of the target ground truth. To create the model, optical data were taken in situ with an iPhone 7 of the dome and aeolian cover macrohabitats and later split into 32 pixel by 32 pixel subsets and annotated for the following variables: microhabitat type and presence/absence of endolithic or BSC biosignatures. To quantify the effect of the split of the macrohabitat into mutually exclusive training, validation and test image chips, we train and evaluate the CNN model across 10 randomized splits (for examples, see Supplementary Fig. 7), and generated recall–precision curves per randomized split (see recall–precision curves in Supplementary Fig. 6).

Annotated images (four per cell in the 10 × 10 cell grid) from each macrohabitat were split randomly into subsets of 70% training, 15% validation and 15% testing, following standard best practices to avoid overfitting and prevent leakage of information between training and validation/testing datasets[118]. After removing images with holes/no samples, 349 images remain in the dome macrohabitat and 378 images remain in the aeolian cover macrohabitat (Supplementary Fig. 8). Predictions on the training set were used to fit the CNN model parameters. Predictions on the validation set were not used to fit the model, but they were used to assess whether the model overfits the training data: if the training accuracy (defined as [true positives + true negatives]/[size of dataset]) and validation accuracy numbers are similar, then no overfitting occurred. If overfitting occurs, it means that the model can predict only the ground truth of the training data by memorization, rather than distinguishing actual salient patterns in the imagery that are generally predictive across the entire macrohabitat. On the other hand, it is possible to rely too much on the validation accuracy, such that it is now the model design (for example, depth of the neural net) that is at risk of overfitting, but in this case on the validation set. Therefore, a third subset, the testing set, was used at the end of each randomized run to measure the final test accuracy of the model. In effect, an accuracy in the test set that is comparable to that of the validation set implies that the model design process was not contaminated by our assessments on the validation set. Therefore, we can conclude that the model generalizes across the entire macrohabitat. See Fig. 5 for results, and Supplementary Fig. 7 for examples of randomized image splits.

For pixel-wise classification of the study site and its surroundings a fully connected CNN, or fully convolutional network, was employed[119]. The orthophotomosaics were split into 708 individual subframes approximately 18.5 m × 18.5 m in extent for input into the fully convolutional network. We used standard data augmentation techniques of rotation, reflection and illumination adjustments to increase the total data volume 18-fold for a total of 12,744 images with ground-truth labels. The images were split into training (75%), validation (12.5%) and testing (12.5%) datasets and statistics were drawn on the distribution of classes within each dataset to ensure consistency in the representation of each class across the datasets (Supplementary Fig. 5). We chose to split the orthophotomosaic scenes into the following 7 classes: AeolianCover, MottledGround, PatternedGround, PolygonRidge, Dome, Objects and Road. Ground-truth labels were created in a geographic information system program using the highest-resolution orthophotomosaic by a remote sensing geologist with on-the-ground knowledge of the field site. A version of the ResNet50 CNN[46] trained on over 1 million images from the ImageNet database was used as the network backbone for our fully convolutional network. Training was carried out on clusters on the Infrastructure for Scientific Applications and Advanced Computing facility at the University of Tennessee, Knoxville, with the Adaptive Moment Estimation (Adam) solver until the loss rate stabilized (typically <40 epochs). We calculated the standard evaluation parameters boundary F1 contour matching score (BF-score), precision and recall, as well as network uncertainty, which was used to formulate a 'confidence score'. Network uncertainty was calculated using the Monte Carlo dropout method[120]. A confidence score was formulated as follows:

$$\text{Confidence score}\ (\%) = \left(\text{TP}_i\,\text{TC}_i + \text{FP}_i\,(100 - \text{FC}_i)\right)/N_i$$

where TP_i and FP_i are true and false positives, respectively, for class i. TC_i and FC_i are the network certainty (0–100%) averaged over all true and false predictions, respectively, for class i, and N_i is the total number of pixels predicted as class i. Evaluation metrics are presented in Extended Data Table 4. While typical evaluation metrics for semantic segmentation, such as the BF-score, measure how accurately a network draws a segmentation boundary around features, the confidence score reports the degree to which a user should trust the network certainty analysis for a given class. The confidence score metric, therefore, enables a more nuanced analysis of network results and feature identification

than typical evaluation metrics. Confidence scores can be high even if the boundaries drawn around objects are inaccurate (that is, BF-scores are low), and conversely confidence scores can be low even if BF-scores are high. If the confidence score is high relative to the BF-score it means that the uncertainty analysis is reporting useful information upon which users should rely (that is, the network is not confident about false predictions and is confident about true predictions), but if the opposite is true users should be wary of the network certainty analysis because the network is probably overly certain of false predictions. Therefore, we report confidence scores values in the main text, but for relevant classes we report BF-score, precision and recall alongside confidence scores for reference (Extended Data Table 4).

Data availability

The data that support the findings of this study are available from the corresponding author upon reasonable request. Sequence data from this study have been deposited to the DDBJ database under BioProject PRJDB14848 with accession numbers DRR425262 to DRR425263.

Code availability

Code and data to reproduce the results of the CNN models can be found at https://github.com/SETI/endo.

References

1. Des Marais, D. J. The biogeochemistry of hypersaline microbial mats. *Adv. Microb. Ecol.* **14**, 251–274 (1995).
2. Belnap, J., Welter, J., Grimm, N., Barger, N. & Ludwig, J. Linkages between microbial and hydrologic processes in arid and semiarid watersheds. *Ecology* **86**, 298–307 (2005).
3. Houghton, J. et al. Spatial variability in photosynthetic and heterotrophic activity drives locale $\delta^{13}C_{org}$ fluctuations and carbonate precipitation in hypersaline microbial mats. *Geobiology* **12**, 557–574 (2014).
4. Allwood, A., Walter, M., Burch, I. & Kamber, B. 3.43 billion-year-old stromatolite reef from the Pilbara Craton of Western Australia: ecosystem-scale insights to early life on Earth. *Precambrian Res.* **158**, 198–227 (2007).
5. Al-Najjar, M. et al. Spatial patterns and links between microbial community composition and function in cyanobacterial mats. *Front. Microbiol.* **5**, 406 (2014).
6. Warren-Rhodes, K., Dungan, J., Piatek, J. & McKay, C. Ecology and spatial pattern of cyanobacterial island patches in the Atacama Desert. *J. Geophys. Res. Biogeosciences* **112**, G04S15 (2007).
7. Allwood, A., Walter, M., Kamber, B., Marshall, C. & Burch, I. Stromatolite reef from the early Archaean era of Australia. *Nature* **441**, 714–718 (2006).
8. Meslier, V. et al. Fundamental drivers for endolithic microbial community assemblies in the hyperarid Atacama Desert. *Environ. Microbiol.* **20**, 1765–1781 (2018).
9. Finstad, K. et al. Microbial community structure and the persistence of cyanobacterial populations in salt crusts of the hyperarid Atacama Desert from genome-resolved metagenomics. *Front. Microbiol.* **8**, 1435 (2017).
10. Wilhelm, M. et al. Constraints on the metabolic activity of microorganisms in Atacama surface soils inferred from refractory biomarkers: Implications for Martian habitability and biomarker detection. *Astrobiology* **18**, 955–966 (2018).
11. Dillon, J. et al. Spatial and temporal variability in a stratified microbial mat community. *FEMS Microbiol. Ecol.* **68**, 46–58 (2009).
12. Rillig, M. & Antonovics, J. Microbial biospherics: the experimental study of ecosystem function and evolution. *Proc. Natl Acad. Sci. USA* **116**, 11093–11098 (2019).
13. Sephton, M. & Carter, J. The chances of detecting life on Mars. *Planet. Space Sci.* **112**, 15–22 (2015).
14. Naveh, Z., & Lieberman, A. S. *Landscape Ecology: Theory and Application* (Springer, 2013).
15. Mony, C., Vandenkoornhuyse, P., Bohannan, B. J. M., Peay, K. & Leibold, M. A. A landscape of opportunities for microbial ecology research. *Front. Microbiol.* **11**, 2964 (2020).
16. Summons, R. et al. Preservation of Martian organic and environmental records: final report of the Mars Biosignature Working Group. *Astrobiology* **11**, 157–181 (2011).
17. Farmer, J. & Des Marais, D. J. Exploring for a record of ancient Martian life. *J. Geophys. Res.* **104**, 26,977–26,995 (1999).
18. Stoker et al. We should search for extant life on Mars in this decade. *Bull. AAS* **53** (2021); https://doi.org/10.3847/25c2cfeb.36ef5e33
19. Jakowsky, B. et al. Mars, the nearest habitable world—a comprehensive program for future Mars exploration. *Bull. AAS* **53** (2021); https://doi.org/10.3847/25c2cfeb.e5222017
20. Hinman, N. et al. Surface morphologies in a Mars analog Ca sulfate salar, High Andes, Northern Chile. *Front. Astron. Space Sci.* **8**, 797591 (2022).
21. Cabrol, N. et al. Record solar UV irradiance in the tropical Andes. *Front. Environ. Sci.* **2**, 19 (2014).
22. Phillips, M.S. et al. Planetary mapping using Deep Learning: a method to evaluate feature identification confidence applied to habitats in Mars-analogue terrain. *Astrobiology* **23** (2023).
23. Wierzchos, J. et al. Adaptation strategies of endolithic chlorophototrophs to survive the hyperarid and extreme solar radiation environment of the Atacama Desert. *Front. Microbiol.* **6**, 934 (2015).
24. Lynch, K. et al. Near-infrared spectroscopy of lacustrine sediments in the Great Salt Lake Desert: an analog study for Martian paleolake basins. *J. Geophys. Res. Planets* **120**, 599–623 (2015).
25. El-Maarry, M., Pommerol, A. & Thomas, N. Analysis of polygonal cracking patterns in chloride-bearing terrains of Mars: indicators of ancient playa settings. *J. Geophys. Res.* **113**, 2263–2278 (2013).
26. Onstott, T. et al. Paleo-rock-hosted life on Earth and the search on Mars: a review and strategy for exploration. *Astrobiology* **19**, 1230–1262 (2019).
27. Davila, A. & Schulze-Makuch, D. The last possible outposts for life on Mars. *Astrobiology* **16**, 159–168 (2016).
28. Osterloo, M. M. et al. Geologic context of proposed chloride-bearing materials on Mars. *J. Geophys. Res.* **115**, E10012 (2010).
29. Flauhaut, J., Martinot, M., Bishop, J.L., Davies, G.R. & Potts, N.J. Remote sensing and in situ mineralogic survey of the Chilean salars: an analog to Mars evaporate deposits? *Icarus* **282**, 152–173 (2017).
30. Bosak, T., Moore, K., Gong, J. & Grotzinger, J. Searching for biosignatures in sedimentary rocks from early Earth and Mars. *Nat. Rev. Earth Environ.* **2**, 490–506 (2021).
31. Balci, N. et al. Biotic and abiotic imprints on Mg-rich stromatolites: lessons from Lake Salda, SW Turkey. *Geomicrobiol. J.* **37**, 401–425 (2020).
32. Williams, A., Buck, B., Soukup, D. & Merkler, D. Geomorphic controls on biological soil crust distribution: a conceptual model from the Mojave Desert (USA). *Geomorphology* **195**, 99–109 (2013).
33. Warren, J. *Evaporites: a Geological Compendium* 2nd edn (Springer, 2016).
34. Wierzchos, J. et al. Microbial colonization of Ca sulfate crusts in the hyperarid core of the Atacama Desert: implications for the search for life on Mars. *Geobiology* **9**, 44–60 (2010).
35. Robinson, C. K. et al. Microbial diversity and the presence of algae in halite endolithic communities are correlated to atmospheric moisture in the hyper-arid zone of the Atacama Desert. *Environ. Microbiol.* **17**, 299–315 (2013).

36. Jørgesen, B. & Des Marais, D. Optical properties of benthic photosynthetic communities: fiber-optic studies of cyanobacterial mats. *Limnol. Oceanogr.* **33**, 99–113 (1988).
37. Szynkiewicz, A., Moore, C., Glamoclija, M., Bustos, D. & Pratt, L. Origin of coarsely crystalline gypsum domes in a saline playa environment at the White Sands National Monument, New Mexico. *J. Geophys. Res.* **115**, F02021 (2010).
38. Walker, J., Spear, J. & Pace, N. Geobiology of a microbial endolithic community in the Yellowstone geothermal environment. *Nature* **434**, 1011–1013 (2005).
39. Rasuk, M. et al. Microbial characterization of microbial ecosystems associated to evaporites domes of gypsum in Salar de Llamara in Atacama Desert. *Microb. Ecol.* **68**, 483–494 (2014).
40. Chen, L., Papandreou, G., Kokkinos, I., Murphy, K. & Yuille, A. Semantic image segmentation with deep convolutional nets and fully connected CRFs. Preprint at *arXiv* https://arxiv.org/abs/1412.7062 (2014).
41. Chan, M. et al. Exploring, mapping and data management integration of habitable environments in astrobiology. *Frontiers in Microbiology* **10**, 147 (2019).
42. Farmer, J. in *From Habitability to Life on Mars* 1–12 (Elsevier, 2018).
43. Hays, L. et al. Biosignature preservation and detection in Mars analog environments. *Astrobiology* **17**, 363–400 (2017).
44. Fairen, A. et al. Astrobiology through the ages of Mars: the study of terrestrial analogues to understand the habitability of Mars. *Astrobiology* **10**, 821 (2010).
45. Green, J. et al. Call for a framework for reporting evidence for life beyond Earth. *Nature* **598**, 575–579 (2021).
46. He, K., Xiangyu, Z., Shaoqing, R. & Jian, S. Delving deep into rectifiers: surpassing human-level performance on ImageNet classification. In *2015 IEEE International Conference on Computer Vision (ICCV)* 1026–1034 (IEEE, 2015).
47. Adams, J. B. & Filice, A. L. Spectral reflectance 0.4 to 2.0 microns of silicate rock powders. *J. Geophys. Res.* **72**, 5705–5715 (1967).
48. National Academies of Sciences, Engineering & Medicine. *Origins, Worlds and Life: a Decadal Strategy for Planetary Science and Astrobiology 2023–2032* (National Academies Press, 2022).
49. Rodríguez Albornoz, C. Geology and Controls on Microbiota of the Salar de Pajonales (7.209.000–7.226.500 N.–510.000–530.000 E), Antofagasta, Northern Chile. Master's thesis, Univ. Católica del Norte Antofagasta (2018).
50. Naranjo, J., Villa, V. & Venegas, C. *Geology of the Salar de Pajonales Area and Cerro Moño. Antofagasta and Atacama Regions* (Geological Maps of Chile Basic Geology Series No. 153 (1: 100.000), National Geological Service, Geology and Mining Subsection, 2013).
51. Schween, J., Hoffmeister, D. & Löhnert, U. Filling the observational gap in the Atacama Desert with a new network of climate stations. *Glob. Planet. Chang.* **184**, 103034 (2020).
52. Gutiérrez, F. & Cooper, A. Surface morphology of gypsum karst. *Treatise Geomorphol.* **6**, 425–437 (2013).
53. Bishop, J. L. et al. Spectral properties of Ca-sulfates: gypsum, bassanite and anhydrite. *Am. Mineral.* **99**, 2105–2115 (2014).
54. Green, A., Berman, M., Switzer, P. & Craig, M. D. A transformation for ordering multispectral data in terms of image quality with implications for noise removal. *IEEE Trans. Geosci. Remote Sens.* **26**, 65–74 (1988).
55. Davis, W., Pater, I. & McKay, C. P. Rain infiltration and crust formation in the extreme arid zone of the Atacama Desert, Chile. *Planet. Space Sci.* **58**, 616–622 (2010).
56. McKay, C. P. et al. Temperature and moisture conditions for life in the extreme arid region of the Atacama Desert: four years of observation including the El Niño of 1997–1998. *Astrobiology* **3**, 393–406 (2003).
57. Warren-Rhodes, K., Rhodes, K., Liu, S., Zhou, P. & McKay, C. Nanoclimate environment of cyanobacterial communities in China's hot and cold hyperarid deserts. *J. Geophys. Res.* **112**, G01016 (2007).
58. Warren-Rhodes, K. et al. Physical ecology of hypolithic communities in the central Namib Desert: the role of fog, rain, rock habitat and light. *J. Geophys. Res.* **118**, 1451–1460 (2013).
59. Lange, O., Kilian, E. & Ziegler, H. Water vapor uptake and photosynthesis of lichens: performance differences in species with green and blue–green algae as phycobionts. *Oecologia* **71**, 104–110 (1986).
60. Lange, O. L., Meyer, A. & Büdel, B. Net photosynthesis activation of a desiccated cyanobacterium without liquid water in high air humidity alone. Experiments with a *Microcoleus sociatus* isolated from a desert soil crust. *Funct. Ecol.* **8**, 52–57 (1994).
61. Palmer, R. & Friedmann, E. I. Water relations and photosynthesis in the cryptoendolithic microbial habitat of hot and cold deserts. *Microb. Ecol.* **18**, 111–118 (1990).
62. Potts, M. & Friedmann, E. Effects of water stress on cryptoendolithic cyanobacteria from hot desert rocks. *Arch. Microbiol.* **130**, 267–271 (1981).
63. Tracy, C. et al. Microclimate and limits to photosynthesis in a diverse community of hypolithic cyanobacteria in northern Australia. *Environ. Microbiol.* **12**, 592–607 (2010).
64. Azúa-Bustos, A. et al. Hypolithic cyanobacteria supported mainly by fog in the coastal range of the Atacama Desert. *Microb. Ecol.* **51**, 568–581 (2011).
65. Rull, F. et al. ExoMars Raman Laser Spectrometer for ExoMars. *Proc. SPIE* **8152**, 81520J (2011).
66. Kontoyannis, C. G., Orkoula, M. & Koutsoukos, P. Quantitative analysis of sulphated calcium carbonates using Raman spectrometry and X-ray powder diffraction. *Analyst* **122**, 33–38 (1997).
67. Lopez-Reyes, G. et al. Analysis of the scientific capabilities of the ExoMars Raman Laser Spectrometer Instrument. *Eur. J. Mineral.* **25**, 721–733 (2013).
68. Hunt, G. Spectral signatures of particulate minerals in the visible and near infrared. *Geophysics* **42**, 501–513 (1977).
69. Bishop, J. L. in *Remote Compositional Analysis: Techniques for Understanding Spectroscopy, Mineralogy, and Geochemistry of Planetary Surfaces* (eds Bishop, J. L. et al.) 68–101 (Cambridge Univ. Press, 2019).
70. Morris, R. V. et al. Evidence for pigmentary hematite on Mars based on optical, magnetic and Mössbauer studies of superparamagnetic (nanocrystalline) hematite. *J. Geophys. Res.* **94**, 2760–2778 (1989).
71. Bishop, J. L., Pieters, C. M. & Burns, R. G. Reflectance and Mössbauer spectroscopy of ferrihydrite–montmorillonite assemblages as Mars soil analog materials. *Geochim. Cosmochim. Acta* **57**, 4583–4595 (1993).
72. Levin, S. A. The problem of pattern and scale in ecology. *Ecology* **73**, 1943–1967 (1992).
73. Underwood, A. J., Chapman, M. G. & Connell, S. D. Observations in ecology: you can't make progress on processes without understanding the patterns. *J. Exp. Mar. Biol. Ecol.* **250**, 97–115 (2000).
74. Turner, M. G. Landscape ecology: the effect of pattern on process. *Annu. Rev. Ecol. Syst.* **20**, 171–197 (1989).
75. Turner, M. G., Gardner, R. H. & O'Neill, R. V. *Landscape Ecology in Theory and Practice* (Springer, 2001).
76. Wiens, J. A., Chr, N., Van Horne, B. & Ims, R. A. Ecological mechanisms and landscape ecology. *Oikos* **66**, 369–380 (1993).
77. Urban, D., O'Neill, R. & Shugart, H. Landscape ecology. *BioScience* **37**, 119–127 (1987).
78. Underwood, A. J. et al. *Experiments in Ecology: their Logical Design and Interpretation using Analysis of Variance* (Cambridge Univ. Press, 1997).

79. Quinn, G. P., & Keough, M. J. *Experimental Design and Data Analysis for Biologists* (Cambridge Univ. Press, 2002).
80. Neyman, J. & Pearson, E. S. On the problem of the most efficient tests of statistical hypotheses. *Philos. Trans. R. Soc. Lond. A* **231**, 289–337 (1933).
81. Zar, J. H. *Biostatistical Analysis* 5th edn (Prentice-Hall/Pearson, 2010).
82. Ripley, B. D. *Journal of the Royal Statistical Society Series B (Methodological)* **39**, 172-212 (1977).
83. Royle, J. A. & Nichols, J. D. Estimating abundance from repeated presence–absence data or point counts. *Ecology* **84**, 777–790 (2003).
84. Krebs, C. *Ecological Methodology* 2nd edn (Addison-Wesley, 1999).
85. Warren-Rhodes, K., Dungan, J., Piatek, J. & McKay, C. Ecology and spatial pattern of cyanobacterial community island patches in the Atacama Desert. *J. Geophys. Res.* **112**, G04S15 (2007).
86. Belnap, J., Phillips, S., Witwicki, D. & Miller, M. Visually assessing the level of development and soil surface stability of cyanobacterially dominated biological soil crusts. *J. Arid Environ.* **72**, 1257–1264 (2008).
87. Warren-Rhodes, K. et al. Hypolithic cyanobacteria, dry limit of photosynthesis, and microbial ecology in the hyperarid Atacama Desert. *Microb. Ecol.* **52**, 389–398 (2006).
88. Yingst, R. et al. 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**, 327–347 (2020).
89. Shen, J., Wyness, A., Claire, M. & Zerkle, A. Spatial variability of microbial communities and salt distributions across a latitudinal gradient in the Atacama Desert. *Microb. Ecol.* **82**, 442–458 (2021).
90. Barrett, J. et al. Variation in biogeochemistry and soil biodiversity across spatial scales in a polar desert ecosystem. *Ecology* **85**, 3105–3118 (2004).
91. Pointing, S. B. et al. Highly specialized microbial diversity in hyper-arid polar desert. *Proc. Natl Acad. Sci. USA* **106**, 19964–19969 (2009).
92. Chiodini, R. et al. Microbial population differentials between mucosal and submucosal intestinal tissues in advanced Crohn's disease of the ileum. *PloS ONE* **10**, e0134382 (2015).
93. Rivas, L. A. et al. A 200-antibody microarray biochip for environmental monitoring: searching for universal microbial biomarkers through immunoprofiling. *Anal. Chem.* **80**, 7970–7979 (2008).
94. Sanchez-Garcia, L. et al. Microbial biomarker transition in high-altitude sinter mounds from El Tatio (Chile) through different stages of hydrothermal activity. *Front. Microbiol.* **9**, 3350 (2019).
95. Parro, V. et al. SOLID3, a multiplex antibody microarray-based optical sensor instrument for in situ life detection in planetary exploration. *Astrobiology* **11**, 15–28 (2011).
96. Parro, V. et al. A microbial oasis in the hypersaline Atacama subsurface discovered by a life detector chip: implications for the search for life on Mars. *Astrobiology* **11**, 969–996 (2011).
97. Blanco, Y., Moreno-Paz, M., Aguirre, J. & Parro, V. in *Hydrocarbon and Lipid Microbiology Protocols* (eds McGenity, T. J. et al.) Ch. 159 (Springer, 2017).
98. Moreno-Paz, M. et al. Detecting nonvolatile life and nonlife-derived organics in a carbonaceous chrondrite analogue with a new multiplex immunoassay and its relevance for planetary exploration. *Astrobiology* **18**, 1041–1056 (2018).
99. Ekwealor, J. & Fisher, K. Life under quartz: hypolithic mosses in the Mojave Desert. *PLoS ONE* **15**, e0235928 (2020).
100. Williams, A., Buck, B. & Beyene, M. Biological soil crusts in the Mojave Desert, USA: micromorphology and pedogenesis. *Soil Sci. Soc. Am.* **76**, 1685–1695 (2012).
101. Archer, S. et al. Endolithic microbial diversity in sandstone and granite from the McMurdo Dry Valleys, Antarctica. *Polar Biol.* **40**, 997–1006 (2017).
102. Noffke, N., Gerdes, G., Klenke, T. & Krumbein, W. Microbially induced sedimentary structures—a new category within the classification of primary sedimentary structures. *J. Sediment. Res.* **71**, 649–656 (2001).
103. Fierer, N. & Jackson, R. The diversity and biogeography of soil bacterial communities. *Proc. Natl Acad. Sci. USA* **103**, 626–631 (2006).
104. Caruso, T. et al. Stochastic and deterministic processes interact in the assembly of desert microbial communities on a global scale. *ISME J.* **5**, 1406–1413 (2011).
105. Valverde et al. Prokaryotic community structure and metabolisms in shallow subsurface of Atacama Desert playas and alluvial fans after heavy rains: repairing and preparing for next dry period. *Front. Microbiol.* **10**, 1641 (2019).
106. Sun, H. Endolithic microbial life in extreme cold climate: snow is required, but perhaps less is more. *Biology* **2**, 693–701 (2013).
107. Maier, S. et al. Photoautotrophic organisms control microbial abundance, diversity and physiology in different types of biological soil crusts. *ISME J.* **12**, 1032–1046 (2018).
108. Roldan, M., Ascaso, C. & Weirzchos, J. Fluorescent fingerprint of endolithic phototrophic cyanobacteria living within halite rocks in the Atacama Desert. *Appl. Environ. Microbiol.* **80**, 2998–3006 (2014).
109. Cockell, C. et al. 0.25 Ga salt deposits preserve geological signatures of habitable conditions and ancient lipids. *Astrobiology* **20**, 864–877 (2019).
110. Ripley, B. D. *Spatial Statistics* (Wiley, 1981).
111. Gelfand, A. E., Diggle, P., Guttorp, P., & Fuentes, M. (eds) *Handbook of Spatial Statistics* (CRC Press, 2010).
112. Dixon, P. M. in *Encyclopedia of Environmetrics*, 1796-1803 (Wiley, 2006).
113. Baddeley, A., Rubak, E. & Turner, R. Spatial point patterns: methodology and applications with R. *J. Stat. Softw.* **75**, 2 (2016).
114. Wood, S. *Generalized Additive Models: an Introduction with R* Ch 3–5 (Chapman and Hall/CRC, 2006).
115. Simon, R. & Wood, N. GAMS in practice: mgcv. In *Generalized Additive Models: an Introduction with R* 2nd ed (eds Blitzstein, J., Faraway, J., Tanner, M. & Zidek, J.) Ch 7 (Chapman and Hall/CRC, 2017).
116. Fang, X. & Chan, K.-S. *Generalized Additive Models with Spatio-temporal Data* (Univ. Iowa); https://stat.uiowa.edu/sites/stat.uiowa.edu/files/techrep/tr396.pdf
117. LeCun, Y. et al. Backpropagation applied to handwritten zip code recognition. *Neural Comput.* **1**, 541–551 (1989).
118. Roberts, M. et al. Common pitfalls and recommendations for using machine learning to detect and prognosticate for COVID-19 using chest radiographs and CT scans. *Nat. Mach. Intell.* **3**, 199–217 (2021).
119. Shelhamer, E., Long J. & Darrell, T. Fully convolutional networks for semantic segmentation. Preprint at *arXiv* https://arxiv.org/abs/1605.06211 (2016).
120. Gal, Y. & Ghahramani, Z. Dropout as a Bayesian approximation: representing model uncertainty in deep learning. Preprint at *arXiv* https://arxiv.org/abs/1506.02142 (2016).
121. Bishop, J. L. & Murad, E. in *Volcano–Ice Interactions on Earth and Mars* (eds Smellie, J. L. & Chapman, M. G.) 357–370 (Special Publication No. 202, Geological Society, 2002).
122. Buzgar, N., Buzatu, A. & Sanislav, I. V. The Raman study of certain sulfates. *An. Stiintificie Univ. Al. I. Cuza IASI Geol.* **55**, 5–23 (2009).
123. Jehlicka, J., Edwards, H. & Oren, A. Raman spectroscopy of microbial pigments. *Appl. Environ. Microbiol.* **80**, 3286–3295 (2013).

Acknowledgements
This study was supported by the NASA Astrobiology Institute (NAI) via grant NNA15BB01A. We acknowledge XRD data and its respective analysis generated from the MAINI's scientific equipment and Centro de Biotecnología at Universidad Católica del Norte. G.C.-D. and C.D. thank BHP Minerals Americas Project 32002137 (2016–2020). We thank K. Phillips (kennydphillips.com) for graphic design. V.P. thanks the Ministry of Science and Innovation (Spain) (grant RTI2018-094368-B-I00), State Agency of Research (MCIN/AEI/10.13039/501100011033) and ERDF 'A way of making Europe' for funding and M. García-Villadangos for technical support.

Author contributions
K.W.-R., M.P., N. Hutchinson, N.A.C. and K.L.R. designed and implemented the experiment. K.W.-R., M.P., C.T.-C., K.L.R., N.A.C., N. Hinman., V.P., P. Sobron, P. Sarazzin, C.J., C.R. and C.D. conducted field operations and sampling. J.M., M.P., C.M., P. Sobron, P. Sarazzin and D.W. collected in situ visible, multi- and hyperspectral data from instruments and instrumented drones. M.P., F.K., D.A., L.N.B., K.L. and K.W.-R. developed deep learning and statistical models. K.W.-R., M.P., K.L.R., V.P., N. Hinman and J.L.B. processed samples and conducted geochemical and mineralogical analyses. K.W.-R., M.P., K.L.R., C.T.-C., N.H., C.D., G.C.-D., M.H.H. and V.P. performed geological and biological analyses. K.W.-R., K.L.R., D.W., G.C.-D., D.A. and C.T.-C. installed, collected and analysed microclimate data. All authors wrote the manuscript. N.A.C. conceived the NAI project study. K.W.-R., K.L.R. and N.H. designed the ecological study. All authors reviewed and edited a manuscript draft.

Competing interests
The authors declare no competing interests.

Additional information
Extended data is available for this paper at https://doi.org/10.1038/s41550-022-01882-x.

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41550-022-01882-x.

Correspondence and requests for materials should be addressed to Kimberley Warren-Rhodes.

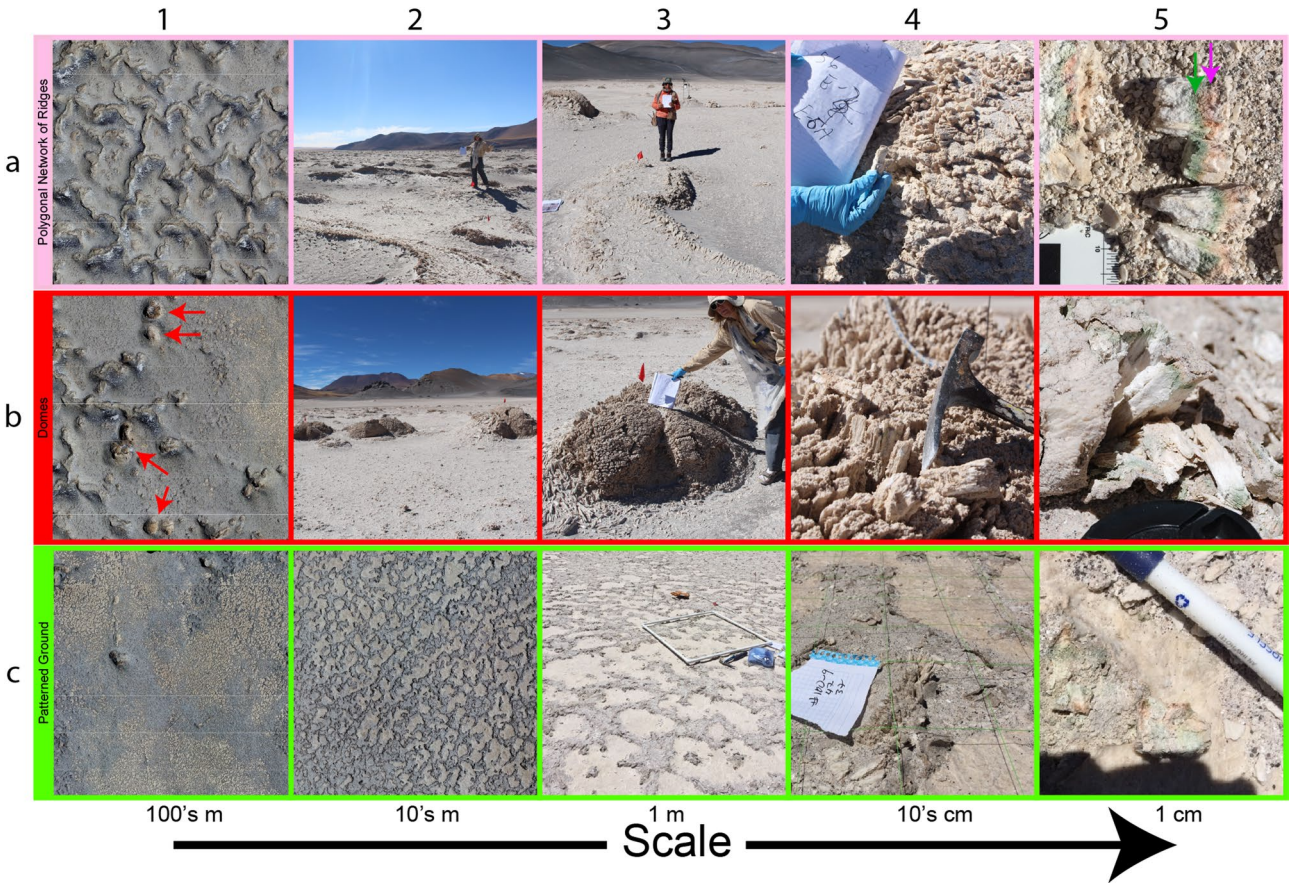
Peer review information *Nature Astronomy* thanks Jesus Martinez-Frias and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

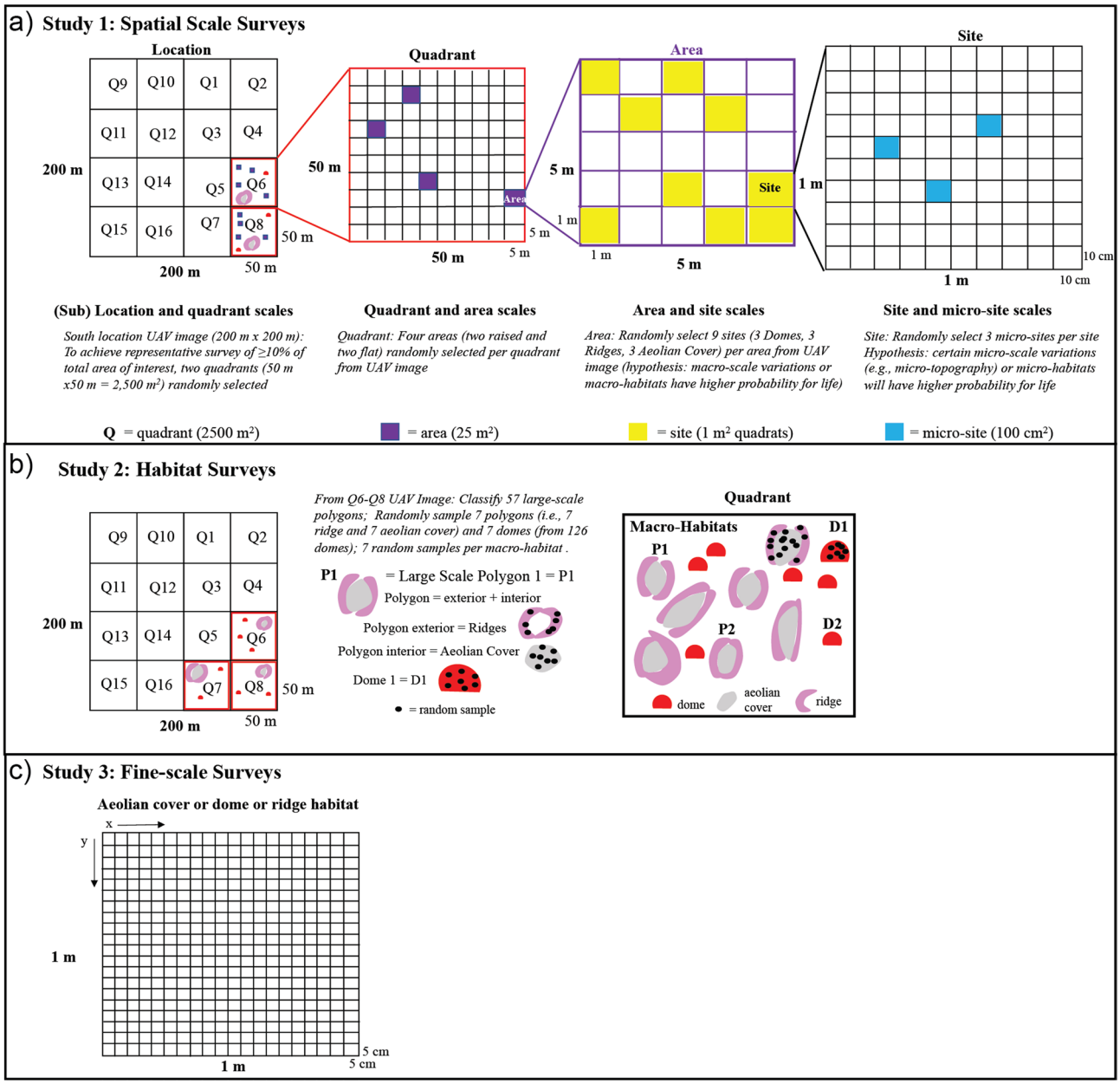
[1]SETI Institute Carl Sagan Center, Mountain View, CA, USA. [2]NASA Ames Research Center, Space Science Division and Astrobiology Division, Moffett Field, CA, USA. [3]Johns Hopkins, Applied Physics Laboratory, Laurel, MD, USA. [4]Universidad Católica del Norte, Antofagasta, Chile. [5]Department of Computer Science, University of Oxford, Oxford, UK. [6]Auckland University of Technology, Auckland, New Zealand. [7]Department of Geosciences, University of Montana, Missoula, MT, USA. [8]MarAlliance, San Francisco, CA, USA. [9]University of Washington, Seattle, WA, USA. [10]Tropical Futures Institute, James Cook University, Singapore, Singapore. [11]Old Dominion University, Norfolk, VA, USA. [12]Centro de Astrobiologia (CAB), INTA-CSIC, Madrid, Spain. [13]Impossible Sensing, St. Louis, MO, USA. [14]Carnegie Mellon University Robotics Institute, Pittsburgh, PA, USA. [15]Honeybee Robotics, Altadena, CA, USA. *A full list of authors and their affiliations appears in the Supplementary Information.
✉e-mail: krhodes@seti.org

https://doi.org/10.1038/s41550-022-01882-x



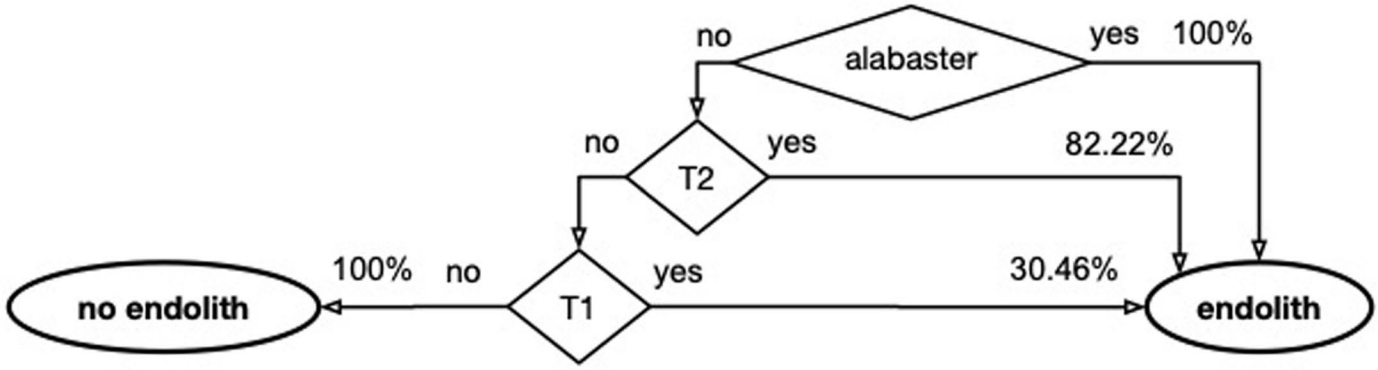
Extended Data Fig. 1 | Geomorphic units and habitats at Dome Field study location from orbital (left) to cm (right) views. a, Large-scale polygonal network pattern (pink boxes). **a1**, Orbital view reveals large-sized polygons (mean diameters in study location = 10 m). **a2**, Large-scale polygon exteriors are composed of ridges and polygon interiors are composed of aeolian cover; **a3**, Ground view of ridge (mean height in study location = 33.5 ± 8.6 cm). Note the flat ridge tops covered in eroded float due to wind erosion and scouring by sand. The sides of ridges are typically composed of crystals and the bottoms by alabaster. **a4**, Arrangement of crystals along one ridge. View of Type II crystals with visible biosignature bands, as seen in **a5** (~0.55 cm below 'eggshell seal'[23] surface duricrust); **a5**, In the close-up view of crystals, microbal laminae become visible. Pink-orange layer: desiccation and radiation-resistance bacteria such as *Salinibacter spp.*, *Halorubrum sp.*, *Pyrococcus sp.*, *Chloroflexota, Thermi* (carotenoids, pink arrow); Green colonized layer: *Cyanobacteria* (chlorophyll pigments, green arrow); **b**, domes (red boxes). **b1**, Distribution of domes (red arrows) in orbital view; **b2**, Dome heights range from ~10 cm to ~3 m, diameters range from ~1 to 7.5 m. **b3**, Ground view of domes, with eroded crystals in left foreground and intact tightly embedded crystals of the dome structure; alabaster lines the bottom of the dome; **b4**, Type I jagged sharp crystals tightly embedded in dome; **b5**, Type II crystal topped with alabaster in a dome. Note brown thin (~0.5 cm) surface duricrust and white powdery alabster efflorescence material with green and pink biosignature layers. **c**, Patterned ground (green boxes); Orbital view of patterned ground geomorphic/macro-habitat unit; **c1**. Bare salar surface is visible as speckled tan features. **c2-c3**, Fractal nature of micro-structure network inhabited by biological soil crust (BSC) communities is apparent from aerial and ground views. Micro-structure network with BSC is the patterned darker material (1–3 cm height) covering the lighter-toned bare salar surface micro-habitat. **c4**, Micro-structure with pinnacles covering the bare open salar surface, which is visible in the bottom right of photo; **c5**, A flipped micro-structure reveals BSC photosynthetic communities with orange/pink and green biosignatures.

https://doi.org/10.1038/s41550-022-01882-x



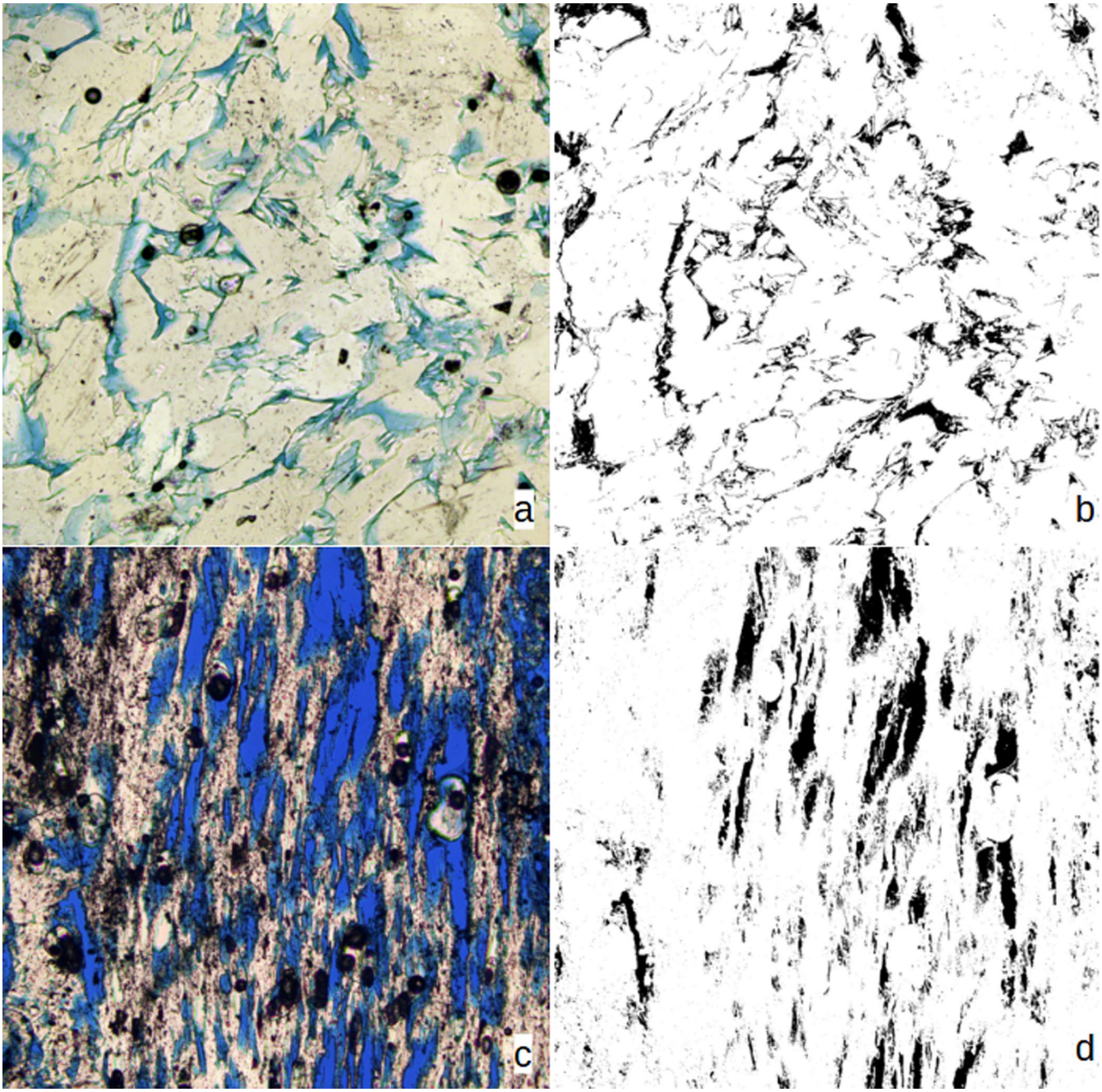
Extended Data Fig. 2 | Microbial Landscape Ecology Methods: Nested Spatial Scale and Habitat Sampling Designs. a) ES-1 nested scale sampling design. **b**) ES-2 habitat study surveys and sampling design. **c**) ES-3 fine-scale microhabitat and biosignature mapping sampling design.

https://doi.org/10.1038/s41550-022-01882-x



Extended Data Fig. 3 | Rules for the probability a sample contains endolithic biosignatures in microhabitats of a dome macrohabitat. Visual representation of the main results of Supplementary Table 1. For example, in the first decision choice, if the microhabitat is alabaster, then the probability of endolithic biosignatures is 100% (%Col is: 27 colonized samples/27 total alabaster samples ×100). In the figure, if the microhabitat is not alabaster, this leads to the next decision choice of whether the microhabitat is a Type II (T2) crystal or not, and so forth.

https://doi.org/10.1038/s41550-022-01882-x



Extended Data Fig. 4 | Thin section photomicrographs (a,c) and binary images of pore networks (b,d) of Alabaster (a,b) and a porous selenite crystal (c,d), respectively. a, b, Alabaster contains a well connected intercrystalline pore network (f = ~9.4%) without preferred orientation. **c, d**, the selenite crystal has a higher total porosity (f = ~11.9%) and larger pores, but pores are parallel to each other (elongated slot pores) and are less well connected in a horizontal direction (growth direction of crystal is up in photomicrograph) Each image is 900 ×900 mm (1750 px 1750 px).

https://doi.org/10.1038/s41550-022-01882-x

Extended Data Table 1 | Relevant ecological survey (black, this study) spatial (red, this study) and comparable rover sampling scales

Unit	Scale	Examples	Example Mars Analogs
Province	Planetary scale	Atacama Desert, Altiplano	Quadrangle, e.g., Syrtis Major
Zone	100s of km	Geologic system within a province	Isidis Basin
Region (orbital)	~100 km^2 (10-100 km)	Salar de Pajonales basin	Jezero Crater
Location (aerial) (location)	~1 km^2 (1-10 km)	Dome Field, Pajonales location	Western Delta of Jezero
Landscape scale (aerial to rover scale) (quadrant, area)	10s to 100s of m to ≤1 km; ~100 m^2	Polygonal networks/large polygons (10-30 m), very large domes, BSC patterned ground units, drainage units, aeolian cover (polygon interior units); macro-habitats	Delta deposits and strata, observed at a distance with rover instruments, such as Mastcam-Z and Ingenuity.
Ground scale (site, quadrat)	~1 m to ≤10 m; ~1 m^2	Scale of individual features, i.e., ridge (part of exterior of large polygon), aeolian cover, most domes (~1 to 7.5m), BSC areas on salar surface; macro-habitats	Outcrops observed nearby with instruments such as SuperCam LIBS (up to 7 m observation distance) and Raman (up to 12 m observation distance).
Micro-scale (micro-site)	1 to 10s of cm; ~1-10 cm^2	Micro-habitats (alabaster, crystals, float, sand, micro-structures and bare salar surface)	High resolution context imaging of outcrops with instruments such as SuperCam VISIR and Mastcam-Z.
Nano-scale	≤ 1cm; ≤1 cm^2	Single crystals, single colonized layers	Individual rocks analyzed with rover instruments, such as PIXL, SHERLOC, and SuperCam RMI.

https://doi.org/10.1038/s41550-022-01882-x

Extended Data Table 2 | Detailed statistical ecology analyses and results

(a) **Probabilities of containing biosignatures (%Col) and chi-squared tests by spatial scale, macrohabitat and microhabitat (ES-1 and ES-2).**

Unit	Biosignature Probabilities (%Col) [1]	Chi-squared test of independence, p-value[2]
Spatial Scales		
Location (Q6 and Q8)	9.2% (ES1, ratio) 11.0±4.5% (ES1, 7 areas, mean±SE)	$\chi^2(1)=2.15$, p=0.143 (Q6 and Q8) (ES1)
Quadrant 6 (areas in Q6)	15.6% (ES1, ratio) 18.5±9.8 (ES1, 3 areas, mean±SE)	χ^2 (2)=6.09, p=0.048 (areas within Q6) (ES1)
Quadrant 8 (areas in Q8)	6.5% (ES1, ratio) 6.5±3.8 (ES1, 4 areas, mean±SE)	χ^2 (3)=7.79, p=0.051 (areas within Q8) (ES1)
Macrohabitats[3]		χ^2 (2)=13.851, p<0.001(ES1)
Ridges (polygon exteriors)	18.4% (ES2, ratio) 18.4±9.7% (ES2, 7 ridges, mean±SE)	
Aeolian cover (polygon interiors)	10.2% (ES2, ratio) 10.2±5.1% (ES2, 7 AC, mean±SE)	
Domes	40.8% (ES2, ratio) 40.8±8.5% (ES2, 7 domes, mean±SE)	
Microhabitats[4]		χ^2 (3)=144.81, p<0.0001 (ES1) χ^2 (2)=70.571, p<0.0001 (ES2)
Alabaster	100% (ES1, ratio) 87.5% (ES2, ratio) 86.3±3.2% (ES2, all 3 types of macrohabitats, mean±SE)	
Crystals	1.8%(ES1, ratio) 1.8%±1.8% (ES1, avg±SE) 15.7% (ES2, ratio) 14.3±9.6% (ES2, all 3 types of macrohabitats, mean±SE)	
Float/Sand[5]	4.7% (ES1, ratio) 4.7±3.3% (ES1, 7 areas, mean±SE) 0% (ES2, ratio)	

1) Biosignature probabilities are presented herein as %Col, as this is the ratio tested for significance in the χ^2 test. 2) χ^2 test of independence of ratios of colonization; 3) %Col of the patterned ground unit is 50% and percent area cover for microstructures of the unit was 40.3%. Patterned ground was excluded here from chi-square tests to include only endolithic morphotypes. 4) In Study ES-1, %Col by BSC was 0% for the bare salar surface microhabitat and 100% for the microstructures; these microhabitats were excluded from chi-square tests to include only endolithic morphotypes; 5) In Study ES-1, float (2.4±1.5%Col) was analyzed separately from sand (12.5±7.2%Col), as well in a combined (when both are present) category (4.7±3.3%Col), as in Study ES-2.

https://doi.org/10.1038/s41550-022-01882-x

Extended Data Table 3 | Microclimate Data at Salar de Pajonales, liquid water availability in alabaster (519 sensor) and Type I crystals (517 sensor, east; 520 sensor, north) in dry and hydrated states; C: colonized; UC: uncolonized; T: temperature (oC); %RH: Percent Relative Humidity

Sensors/ aspect, microhabitat, colonization	Mean annual T	Mean annual %RH	Mean daily min %RH	Mean daily max %RH	Mean daily min T	Mean daily max T	Pre-rain T "dry" state	Pre-rain %RH "dry state"	Yearly total # hrs liquid water	17 mm rain # hrs liquid water
519, south, C, Alabaster	4.2 ± 6.1	38.4 ± 24.5	28.1 ± 22.1	46.8 ± 26.0	-11.3 ± -6.3	23.4 ± 8.1	10.2± 3.1	19.5 ± 7.3	560	474
517, east, UC, Type-I Crystal	6.0 ± 5.8	24.4 ± 17.7	9.2 ± 13.4	40.4 ± 22.1	-11.9 ± 6.3	29.1 ± 8.3	11.6 ±3.2	20.8 ± 13.8	275	184
520, north, UC, Type-I Crystal	4.6 ± 5.2	25.4 ± 19.6	12.9 ± 16.5	37.0 ± 22.7	-11.2 ± 6.2	24.5 ± 6.6	9.3 ±3.2	21.8 ± 12.0	288	210
ANOVA (p-value)	<0.001	<0.001	<0.001	<0.001	ns	<0.001	<0.001	<0.001	--	--

Mean annual and mean daily data are from 11/3/18-10/20/19; Pre-rain T and %RH are calculated from 11/1/18-01/27/19 to reflect the summer hot, dry months prior to rain onset and are indicative of microhabitat 'dry state'; total # hrs RH ≥ 95% indicates the number of hours of liquid water present in the microhabitat for the entire year and for the 17-mm rain (1/28/19-2/20/19) event, respectively. Note that Type I crystal microhabitats at different orientations and light regimes (east, 517 sensor and north, 520 sensor) on the same dome had similar liquid water availability. *ns: no significance. Relative humidity and temperature records were summarized to identify the high and low temperatures and relative humidity of each day in RStudio (3.6.3 version) and analysed with respect to three metrics: (1) mean annual, (2) mean daily minimum and (3) mean daily maximum. To test whether the mean annual, daily maximum and daily minimum temperature and relative humidity differ for each sensor, a two-way ANOVA was performed by considering microhabitats (519, 517 and 520 sensors) and time (day) as fixed factors (relative humidity data was previously log-transformed to achieve normality assumption). Pre-rain T and %RH "dry state" data were analysed as previously indicated. Tukey multiple comparison tests were carried out to determine significant differences between microhabitats.

https://doi.org/10.1038/s41550-022-01882-x

Extended Data Table 4 | Comparison of evaluation metrics of neural network semantic segmentation results on orthophoto mosaics at 6.9 cm/pixel and 23.9 cm/pixel ground sampling distances (GSD)

	6.9 cm/pixel GSD				23.9 cm/pixel GSD			
	BF-Score	Precision	Recall	Confidence	BF-Score	Precision	Recall	Confidence
PolygonRidge	91.0%	92.3%	88.8%	91.3%	35.6%	37.1%	32.9%	79.8%
Dome	80.7%	79.7%	80.9%	79.4%	26.0%	22.6%	28.6%	54.8%
PatternedGround	57.1%	54.6%	56.7%	78.8%	28.4%	30.0%	23.6%	64.3%